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 from the first submission is used.

Reviewer#1 : Wolfgang Wagner

General comments
This is an excellent paper, so I keep my review short.
This study is one of the first evaluations of ECMWF's latest reanalysis ERA-5. ERA-5 brings a number of important improvements over the widely used ERA-Interim reanalysis, including more detailed process descriptions and finer spatial- and temporal resolution. The authors evaluate ERA-5 by forcing a land surface model with ERA-Interim and ERA-5 data. To single out the effects of the enhanced precipitation estimates in ERA-5, the authors add a third forcing data set composed of ERA-5 data with only precipitation coming from ERA-Interim. The authors comprehensively evaluate the model simulations over North America with an impressive number of reference data, from in situ networks (runoff, soil moisture, evaporation, snow depth..) and remote sensing (soil moisture, LAI, ...). The results are realistic, pointing to consistent improvements of ERA-5 over ERA-Intermin in particular for the hydrologic components.

Dear Wolfgang Wagner, many thanks for reviewing the manuscript and for highlighting its relevance and interest. Your comments and suggestion led to an improved version of the manuscript, in particular with respect to section 4.

Specific comments
1.1 [Lines 254: Please explain, why discarding stations with drainage areas differing by more than 20 % from the simulated one makes sense.]

Author's response to 1.1:
The following explanation has been added to the revised version of the manuscript:

P.8, L.254, "This criterion aims to ensure a meaningful comparison between observed and simulated values. It is necessary for coping with the significant distortions in the model representation of the river network that are caused by the coarse spatial resolution of the CTRIP global river network (0.5°x0.5°)."

1.2 [Section 4: This section is a bit weak. Consider e.g. to summarize key findings with respect to each of the analysed processes and data sets. I also think that too much of the present discussion deals with other on-going work of the authors. Please try to discuss relevant links with similar undertakings currently on-going in the US, Japan,China, etc..]

Author's response to 1.2
In agreement with comment 1.2, the revised version of the manuscript have the following modifications:

P.16, L.350-352: "Albergel et al., 2017, 2018 (in prep.) recently presented a Land Data Assimilation System (LDAS-Monde) able to sequentially assimilate satellite derived estimates of surface soil moisture and LAI."
is now:
"ERA-5 has a great potential to further improve the representation of land surface variables if used to force offline LDAS. In the past recent years, several LDAS have emerged at different spatial scales, (i) regional like the Coupled Land Vegetation LDAS (CLVLDAS, Sawada and Koike, 2014,

Sawada et al., 2015), the Famine Early Warning Systems Network (FEWSNET) LDAS (FLDAS, McNally et al., 2017), (ii) continental like the North American LDAS (NLDAS, Mitchell et al., 2004; Xia etal., 2012), the National Climate Assessment LDAS (NCA-LDAS Kumar et al., 2018) as well as at (iii) global scale like the Global Land Data assimilation (GLDAS, Rodell et al., 2004) and more recently LDAS-Monde (Albergel et al., 2017, 2018 in prep). LDAS-Monde is a global capacity system able to sequentially assimilate satellite derived estimates of surface soil moisture and LAI."

New references:

-Albergel, C., S. Munier, A. Bocher, C. Draper, D. J. Leroux, A. L. Barbu, J.-C. Calvet: LDAS-Monde global capacity integration of satellite derived observations applied over North America: assessment, limitations and perspectives. to be sumitted to Remote Sensing, Special Issue "Assimilation of Remote Sensing Data into Earth System Models", 2018

-Kumar, S.V., M. Jasinski, D. Mocko, M. Rodell, J. Borak, B. Li, H. Kato Beaudoing, and C.D. Peters-Lidard: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment. J. Hydrometeor., 0, https://doi.org/10.1175/JHM-D-17-0125.1

-McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., Funk, C., Peters-Lidard, C. D. and Verdin, J. P.: A land data assimilation system for sub-Saharan Africa food and water security applications. Scientific Data, 4, 170012, :10.1038/sdata.2017.12, 2017.

-Mitchell, K. E., et al. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, J. Geophys. Res.,109, D07S90, 2004. doi:10.1029/2003JD003823

-Muñoz-Sabater, Joaquín, Emanuel Dutra, Gianpaolo Balsamo, Souhail Boussetta, Ervin Zsoter, Clement Albergel, Anna Agusti-Panareda: ERA5-Land: an improved version of the ERA5 reanalysis land component. Joint ISWG and LSA-SAF Workshop, 26-28 June 2018, Lisbon, Portugal.

-Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, The Global Land Data Assimilation System, Bull. Amer. Meteor. Soc., 85(3), 381–394, 2004.

-Sawada, Y., T. Koike, and J. P. Walker, A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics, J. Geophys. Res.Atmos., 120, doi: 10.1002/2014JD022895, 2015.

-Sawada, Y., and T. Koike, Simultaneous estimation of both hydrological and ecological parameters in an ecohydrological model by assimilating microwave signal, J. Geophys. Res. Atmos., 119, doi:10.1002/2014JD021536, 2014.

Xia, Y., et al. 2012, Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, J. Geophys. Res., 117, D03109, doi:*10.1029/2011JD016048*, 2012.

1.3 [Albergel et al. (2018) not in the references]

Author's response to 1.3: it is now corrected in the revised version of the manuscript.

1.4 [Figure 4: "For sake of clarity"]

Author's response to 1.4: it is now corrected in the revised version of the manuscript.

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

Reviewer#2

This paper documents the improvements in land surface simulations driven by ERA-5 relative to ERA-Interim. The work is relevant from the standpoint of documenting the improvements in ERA-5 as it relates to hydrology. The manuscript is somewhat sloppily prepared (a figure was missing!). My comments and suggestions are below.

We thanks anonymous Reviewer 2 for his/her review of the manuscript and for highlighting the relevance of the study for documenting the improvements in ERA-5 as it relates to hydrology. It is the main objective of the study.

We are sorry that a figure (figure 5 according to Reviewer#2 first major comment) was missing in the version of the manuscript he/she had. From the pdf file available on HESSD website, this figure is available, at least when downloaded from all co-author's institutes from the date Reviewer#2's comments were posted on-line (03/05/2018).

Reviewer#2 has made several fruitful comments/corrections/suggestions that led to an improved version of the manuscript. Again we would like to thanks Reviewer#2 for his/her work.

Responses to the Reviewer are available in the supplement.

Major comments:

2.1. [The text has numerous language issues and grammar mistakes, some of which are listed below. I ran out of steam documenting all of them. I assume that the author would take a fresh and careful look at the manuscript to correct them all (including the ones that are not listed). Note that Figure 5 was missing in the manuscript version that I reviewed.]

Author's response to 2.1
Many thanks for correcting some language issues and grammar mistakes, all the points listed by Reviewer#2 were corrected and a fresh and careful look at the revised version of the manuscript has been taken. We are sorry that figure 5 was missing in the version of the manuscript Reviewer#2 had. From the pdf file available on HESSD website, this figure is available, at least when downloaded from all co-author's institutes from the date Reviewer#2's comments were posted on-line (03/05/2018).

2.2. [The paper is really an offline LDAS simulation, but makes no mention of other LDAS work. The literature review should encompass the recent work in this regard that have considered the assimilation of land measurements (NCA-LDAS, for.e.g). I think the paper would be more powerful and of broader interest if the authors can document how the ERA-5 forced system compares with the such LDAS efforts. For example, does ERA-5 have comparable skills to NLDAS2, the defacto standard land product over CONUS? How does it compare to MERRA2 and ERA-Interim-Land? Without such comparisons, the paper sounds more as a technical report the impact of ERA related changes.]

Author's response to 2.2
This study only considers offline simulations and is not "[...] really an offline LDAS simulation [...]"; no assimilation of any land measurements is done in this work. The fact that only open loop

(offline) simulations are considered is what permits to express an evaluation on the quality of the used forcing being ERA-5 and ERA-Interim, while in case there would be an LDAS attached the causal attribution would become more complex. Furthermore, LDAS's notions appear only in the discussions and conclusions section L.510 as a next working step.

Although we agree that comparisons with other products are of interest, it goes beyond the scope of this study that is documenting the improvements in ERA-5 with respect to ERA-Interim as it relates to hydrology. It is likely that future work from the same group of authors will consider such comparisons especially with the shortcoming land version of ERA-5; ERA5-land (Sabater et al., 2018, workshop paper). We understand however the importance of mentioning other LDAS work and the following sentence has been modified:

P.16, L.350-352: "Albergel et al., 2017, 2018 (in prep.) recently presented a Land Data Assimilation System (LDAS-Monde) able to sequentially assimilate satellite derived estimates of surface soil moisture and LAI."
is now:
"ERA-5 has a great potential to further improve the representation of land surface variables if used to force offline LDAS. In the past recent years, several LDAS have emerged at different spatial scales, (i) regional like the Coupled Land Vegetation LDAS (CLVLDAS, Sawada and Koike, 2014, Sawada et al., 2015), the Famine Early Warning Systems Network (FEWSNET) LDAS (FLDAS, McNally et al., 2017), (ii) continental like the North American LDAS (NLDAS, Mitchell et al., 2004; Xia etal., 2012), the National Climate Assessment LDAS (NCA-LDAS Kumar et al., 2018) as well as at (iii) global scale like the Global Land Data assimilation (GLDAS, Rodell et al., 2004) and more recently LDAS-Monde (Albergel et al., 2017, 2018 in prep). LDAS-Monde is a global capacity system able to sequentially assimilate satellite derived estimates of surface soil moisture and LAI."

New references:
-Albergel, C., S. Munier, A. Bocher, C. Draper, D. J. Leroux, A. L. Barbu, J.-C. Calvet: LDAS-Monde global capacity integration of satellite derived observations applied over North America: assessment, limitations and perspectives. to be sumitted to Remote Sensing, Special Issue "Assimilation of Remote Sensing Data into Earth System Models", 2018

-Kumar, S.V., M. Jasinski, D. Mocko, M. Rodell, J. Borak, B. Li, H. Kato Beaudoing, and C.D. Peters-Lidard: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment. J. Hydrometeor., 0, https://doi.org/10.1175/JHM-D-17-0125.1

-McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., Funk, C., Peters-Lidard, C. D. and Verdin, J. P.: A land data assimilation system for sub-Saharan Africa food and water security applications. Scientific Data, 4, 170012, :10.1038/sdata.2017.12, 2017.

-Mitchell, K. E., et al. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, J. Geophys. Res.,109, D07S90, 2004. doi:10.1029/2003JD003823

-Muñoz-Sabater, Joaquín, Emanuel Dutra, Gianpaolo Balsamo, Souhail Boussetta, Ervin Zsoter, Clement Albergel, Anna Agusti-Panareda: ERA5-Land: an improved version of the ERA5 reanalysis land component. Joint ISWG and LSA-SAF Workshop, 26-28 June 2018, Lisbon, Portugal.

-Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, The Global Land Data Assimilation System, Bull. Amer. Meteor. Soc., 85(3), 381–394, 2004.

-Sawada, Y., T. Koike, and J. P. Walker, A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics, J. Geophys. Res.Atmos., 120, doi: 10.1002/2014JD022895, 2015.

-Sawada, Y., and T. Koike, Simultaneous estimation of both hydrological and ecological parameters in an ecohydrological model by assimilating microwave signal, J. Geophys. Res. Atmos., 119, doi:10.1002/2014JD021536, 2014.

-Xia, Y., et al. 2012, Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, J. Geophys. Res., 117, D03109, doi:*10.1029/2011JD016048*, 2012.

2.3. [I found the use of metrics to be a bit convoluted and inconsistent. For example, why use NIC for metrics such as R, instead of simply taking a difference? NIC is more useful when the dynamic range of the metrics are really large (NSE ranges from -infinity to 1, so a large negative value would blow up a domain average). I suggest sticking with simple differences so that the impact on the model runs are more intuitive. Why use ubRMSED to look at impacts on fluxes? (It is used for soil moisture because of the large climatological differences)Also, I would change the sign of N_MAE and N_ubRMSD to be the same as that of NIC values (Positive value indicating improvements).]

Author's response to 2.3

Agreed, more consistency can now be found across the metrics and several changes have been made in the revised versions of the manuscript including:
- NIC is applied to NSE, only, and the impact on R is assessed using R differences,
- RMSD (instead of ubRMSD) is now used to assess the impact on fluxes (it is kept for soil moisture, for the reason explained by Reviewer#2, and snow evaluation as we believe it is still very informative since it shows the improvements on the "random" component of the error).
- Changing the sign of Nmae and NubRMSE is a bit contra-intuitive, instead we have decided to remove those metrics. Snow impact is now assesed using bias, ubRMSD and R.
- For each metrics, the 95% confidence interval of the median derived from a 10000 samples bootstrapping is provided.

It led to several changes in section 3.1 detailed below (text, tables and figures). A new Table (Table III) has been added to present scores for the snow evaluation, figure 6 (for snow evaluation) has been modified, figure 7 (panels c and d) now shows RMSD (instead of ubRMSD) and figure 11 (now figure 10) shows score differences (instead of N_RMSD and NIC_R). Please see also Author's response to 2.24.

The whole new section 3.1 is now (modification are highlighted in yellow):
"This section presents the results of the comparison versus in situ observations of land surface variables from model simulations using either ei_S, e5ei_S or e5_S starting with soil moisture. The statistical scores for 2010–16 surface soil moisture from ei_S, e5ei_S and e5_S are presented in Table II. Median R values on volumetric time-series (anomaly time series) along with their 95% confidence intervals are 0.66±0.02 (0.53±0.02), 0.69±0.02 (0.54±0.04) and 0.71±0.02 (0.58±0.03) while median ubRMSD are 0.052±0.003, 0.052±0.002 and 0.050±0.003 for ei_S, e5ei_S and and e5_S, respectively. These results underline the better capability of the ISBA LSM to represent surface soil moisture variability when forced by ERA-5 reanalysis. Also the latest configuration (e5_S) presents more stations with better R values on volumetric time-series (anomaly time series) than both ei_S and e5ei; respectively 60% and 75% (out of 110 and 107 stations, respectively). This is also reflected on figure 2 illustrating correlations values on volumetric time-series (fig.2a) and anomaly time-series (fig.2b) on maps. Stars symbols represent stations for which ISBA LSM performs best when forced ERA-Interim, circles when it is forced by ERA-5 with ERA-Interim precipitations and downward pointing triangles when it is forced by all ERA-5 atmospheric variables. Both maps on figure 2 are dominated by downward pointing triangles. Fig.2c(d) shows

histograms of R differences on volumetric (anomaly) time-series, for soil moisture from e5_S (in red) e5ei_S (in green) with respect to ei_S, median values of the differences are reported, also.

172 out of 344 gauging stations retained for the evaluation according to the criteria described in the methodology section presents NSE scores in the [-1, 1] interval. Figure 3 represents performance of each dataset for this pool of stations. Fig3.a is a scatterplot of NSE scores between in situ and simulated river discharges *Q*; NSE scores for *Q* simulated with either ERA-5 but ERA-Interim precipitations (e5ei_S, green crosses) or ERA-5 (e5_S, red dots) function of NSE scores for *Q* simulated using ERA-Interim (ei_S). When considering e5_S, almost all the red dots are above the 1:1 diagonal suggesting a general improvement from the use of e5_S. For a large part, e5ei_S green crosses are above this diagonal, suggesting that the improvement in e5_S does not only comes from precipitation but from other variables, also. Median NSE values are 0.06±0.06, 0.12±0.07 and 0.24±0.05 for ei_S, e5ei_S and e5_S, respectively. Fig.3b shows an histogram of river discharges ratio for ei_S (Qr_ei in blues), e5ei_S (Qr_e5ei in green) and e5_S (Qr_e5 in red), median values are 0.67, 075 and 0.77, respectively. While all three experiments underestimate *Q* (a value of 1 being a perfect match), the use of e5ei_S and e5_S leads to better results. Finally, figure 3c illustrates hydrographs for a river station in Lousiana (33.08°N, -93.85°W) representing scaled *Q* (using either observed or simulated drainage areas), in situ data (black crosses), simulated river discharges from ei_S (blue solid line), e5ei_S (green solid line) and e5_S (red solid line). From this hydrograph, the added value of e5_S is clear, particularly for the 2011 and 2015 main events. NSE scores are 0.47, 0.61 and 0.76 for ei_S, e5ei_S and e5_S, respectively. Figure 4 illustrates the added value of using e5_S (a) or e5ei_S (b) with respect to ei_S. For 156 out of the pool of 172 stations NIC_{NSE} values computed using e5_S with respect to ei_S are positive (large blue circles) showing an general improvement from the use of e5_S (representing 91% of the stations) with a median NIC_{NSE} value of 14%±0.05. When considering e5ei_S versus ei_S, they are still 118 (69%) with a median NIC_{NSE} value of 4%±0.02 suggesting that the improvement in e5_S does not only comes from precipitation but from other variables, also. It is also worth-noticing that stations where a score degradation is observed (large red circles) are located in areas known for irrigation which is not represented in ISBA. All scores computed for seasons (December-January-Februray, March-April-May, Jun-July-August, September-October-November) suggest the same ranking (not shown).

The mean snow depth bias of ei_S (see Figure 5) highlights a clear underestimation of winter snow depth accumulation mainly over the Rocky Mountains. This is likely a result of the underestimation of snowfall by ei_S associated with an overestimation of snow melt due to the coarse resolution of the ei_S reflected in a smooth topography. The replacement of all forcing variables by e5_S but keeping ei_S precipitation (e5ei_S, Fig.5b) shows a slight increase in snow depth. This result justifies the above hypothesis that part of the snow underestimation is also due to temperature issues linked with a coarse model orography. Moving to the full e5_S forcing there is a clear increase of snow depth, when compared with both ei_S and e5ei_S forced simulations resulting from an increase in snowfall in e5_S. Figure 6 presents the mean seasonal cycle of bias and ubRMSD (fig.6a) and correlations (fig.6b) over 2010-2016. In addition to the added values of e5_S in terms of the mean snow depth already presented in figure 5, the temporal variability and random errors are also improved. Comparably with what was discussed for the mean bias, e5ei_S shows some benefits when compared with ei_S in terms of ubRMSD and correlation (median bias, ubRMSD and R values of e5_ei over the whole period are; -1.70±0.33 cm., 7.40±0.65 cm. and 0.77± 0.01, respectively, for ei_S they are; -2.11±0.33 cm., 7.58±0.65 cm. and 0.75± 0.01, respectively) while e5_S has a clear improvement in ubRMSD and correlation (median bias, ubRMSD and R values of e5_ei over the whole period are; -0.64±0.19 cm., 7.00±0.65 cm. and 0.82± 0.01, respectively). The improvements on the snow depth simulations are consistent throughout the entire snow covered season (see Fig.6a and b) with a maximum improvement from January to March. These results highlight the cumulative effect of the forcing quality on the snow depth simulation. Finally Table III presents scores from the comparison of snow depth with in situ measurements, median Bias, ubRMSD and R values are given for the three seasons affected by snow (September-October-November, December-January-February and Mars-April-May) and for the whole period. e5_S always presents better scores when compared to ei_S and it is always the configuration presenting the highest percentage of stations with the best scores. Looking at the 95% confidence interval, for the correlation and bias it is clear that the changes are significant.

Results from the comparisons between ei_S, e5ei_S, e5_S and in situ sensible and latent flux measurements are presented in table IV and illustrated by figure 7 and 8. 37 stations present significant correlation values (at p-value < 0.05). For sensible heat flux, median correlation and RMSD values are 0.62±0.11, 0.62±0.11 and 0.65±0.11, 39.58±3.71 $W.m^{-2}$, 32.89±3.86 $W.m^{-2}$ and 32.73±2.61 $W.m^{-2}$ for ei_S, e5ei_S and e5_S, respectively. For latent heat flux, they are 0.63±0.05, 0.62±0.07 and 0.70±0.04, 39.00±5.38 $W.m^{-2}$, 37.12±4.37 $W.m^{-2}$ and 36.66±4.94 $W.m^{-2}$. As for surface soil moisture, river discharge and snow depth, e5_S presents better results than e5ei_S and ei_S. At the station level, figure 7 illustrates scatter plots of correlations and RMSD for sensible and latent heat flux from ei_S, e5ei_S, e5_S against in situ measurements of sensible (fig.7a for correlation, fig.7c for RMSD) and latent (fig.7b for correlation, fig.7d for RMSD) heat flux. Scores for either e5ei_S (green dots) or e5_S (in red) are presented function of those for ei_S. When looking at the correlations, almost all of e5_S and e5ei_S symbols (in red and green, respectively on fig.7a, fig.7c) are above the 1:1 diagonal indicating that e5_S and e5ei_S better represent sensible and latent heat flux than ei_S. Same tendency is observed for RMSD with most of the symbol below the 1:1 diagonal. If RMSD values are comparable for e5_S and e5ei_S, R values are clearly higher for e5_S. "

Table I : Evaluation datasets and associated metrics used in this study.

Datasets used for the evaluation	Source	Metrics associated
In situ measurements of soil moisture (USCRN, Bell et al., 2013)	https://www.ncdc.noaa.gov/crn	R (on both volumetric and anomaly time-series) ubRMSD
In situ measurements of streamflow (USGS)	https://nwis.waterdata.usgs.gov/nwis	Nash Efficiency (NSE), Normalized Information Contribution (NIC) based on NSE, Ratio of simulated and observed streamflow (Q)
In situ measurements of snow depth (GHCN, Menne et al., 2012a, b)	https://www.ncdc.noaa.gov/climate-monitoring/	R, bias and ubRMSD
In situ measurements of sensible and latent heat fluxes (FLUXNET-2015)	http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/	R, RMSD
Satellite derived surface soil moisture (ESA CCI SSM v4, Dorigo et a., 2015, 2017)	http://www.esa-soilmoisture-cci.org	R (on both volumetric and anomaly time-series)
Satellite derived Leaf Area Index (GEOV1, Baret et al., 2013)	http://land.copernicus.eu/global/	R and RMSD
Satellite-driven model estimates of land evapotranspiration (GLEAM, Martns et al., 2017)	http://www.gleam.eu	R and RMSD
Upscaled estimates of Gross Primary Production (GPP, Jung et al., 2017)	https://www.bgc-jenna.mpg.de/geodb/projects/Home.php	R and RMSD

Table II: Comparison of surface soil moisture with in situ observations for ei_S, e5ei_S and e5_S over 2010-2016 (April to September months are considered). Median correlations R (on volumetric and anomaly time series) and ubRMSD are given for the USCRN. Scores are given for significant correlations with p-values <0.05.

	Median R* on volumetric time series, 95 % Confidence Interval** (% of stations for which this configuration is the best)	Median R*** on anomalies time series, 95 % Confidence Interval** (% of stations for which this configuration is the best)	Median ubRMSD* (m^3m^{-3}), 95 % Confidence Interval** (% of stations for which this configuration is the best)
ei_S	0.66±0.02 (20 %)	0.53±0.02 (15 %)	0.052±0.003 (19 %)
e5ei_S	0.69±0.02 (20 %)	0.54±0.04 (10 %)	0.052±0.002 (24 %)
e5_S	0.71±0.02 (60 %)	0.58±0.03 (75 %)	0.050±0.003 (57 %)

*** only for stations presenting significant R values on volumetric time series (p-value<0.05): 110 stations**
**** 95% confidence interval of the median derived from a 10000 samples bootstrapping**
***** only for stations presenting significant R values on anomaly time series (p-value<0.05): 107 stations**

Table III: Comparison of snow depth with in situ measurements, median Bias, ubRMSD and R values are given for the three seasons affected by snow (SON, DJF, MAM) and for the whole period (All). SON, DJF and MAM stand for September-October-November, December-January-February and Mars-April-May, respectively.

		Median bias (cm)*, 95 % Confidence Interval** (% of stations for which this configuration is the best)	Median ubRMSD (cm)*, 95 % Confidence Interval** (% of stations for which this configuration is the best)	Median R*, 95 % Confidence Interval** (% of stations for which this configuration is the best)
ei_S	SON	-0.27±0.04 (13 %)	2.05±0.17 (13 %)	0.70±0.01 (21 %)
	DJF	-6.28±0.86 (11 %)	10.34±0.63 (17 %)	0.72± 0.01 (20 %)
	MAM	-1.90±0.33 (15 %)	7.82±0.79 (17 %)	0.65± 0.01 (18 %)
	All	-2.11±0.33 (11 %)	7.58±0.65 (14 %)	0.75± 0.01 (19 %)
e5ei_S	SON	-0.25±0.04 (12 %)	2.03±0.15 (10 %)	0.74± 0.01 (23 %)
	DJF	-4.84±0.80 (14 %)	9.98±0.50 (14 %)	0.75± 0.01 (21 %)
	MAM	-1.49±0.33 (14 %)	7.61±0.76 (13 %)	0.69±0.02 (22 %)
	All	-1.70±0.33 (14 %)	7.40±0.65 (14 %)	0.77± 0.01 (20 %)
e5_S	SON	-014±0.03(76 %)	1.83±0.14 (77 %)	0.79± 0.01 (56 %)
	DJF	-1.70±0.44 (75 %)	9.64±0.46 (69 %)	0.80± 0.01 (59 %)
	MAM	-0.57±0.22 (71 %)	7.43±0.79 (70 %)	0.76± 0.01 (60 %)
	All	-0.64±0.19 (75 %)	7.00±0.65 (72 %)	0.82± 0.01 (61 %)

*** only for stations presenting more than 80% of (daily) data; 1901 out of 2056 stations.**
**** 95% confidence interval of the median derived from a 10000 samples bootstrapping**

Table IV: Comparison of sensible (H) and latent (LE) heat flux with in situ observations for ei_S, e5ei_S and e5_S. Median correlations (R) and median RMSD are given for the fluxnet stations. Scores are given for significant correlations with p-values <0.05.

	H Median R*, 95 % Confidence Interval** (% of stations for which this configuration is the best)	H Median RMSD* W.m^{-2}, 95 % Confidence Interval** (% of stations for which this configuration is the best)	LE Median R*, 95 % Confidence Interval** (% of stations for which this configuration is the best)	LE Median RMSD* W.m^{-2}, 95 % Confidence Interval** (% of stations for which this configuration is the best)
ei_S	0.62±0.11 (8 %)	39.58±3.71 (5 %)	0.63±0.05 (8 %)	39.00±5.38 (16 %)
e5ei_S	0.62±0.11(27 %)	32.89±3.86 (27%)	0.62±0.07 (11 %)	37.12±4.37 (22 %)
e5_S	0.65±0.11 (65 %)	32.73±2.61 (68 %)	0.70±0.04 (81 %)	36.66±4.94 (62 %)

*** only for stations presenting significant R values (p-value<0.05): 37 stations**
**** 95% confidence interval of the median derived from a 10000 samples bootstrapping**

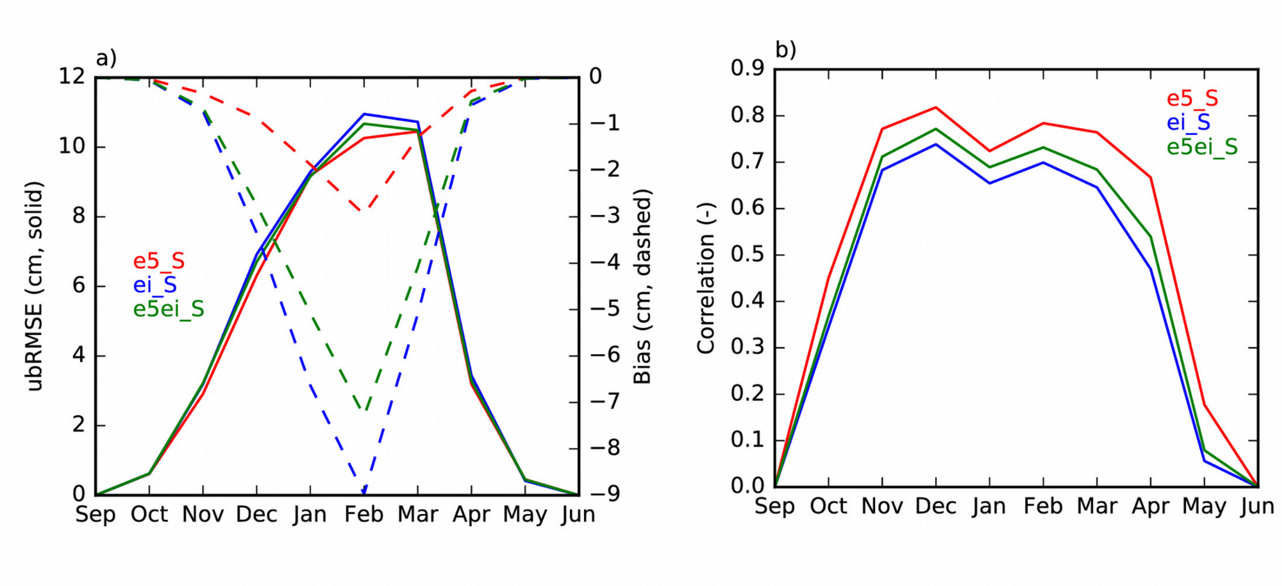

Figure 6: (a) Mean seasonal cycle of the bias (dashed lines) and ubRMSD (solid lines) averaged over all stations and (b) the mean seasonal cycle of the correlations for ei_S (in blue), e5ei_S (in green) and e5_S (in red).

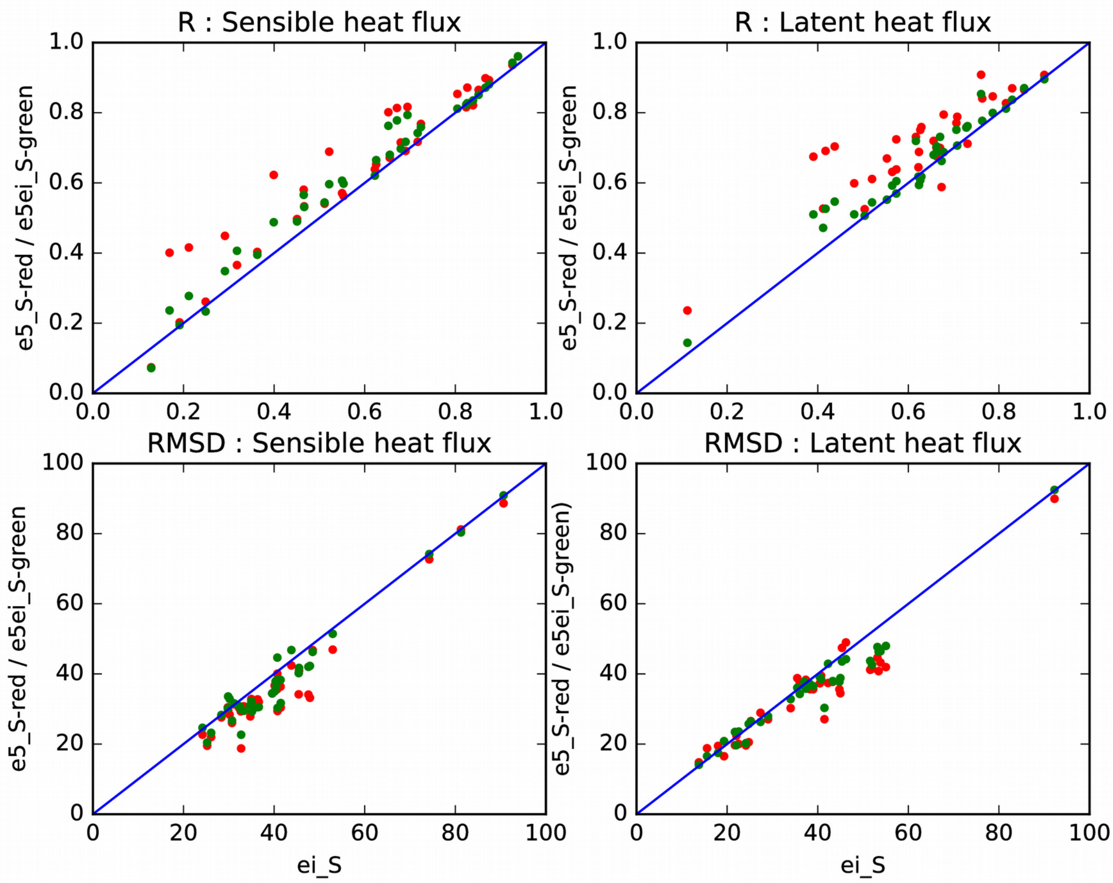

Figure 7: Scatterplots illustrating evaluation of ei_S, e5ei_S, e5_S against in situ measurements of sensible (a for correlation, c for RMSD) and latent (b for correlation, d for RMSD) heat flux. Scores for either e5ei_S (green dots) or e5_S (in red) are presented as function of those for ei_S.

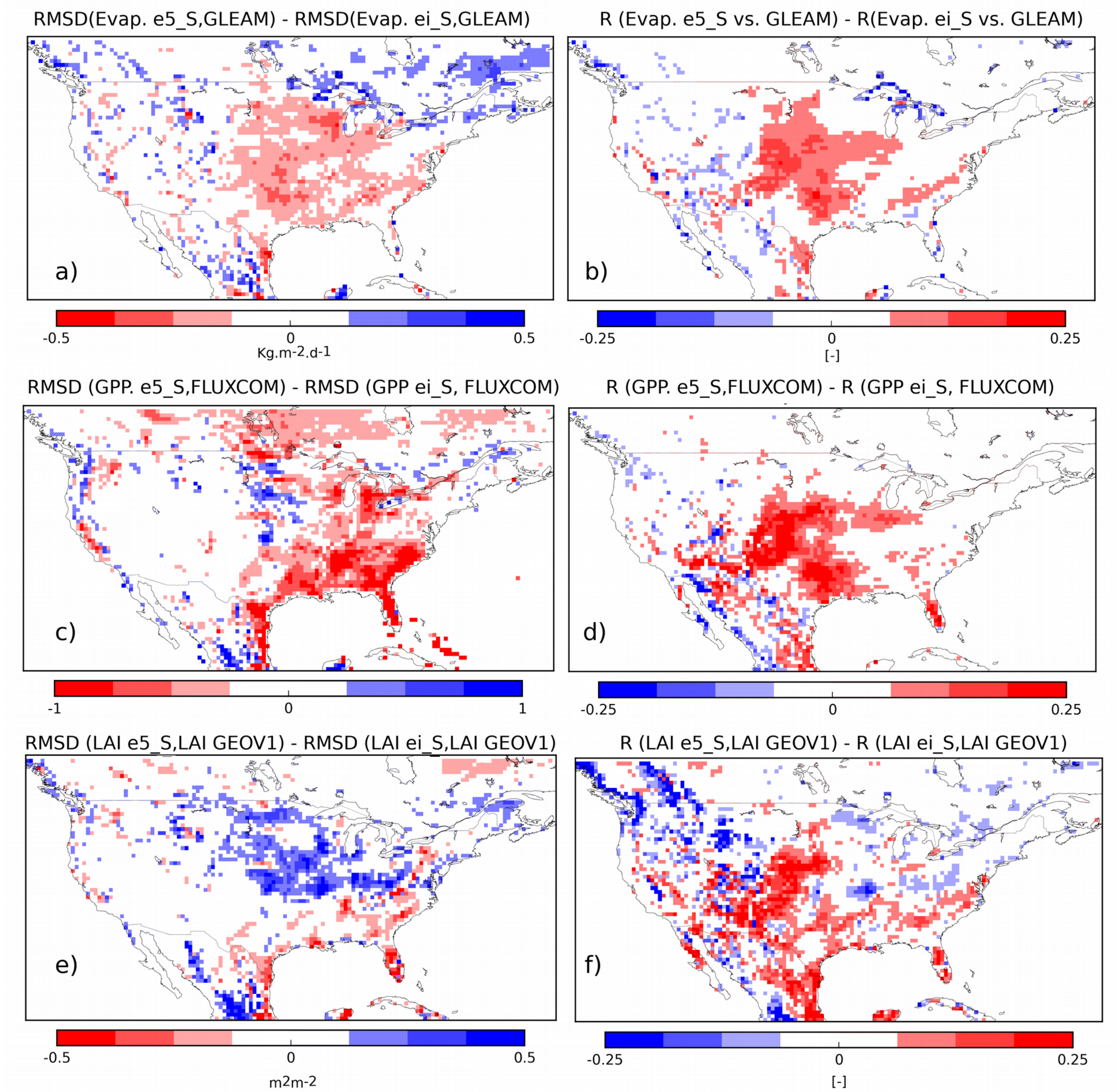

Figure 10: RMSD differences (a, c, e) and Correlation differences (b, d, f) for e5_S simulations with respect to ei_S simulations for three land surface variables: evapotranspiration, Gross Primary Production and Leaf Area Index from top to bottom. Areas in red represent an improvement from the use of ERA-5.

2.4. [Section 3.1: I think the descriptions need to tone down the language on how much improvements are actually gained. From table 1, it looks like the improvements are quite small though they are systematic with the new version. I think it is important to quantify the magnitude of improvements (showing their spatial distribution through, for e.g., histograms).]

Author’s response to 2.4

In agreement with Reviewer#2’s comment, some parts of section 3.1 are now re-worded in the revised version of the manuscript (please see also Author’s response to 2.3, 2.24). The idea that improvements, even when they are quite small are systematic is now mentioned in the conclusion, also.

We believe that several figures already quantify the magnitude of improvements (e.g. figure 3.a, 4, 5, 6 and 7). Two histograms of R differences, as suggested by Reviewer#2, where added to figure 2 (panels c & d) to show the spatial distribution of the improvement on correlation for soil moisture (for both volumetric and anomaly time series). Please see below new figure 2.

P.12, L.388: the following sentence has been added to describe the new figure: " Fig.2c(d) shows histograms of R differences on volumetric (anomaly) time-series, for soil moisture from e5_S (in red) e5ei_S (in green) with respect to ei_S, median values of the differences are reported, also."

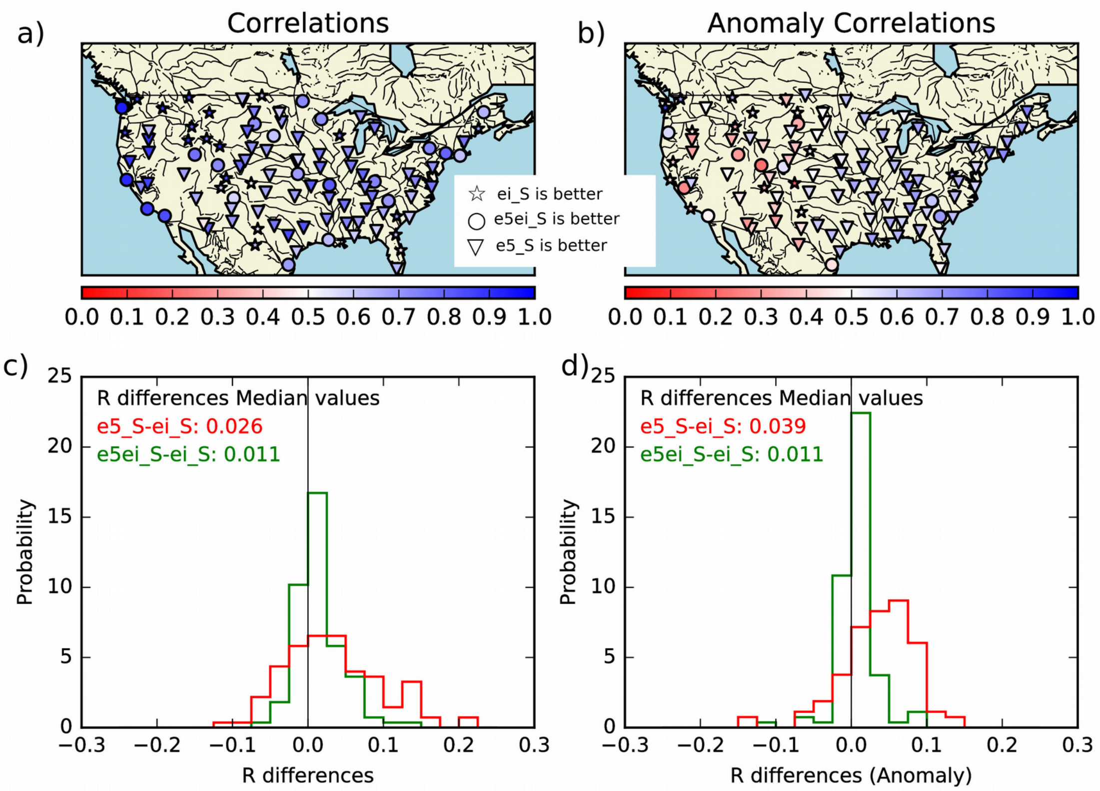

Figure 2 : Maps of correlation (R) on volumetric time-series (a) and anomaly time-series (b) between in situ measurements at 5 cm depth from the USCRN network and the ISBA Land Surface Model within the SURFEX modeling platform forced by either ERA-Interim (ei_S), ERA-5 with ERA-Interim precipitations (e5ei_S) and ERA-5 (e5_S). For each stations presenting significant R (p-values < 0.05) simulation that presents the better R values is represented. Stars symbols are when ei_S, presents the best value, circles when it e5ei_S and downward pointing triangles when it is e5_S. (c) Shows histograms of R differences on volumetric time-series, R(e5_S)-R(ei_S) in red and R(e5ei_S)-R(ei_S) in green, median values of the differences are reported, also. (d) Same as (c) for R values on anomaly time-series.

Minor comments:

2.5. [Fix the sentence starting with 'ERA-5 important changes ..' to something like ' ERA-5 has important changes relative to ERA-Interim former atmospheric reanalysis including...']

Author's response to 2.5
It has now been fixed in the revised version of the manuscript

2.6. [Change the sentence 'ERA-5 is forseen .. ' to something like 'As ERA-5 is expected to replace ERA-Interim reanalysis, this study assesses whether']

Author's response to 2.6
P.2, L.65-69 "It will eventually replace ERA-Interim reanalysis. Assessing ERA-5 ability to force a LSM with respect to ERA-Interim is therefore highly relevant. To that end, ERA-5, ERA-Interim as well as a combination of both (ERA-5 with precipitation of ERA-Interim) are used to constrain [...]"
<u>is now:</u>
"As ERA-5 will eventually replace ERA-Interim reanalysis assessing its ability to force a LSM with respect to ERA-Interim is highly relevant. In this study, ERA-5, ERA-Interim as well as a combination of both (ERA-5 with precipitation of ERA-Interim) are used to constrain [...]"

2.7. [Change the sentence 'ERA-5 impact on the ISBA ..' to 'ERA-5's impact on ISBA LSM relative to ERA-Interim is evaluated using remote sensing...']

Author's response to 2.7
P.1, L.22-24: "ERA-5 impact on the ISBA LSM with respect to ERA-Interim is assessed over a data-rich area: North America. A comprehensive evaluation of ERA-5 impact is conducted using remote sensing and in-situ observations covering a substantial part of the land surface storage and fluxes."
<u>is now :</u>
"ERA-5 impact on ISBA LSM relative to ERA-Interim is evaluated using remote sensing and in-situ observations covering a substantial part of the land surface storage and fluxes over the CONtinuous US (CONUS) domain."

2.8. [Line 34 – Fix 'Interim ..' to 'Interim.' (only one period).]

Author's response to 2.8 : done

2.9. [Line 36: change 'extend' to 'extent']

Author's response to 2.9 : done

2.10. [Line 46: Change 'essentials' to 'essential']

Author's response to 2.10 : done

2.11. [Line 52: Change 'progresses' to 'progress']

Author's response to 2.11 : done

2.12. [Line 55: Add a comma after 'decade'.]

Author's response to 2.12 : done

2.13. [Lines 58-60: MERRA is retired. More appropriate to refer to MERRA2 papers. Given that this paper focuses on land-only simulations, there should be a description of LDAS analysis forced by observed precipitation (and meteorology) such as NLDAS, GLDAS ,etc.]

Author's response to 2.12
Following changes have been made in the revised version of the manuscript:
"Amongst them are NASA's Modern Era Retrospective-analysis for Research and Applications (MERRA; Rienecker et al., 2011) as well as ECMWF's (European Centre for Medium-Range Weather Forecasts) Interim reanalysis (ERA-Interim; Dee et al., 2011). Their offline use in LSMs

led to global Land Surface Variables (LSVs) reanalysis datasets that can support e.g. water resources analysis (Schellekens et al., 2017), like MERRA-Land (Reichle, 2011) and ERA-Interim/Land (Balsamo et al., 2015)."

is now:

"Amongst them are NASA's Modern Era Retrospective-analysis for Research and Applications (MERRA; Rienecker et al., 2011 and MERRA2; Gelaro et al. 2016,) as well as ECMWF's (European Centre for Medium-Range Weather Forecasts) Interim reanalysis (ERA-Interim; Dee et al., 2011). Their offline use in either LSMs or Land Data Assimilation System (LDAS), with or without meteorological corrections (e.g., precipitations) led to global land surface variables (LSVs) reanalysis datasets that can support e.g. water resources analysis (Schellekens et al., 2017), like MERRA-Land and MERRA2-Land (Reichle, 2011; 2017), ERA-Interim/Land (Balsamo et al., 2015), the forthcoming ERA5-Land (Muñoz-Sabater et al., 2018), the North American LDAS (NLDAS, Mitchel et al., 2004), the Global LDAS (GLDAS, Rodell et al., 2004) and LDAS-Monde (Albergel et al., 2017)."

References

-Gelaro, R., W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A.M. da Silva, W. Gu, G. Kim, R. Koster, R. Lucchesi, D. Merkova, J.E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, and B. Zhao, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate,* 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1

-Mitchell, K. E., et al. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, J. Geophys. Res.,109, D07S90, 2004. doi:10.1029/2003JD003823

-Muñoz-Sabater, Joaquín, Emanuel Dutra, Gianpaolo Balsamo, Souhail Boussetta, Ervin Zsoter, Clement Albergel, Anna Agusti-Panareda: ERA5-Land: an improved version of the ERA5 reanalysis land component. Joint ISWG and LSA-SAF Workshop, 26-28 June 2018, Lisbon, Portugal.

-Reichle, R.H., C.S. Draper, Q. Liu, M. Girotto, S.P. Mahanama, R.D. Koster, and G.J. De Lannoy, 2017: Assessment of MERRA-2 Land Surface Hydrology Estimates. *J. Climate,* 30, 2937–2960, https://doi.org/10.1175/JCLI-D-16-0720.1

-Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, The Global Land Data Assimilation System, Bull. Amer. Meteor. Soc., 85(3), 381–394, 2004.

2.14. [Lines 65-68: Similar to abstract, these sentences are awkwardly written.]

Author's response to 2.14 :

P.2, L.65-68, "ERA-5 important changes relative to ERA-Interim former atmospheric reanalysis include a higher spatial and temporal resolution as well as a better global balance of precipitation and evaporation."

is now :

"ERA-5 has important changes relative to ERA-Interim former atmospheric reanalysis including a higher spatial and temporal resolution as well as a better global balance of precipitation and evaporation."

2.15. [Lines 96: Change to say 'Section 2 presents the details of two atmospheric ..']

Author's response to 2.15 : done

2.16. [Line 120: Change to say 'which allows it to use ..']

Author’s response to 2.16 : done

2.17. [Line 132: Add a comma after ‘study’]

Author’s response to 2.17 :
Sentence is now : “This study makes use of the CO_2-responsive version of the ISBA LSM included in the open-access SURFEX modelling platform of Météo-France (Masson et al., 2013).”

2.18. [Section 2.3: I would say ‘interpolated to’ rather than ‘interpolated at’: What interpolation methods were used?]

Author’s response to 2.18 :
A bi-linear interpolation from the native reanalysis grid to the regular grid, it is now added in the revised version of the manuscript (L.202).

2.19. [Line 217: Kumar et al. (2009) is not in the list of references.]

Author’s response to 2.19
Reference to Kumar et al. (2009) is now in the list of references along with Kumar et al. (2018) that we find appropriate in this context.

- Kumar, S.V, R. H. Reichle, R. D. Koster, W. T. Crow, and C. Peters-Lidard. 2009. "Role of Subsurface Physics in the Assimilation of Surface Soil Moisture Observations." *J. Hydrometeor*, 10 (6): 1534-1547 [10.1175/2009JHM1134.1]
- Kumar, S.V., M. Jasinski, D. Mocko, M. Rodell, J. Borak, B. Li, H. Kato Beaudoing, and C.D. Peters-Lidard: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment. J. Hydrometeor., 0, https://doi.org/10.1175/JHM-D-17-0125.1

2.20. [Line 237: I would not say ‘artificially increasing the perceived agreement’ – Just that the skill values are higher because it includes the seasonal cycle.]

Author’s response to 2.20 : done
P.7, L.236-239: “Soil moisture time series usually show a strong seasonal pattern possibly artificially increasing the perceived agreement between modeled and observed data sets.” is now : “Soil moisture time series usually show a strong seasonal pattern possibly increasing the skill values between modeled and observed data sets.”

2.21. [Line 239: ‘Monthly averaged are also computed’ (?)]

Author’s response to 2.21 : Please see Author’s answer to 2.23 which provides clarification on this paragraph.

2.22. [Line 240: Change ‘week’ to ‘weeks’.]

Author’s response to 2.22 : done

2.23. [Lines 239-242: It sounds like this you are really computing the z-scores rather than anomalies, since you are scaling the differences with standard deviation]

Author’s response to 2.23 :

According to Reviewer#2's comments 2.21 and 2.23, this paragraph has been clarified as follow:

P.7/8, L236-239: "To avoid seasonal effects, time series of anomalies from a moving monthly averaged are also computed. At each grid and observation points, the difference to the mean is calculated using a sliding window of five week and the difference is scaled by the standard deviation as in Albergel et al., (2013b). Anomaly time series reflect the time-integrated impact of antecedent meteorological forcing."
<u>is now:</u>
"To avoid seasonal effects, monthly anomaly time-series are calculated. At each grid and observation point, the difference from the mean is produced for a sliding window of five weeks, and the difference is scaled to the standard deviation as in Albergel et al., (2013b). For each surface soil moisture estimate at day (i), a period F is defined, with F= [i−17,i+17](corresponding to a five-week window). If at least five measurements are available in this period, the average soil moisture value and the standard deviation are calculated. Anomaly time series reflect the time-integrated impact of antecedent meteorological forcing."

2.24. [Line 245: What significance test is done to compute the p-values? This varies depending on the metric of interest. In particular, since several derived metrics (NICs) are used here, how did you compute the statistical significance?]

Author's response to 2.24 :
The p-values is applied on correlation values and only stations with significant correlation values (at p-values < 0.05) are retained, it is now clarified in the revised version of the manuscript. Table II on soil moisture evaluation now shows the 95% confidence interval for all metrics (95% confidence interval of the median derived from a 10000 samples bootstrapping). Please see Author's response to 2.3, also.

2.25. [Line 265: Change from 'an NSE' to 'a NSE']

Author's response to 2.25 : done

2.26. [Line 323: Change 'exercises' to 'studies']

Author's response to 2.26 : done

2.27. [Line 347: Change 'equivalents' to 'equivalent']

Author's response to 2.27 : done

2.28. [Figure 1: As the authors describe, this figure is not very useful. The lines are too close to each other in most part. It will be easier to see them if you plot the differences (relative to ei_S ; then you only have two lines). Another option is to show a seasonal cycle rather than the entire time series.]

Author's response to 2.28
Agreed, seasonal cycles would prove better, a new figure 1 along with a new caption (please see below) has been produced. Text has also been slightly modified to match with the new figure.

P.12, L.364-365: "Averaged time-series of the six main land surface variables evaluated in this study over the whole domain for 2010-2016 are illustrated on figure 1, [...]"
<u>is now:</u>

“Seasonal time-series of the six main land surface variables evaluated in this study over the whole domain for 2010-2016 are illustrated on figure 1, [...]”

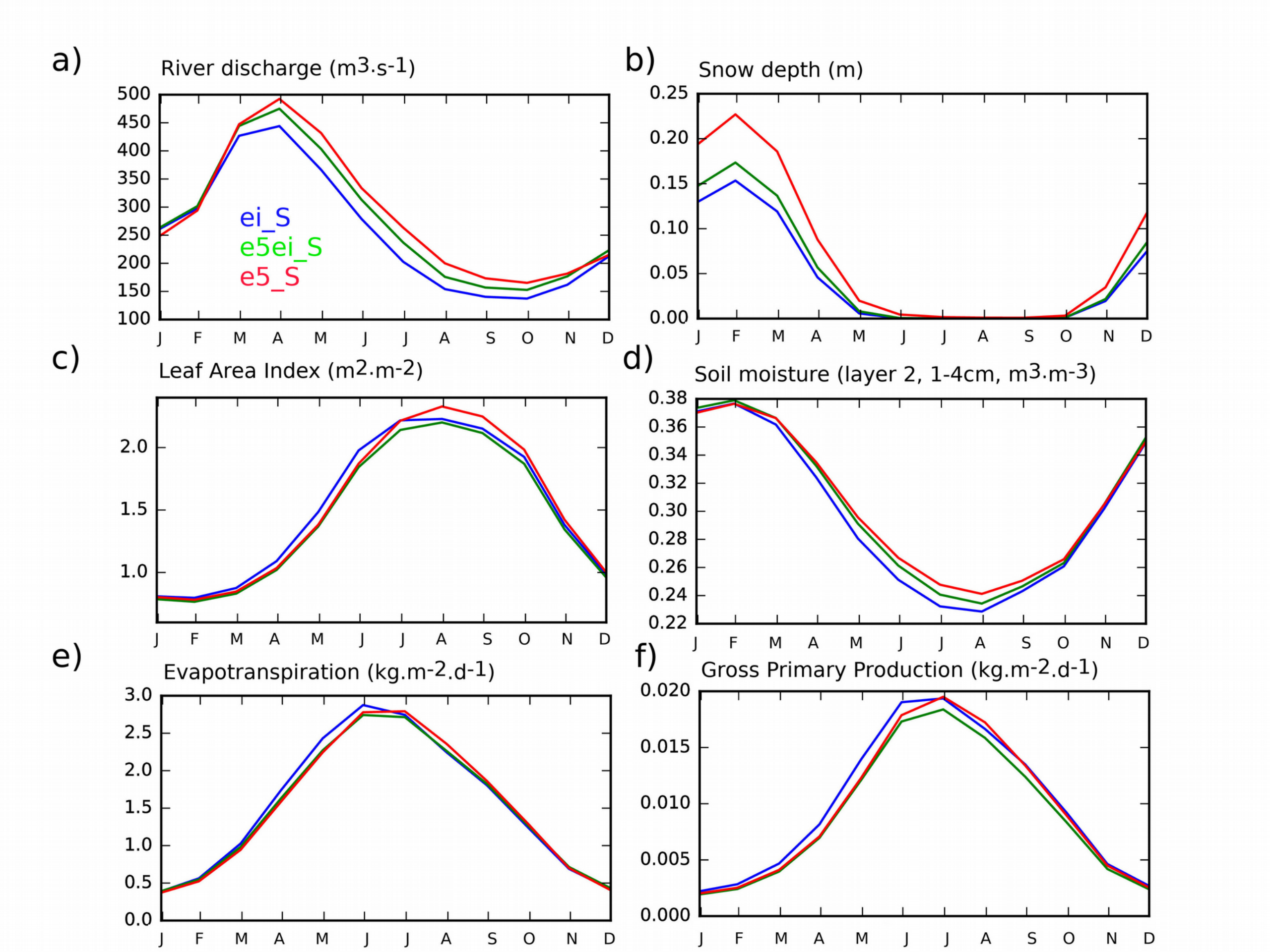

Illustration 1: Seasonal cycle for the 6 main land surface variables evaluated in this study over the whole domain for 2010-2016: (a) river discharge, (b) snow depth, (c) leaf area index , (d) liquid soil moisture in the second layer of soil (1-4 cm depth), (e) evapotranspiration and (f) gross primary production. Land surface variables simulated with SURFEX forced by ERA-Interim (ei_S) are in blue, by ERA-5 (e5_S) with precipitation from ERA-Interim (e5ei_S) in blue and by ERA-5 in red.

2.29. [Lines 390-395: Say NSE rather than ‘efficiency’]

Author’s response to 2.29 : done

2.30. [Lines 465: Change ‘Aprils’ to ‘April’]

Author’s response to 2.30 : done

2.31. [Lines 479-480: What does ‘lasting dataset’ mean?]

Author’s response to 2.30 :
“the three lasting dataset” has now been removed from the revised version of the manuscript.

P.15, L.476-480, “Figure 10 illustrates seasonal scores between ISBA LSM forced by either ERA-Interim (ei_S in blue) ERA-5 but ERA-Interim precipitation (e5ei in green) or ERA-5 (e5_S in red) ***and the three lasting dataset***; (fig10.a, fig10.b) evapotranspiration estimates from the GLEAM

project over 2010-2016, (fig10.c, fig10.d) upscaled GPP from the FLUXCOM project over 2010-2013 and (fig10.e, fig10.f) LAI estimates from the Copernicus GLS project over 2010-2016. Left column (fig10.a, c and e) are for RMSD and right column (fig8.b, d, e) for correlations."
<u>is now :</u>
"Figure 9 illustrates seasonal scores between ISBA LSM forced by either ERA-Interim (ei_S in blue) ERA-5 but ERA-Interim precipitation (e5ei in green) or ERA-5 (e5_S in red) for; (fig9.a, fig9.b) evapotranspiration estimates from the GLEAM project over 2010-2016, (fig9.c, fig9.d) upscaled GPP from the FLUXCOM project over 2010-2013 and (fig9.e, fig9.f) LAI estimates from the Copernicus GLS project over 2010-2016. Left column (fig9.a, c and e) are for RMSDs and right column (fig9.b, d, e) for correlations."

2.32. [Lines 509-510: Fix – 'It is however acknowledge that ..']

Author's answer to 2.32: done

ERA-5 and ERA-Interim driven ISBA land surface model simulations : Which one performs better?

Clement Albergel[1], Emanuel Dutra [2], Simon Munier[1], Jean-Christophe Calvet[1], Joaquin Munoz-Sabater[3], Patricia de Rosnay[3], Gianpaolo Balsamo[3]

[1]CNRM UMR 3589, Météo-France/CNRS, Toulouse, France

[2]Instituto Dom Luiz, IDL, Faculty of Sciences, University of Lisbon, Portugal

[3]ECMWF, Reading, UK

Corresponding author : Clement Albergel, clement.albergel@meteo.fr

Abstract – The European Centre for Medium Range Weather Forecast (ECMWF) recently released a first 7-year segment of its latest atmospheric reanalysis: ERA-5 over 2010-2016. ERA-5 has important changes relative to ERA-Interim former atmospheric reanalysis includinge a higher spatial and temporal resolution as well as a more recent model and data assimilation system. ERA-5 is foreseen to replace ERA-Interim reanalysis and one of the main goals of this study is to assess whether ERA-5 can enhance the simulation performances with respect to ERA-Interim when it is used to force a Land-Surface-Model (LSM). To that end, both ERA-5 and ERA-Interim are used to force the ISBA (Interactions between Soil, Biosphere, and Atmosphere) LSM fully coupled with the Total Runoff Integrating Pathways (TRIP) scheme adapted for the CNRM (Centre National de Recherches Météorologiques) continental hydrological system within the SURFEX (SURFace Externalisée) modelling platform of Météo-France. Simulations cover the 2010-2016 period at half a degree spatial resolution.

ERA-5 impact on ISBA LSM relative to ERA-Interim is evaluated using remote sensing and in-situ observations covering a substantial part of the land surface storage and fluxes over the CONtinuous US (CONUS) domain.~~ERA-5 impact on the ISBA LSM with respect to ERA-Interim is assessed over a data-rich area: North America. A comprehensive evaluation of ERA-5 impact is conducted using remote sensing and in-situ observations covering a substantial part of the land surface storage and fluxes.~~ The remote sensing observations include: (i) satellite-driven model estimates of land evapotranspiration , (ii) upscaled ground-based observations of gross primary production, (iii) satellite derived estimates of surface soil moisture as well as (iv) satellite derived estimates of Leaf Area Index (LAI). The in-situ observations cover (i) soil moisture, (ii) turbulent heat fluxes, (iii) river discharges and (iv) snow depth. ERA-5 leads to a consistent improvement over ERA-Interim as verified with the use of these 8 independent observations of different land status and of the model simulations forced by ERA-5 when compared with ERA-Interim.~~;~~ This is particularly evident for the land surface variables linked to the terrestrial hydrological cycle while variables linked to

vegetation are less impacted. Results also indicate that while precipitation provides, to a large extent~~d~~, improvements in surface fields (e.g. large improvement in the representation of river discharge and snow depth), the other atmospheric variables play an important role, contributing to the overall improvements. These results highlight the importance of enhanced meteorological forcing quality provided by the new ERA-5 reanalysis, which will pave the way for a new generation of land-surface developments and applications.

1. **Introduction**

Observing and simulating the response of land biophysical variables to extreme events is a major scientific challenge in relation to the adaptation to climate change. To that end, Land Surface Models (LSMs) constrained by high quality gridded atmospheric variables and coupled with river routing models are essential~~s~~ (Schellekens et al., 2017, Dirmeyer et al., 2006). Such LSMs should represent land surface biogeophysical variables like surface and root zone soil moisture (SSM and RZSM, respectively), biomass and Leaf Area Index (LAI) in a way that is fully consistent with the representation of surface and energy flux as well as river discharge simulations. Land surface simulations, such as those from the Global Soil Wetness Project (GSWP, Dirmeyer et al., 2002, 2006; Dirmeyer, 2011), combined with seasonal forecasting systems have been of paramount importance in triggering progress~~es~~ in land-related predictability as documented in the Global Land–Atmosphere Coupling Experiments (GLACE, Koster et al., 2009a, 2011). The land surface state estimates used in those studies were generally obtained with offline (or stand-alone) model simulations, forced by 3-hourly meteorological fields from atmospheric reanalysis. In the past decade, several improved global atmospheric reanalysis of the satellite era (1979-onwards) have been produced that enable new applications of offline land surface simulations. ~~Amongst them are NASA's Modern Era Retrospective-analysis for Research and Applications (MERRA; Rienecker et al., 2011) as well as ECMWF's (European Centre for Medium-Range Weather Forecasts) Interim reanalysis (ERA-Interim; Dee et al., 2011). Their offline use in LSMs led to global land surface variables (LSVs) reanalysis datasets that can support e.g. water resources analysis (Schellekens et al., 2017), like MERRA-Land (Reichle, 2011) and ERA-Interim/Land (Balsamo et al., 2015).~~Amongst them are NASA's Modern Era Retrospective-analysis for Research and Applications (MERRA; Rienecker et al., 2011 and MERRA2; Gelaro et al. 2016,) as well as ECMWF's (European Centre for Medium-Range Weather Forecasts) Interim reanalysis (ERA-Interim; Dee et al., 2011). Their offline use in either LSMs or Land Data Assimilation System (LDAS), with or without meteorological corrections (e.g., precipitations) led to global Land Surface Variables (LSVs) reanalysis datasets that can support e.g. water resources analysis (Schellekens et al., 2017), like MERRA-Land and MERRA2-Land (Reichle, 2011; 2017), ERA-Interim/Land (Balsamo et al., 2015), the forthcoming ERA5-Land (Muñoz-Sabater et al., 2018), the North American LDAS (NLDAS, Mitchel et al., 2004), the Global LDAS (GLDAS, Rodell et al., 2004) and LDAS-Monde (Albergel et al., 2017).The quality of those offline land surface simulations relies on the accuracy of the forcing and of the realism of the land surface model itself (Balsamo et al., 2015).

ECMWF recently released the first 7-year segment of its latest atmospheric reanalysis: ERA-5 over 2010-2016. ERA-5 has important changes relative to ERA-Interim former atmospheric reanalysis includinge a higher spatial and temporal resolution as well as a better global balance of precipitation and evaporation. ~~It will eventually replace ERA-Interim reanalysis. Assessing ERA-5 ability to force a LSM with respect to ERA-Interim is therefore highly relevant. To that end, ERA-5, ERA-Interim as well as a combination of both (ERA-5 with precipitation of ERA-Interim) are used to constrain~~ As ERA-5 will eventually replace ERA-Interim reanalysis assessing its ability to force a LSM with respect to ERA-Interim is highly relevant. In this study, ERA-5, ERA-Interim as well as a combination of both (ERA-5 with precipitation of ERA-Interim) are used to constrain the CO_2-responsive version of the Interactions between Soil, Biosphere, and Atmosphere (ISBA, Noilhan and Mahfouf, 1996; Calvet et al., 1998, 2004 ; Gibelin et al., 2006) LSM fully coupled with the CNRM (Centre National de Recherches Météorologiques) version of the Total Runoff Integrating Pathways (TRIP, Oki et al., 1998) continental hydrological system (CTRIP hereafter, Decharme et al., 2010) within the SURFEX (SURFace Externalisée, Masson et al., 2013) modeling system of Météo-France. ISBA models leaf-scale physiological processes and plant growth, while transfer of water and heat through the soil rely on a multilayer diffusion scheme.

In this study SURFEX is applied over a data-rich area: North America (longitudes from 130.0ºW to 60.0ºW, latitudes from 20.0º to 55.0º N) for the period 2010-2016. ERA-5 added values with respect to ERA-Interim is assessed by providing verification and diagnostics comparing ISBA land surface variables outputs when forced by either ERA-5, ERA-Interim, ERA-5 with ERA-Interim precipitations to several in-situ measurement data sets as well as satellite-derived estimates of Earth Observations. Namely, In-situ measurements of (i) soil moisture from the USCRN network (US Climate Reference Network, Bell et al., 2013) spanning all over the United States of America and (ii) of turbulent heat fluxes from FLUXNET-2015 (http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/) are used in the evaluation, together with (iii) river discharges from the United States Geophysical Survey (USGS, https://waterwatch.usgs.gov/) and (iv) snow depth measurements from the Global Historical Climatology Network (GHCN, Menne et al., 2012a, 2012b). Are also used: (i) satellite-driven model estimates of land evapotranspiration from the Global Land Evaporation

Amsterdam Model (GLEAM, Martens et al., 2017) project, (ii) upscaled ground-based observations of gross primary production from the FLUXCOM project (Jung et al., 2017), (iii) satellite derived estimates of SSM from the Climate Change Initiative (CCI) program of the European Space Agency (ESA-CCI SSMv4 Dorigo et al., 2015, 2017) as well as (iv) satellite derived estimates of LAI from the Copernicus Global Land Service program (CGLS, http://land.copernicus.eu/global/).

Section 2 presents the details of two atmospheric reanalyses data sets, ERA-Interim and ERA-5, the SURFEX model configuration as well as the evaluation strategy with the observational data sets. Section 3 provides a set of statistical diagnostics to assess and evaluate ERA-5 impact in ISBA with respect to ERA-Interim. Finally, section 4 provides perspectives and future research directions.

2. **Methodology**

2.1. *ERA-Interim and ERA-5 reanalysis*

ERA-Interim is a global atmospheric reanalysis produced by ECMWF (Dee et al. 2011). It uses Integrated Forecast System (IFS) version 31r1 (more information at https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/ifs-documentation) with a spatial resolution of about 80 km (T255) and with analyses available for 0000, 0600, 1200, and 1800 UTC. It covers the period from 1 January 1979 onward and continues to be extended forward in near-real time (with a delay of approximately 1 month). Reanalyses merge observations and model forecasts in data assimilation methods to provide an accurate and reliable description of the climate over the last few decades. Berrisford et al. (2009) provide a detailed description of the ERA-Interim product archive. ERA-5 (Hersbach, 2016) is the latest and fifth generation of European reanalyses produced by the ECMWF and a key element of the EC-funded Copernicus Climate Change Service (C3S). It is expected that ERA-5 will replace the production of the current ERA-Interim reanalysis (Dee et al., 2011) before the end of 2018, from 1979 to close to Near Real Time (NRT) period, i.e., in ERA-5 regular routine updates will be conducted to keep close to NRT. In a second phase, an extension back to 1950 is also expected. ERA-5 adds different characteristics to ERA-Interim reanalysis, which makes it richer in term of climate information.

ERA-5 uses one of the most recent versions of the Earth system model and data assimilation methods applied at ECMWF, which makes it able to use modern parameterizations of Earth processes compared to older versions used in ERA-Interim. For instance, developments were done at ECMWF, which allows ~~ERA-5 to apply~~it to use a variational bias scheme not only to satellite observations, but also to ozone, aircraft and surface pressure data. ERA-5 benefits also of reprocessed data sets that were not ready yet during the production of ERA-Interim. Two other important features of ERA-5 are the improved temporal and spatial resolution, from 6-hourly in ERA-Interim to hourly analysis in ERA-5, and from 79 km in the horizontal dimension and 60 levels in the vertical, to 31 km and 137 levels in ERA-5. Finally, ERA-5 also provides an estimate of uncertainty through the use of a 10-member Ensemble of Data Assimilations (EDA) at a coarser resolution (63 km horizontal resolution) and 3-hourly frequency.

2.2. *SURFEX modeling system*

2.2.1. *The ISBA Land-Surface-Model*

~~In t~~This study makes use of the CO_2-responsive version of the ISBA LSM ~~is~~ included in the open-access SURFEX modelling platform of Météo-France (Masson et al., 2013). The most recent version of SURFEX (version 8.1) is used ~~in this study~~ with the “NIT” biomass option for ISBA. The latter simulates the diurnal cycle of water and carbon fluxes, plant growth and key vegetation variables like LAI and above ground biomass on a daily basis. It can be coupled to the CTRIP river routing model in order to simulate the streamflow. In this version of ISBA, a single-source energy budget of a soil–vegetation composite is computed. Also, the ISBA parameters are defined for 12 generic land surface patches, which include nine plant functional types (needle leaf trees, evergreen broadleaf trees, deciduous broadleaf trees, C3 crops, C4 crops, C4 irrigated crops, herbaceous, tropical herbaceous, and wetlands), bare soil, rocks, and permanent snow and ice surfaces. A more comprehensive model description can be found in Masson et al. (2013).

ISBA accounts for the atmospheric CO_2 concentration on stomatal aperture (Calvet et al., 1998, 2004; Gibelin et al., 2006). Also, photosynthesis and its coupling with stomatal conductance at a leaf level are accounted for. The vegetation net assimilation of CO_2 is estimated and used as an input to a simple vegetation growth sub-model able to predict LAI: photosynthesis drives the dynamic evolution of the vegetation biomass and LAI variables in response to atmospheric and climate conditions. During the growing phase, enhanced photosynthesis corresponds to a CO_2 uptake, which leads to vegetation growth. In contrast, lack of photosynthesis leads to higher mortality rates. The gross primary production (GPP) is defined as the carbon uptake while the ecosystem respiration (RECO) is the release of CO_2, the difference between these two quantities being the net ecosystem CO_2 exchange (NEE). Evaporation due to (i) plant transpiration, (ii) liquid water intercepted by leaves, (iii) liquid water contained in top soil layers, and (iv) the sublimation of the snow and soil ice are combined to represent the total evaporative flux.

ISBA 12-layers explicit snow scheme (Boon et Etchevers, 2001, Decharme et al., 2016) as well as its multi-layer soil diffusion scheme (ISBA-Dif) are used. The later is based on the mixed form of the Richards equation (Richards, 1931) and explicitly solves the one-dimensional Fourier law. It also incorporates soil freezing processes developed by Boone et al. (2000) and Decharme et al.

(2013). The total soil profile is vertically discretized; both the temperature and moisture of each soil layer are computed accordingly to their textural and hydrological characteristics. The Brookes and Corey model (Brooks and Corey, 1966) determines the closed-form equations between the soil moisture and the soil hydrodynamic parameters, including the hydraulic conductivity and the soil matrix potential (Decharme et al., 2013). The defaults discretization with 14 layers over 12 m depth is used. The lower boundary of each layer being: 0.01, 0.04, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.5,2, 3, 5, 8 and 12 m deep (see Fig. 1 of Decharme et al., 2011). Amounts of clay, sand and organic carbon in the soil determine the by thermal and hydrodynamic soil properties (Decharme et al., 2016). They are taken from the Harmonised World Soil Database (HWSD, Wieder et al., 2014). As for hydrology, the infiltration, surface evaporation and total runoff are accounted for in the soil water balance. The infiltration rate defines the discrepancy between the surface runoff and the throughfall rate. The later being defined as the sum of rainfall not intercepted by the canopy, dripping from the canopy (i.e., interception reservoir) as well as snow melt water. The soil evaporation affects only the superficial layer (top 1 cm) and is proportional to its relative humidity. Transpiration water from the root zone (the region where the roots are asymptotically distributed) follows the equations in Jackson et al. (1996). Canal et al. (2014) provide more information on the root density profile.

Both the surface runoff (the lateral subsurface flow in the topsoil) and a free drainage condition at the bottom soil layer contribute to ISBA total runoff. The Dunne runoff (i.e. when no further soil moisture storage is available) and lateral subsurface flow from a sub-grid distribution of the topography are computed using a basic TOPMODEL approach. The Horton runoff (i.e. when rainfall has exceeded infiltration capacity) is estimated from the maximum soil infiltration capacity and a sub-grid exponential distribution of the rainfall intensity.

2.2.2. *The CTRIP hydrological system*

CTRIP is driven by three prognostic equations corresponding to (i) the groundwater, (ii) the surface stream water and (ii) the seasonal floodplains. Streamflow velocity is computed using the Manning formula as described in Decharme et al. (2010). When the river water level overtops the river-bank, it fills up the floodplain reservoir which empties when the water level drops below this threshold (Decharme et al., 2012). Occurrence of flooding impacts the ISBA soil hydrology through infiltration and it influences the overlying atmosphere via free surface water evaporation and precipitation interception, also. The groundwater scheme is based on the two-dimensional groundwater flow equation for the piezometric head (Vergnes and Decharme,2012). Its coupling with ISBA enables accounting for the presence of a water table under the soil moisture column. It allows upward capillary fluxes into the soil (Vergnes et al., 2014). CTRIP is coupled to ISBA through OASIS-MCT (Voldoire et al., 2017). Once a day, ISBA provides CTRIP with updates on runoff, drainage, groundwater and floodplain recharges, and CTRIP feedbacks to ISBA the water table depth or rise, floodplain fraction, and flood potential infiltration. The current CTRIP version consists of a global streamflow network at 0.5°x0.5° spatial resolution.

2.3. *Evaluation strategy and data sets*

Three experiments are considered for the evaluation; (i) SURFEX forced by ERA-Interim, all atmospheric variables interpolated to~~at~~ 0.5ºx0.5º spatial resolution (referred as ei_S hereafter, the benchmark experiment), (ii) SURFEX forced by ERA-5 all atmospheric variables interpolated at 0.5ºx0.5º spatial resolution except precipitation (rain and snow interpolated to hourly time steps assuming a constant flux) that comes from ERA-Interim (referred as e5ei_S hereafter). (iii) SURFEX forced by ERA-5, all atmospheric variables interpolated at 0.5ºx0.5º spatial resolution (referred as e5_S hereafter). A bi-linear interpolation from the native reanalysis grid to the regular grid has been used. For all three experiments, the first year (2010) was spun up 20 times to allow the model to reach equilibrium. Comparing e5_S to ei_S provides the overall improvements from ERA-Interim to ERA-5. The idealized e5ei_S simulation was carried out to assess the role of precipitation changes from ERA-Interim to ERA-5.

This study makes use of several in-situ measurement data sets as well as satellite-derived estimates of Earth Observations that are described in the next two sections. The different performance metrics used for the evaluation are described, also. Their choice is of crucial interest; it is governed by the nature of the variable itself and is influenced by the purpose of the investigation and its sensitivity to the considered variables (Stanski et al. 1989). No single metric or statistic can capture all the attributes of environmental variables. Some are robust in respect to some attributes while insensitive to others (Entekhabi et al. 2010). While performance metrics like the correlation coefficient, unbiased root mean squared differences, root mean squared differences, efficiency score (depending on the considered variable) are first applied to the three simulations independently, metrics like the Normalized Information Contribution (NIC, e.g. Kumar et al., 2009) are then used to quantify improvement or degradation from a data set to another. Table I summarises the different dataset used for the evaluation as well as the performance metrics used.

2.3.1. *In situ measurement of soil moisture, river discharges, snow depth and fluxes*

USCRN is a network of climate-monitoring stations maintained and operated by the National Oceanic and Atmospheric Administration (NOAA). It aims at providing climate-science-quality measurements of air temperature and precipitation. To increase the network's capability of monitoring soil processes and drought, soil observations were added to USCRN instrumentation. In 2011, the USCRN team completed at each USCRN station in the conterminous United States the installation of triplicate-configuration soil moisture and soil temperature probes at five standards depths (5, 10, 20, 50, and 100 cm) as prescribed by the World Meteorological Organization. 111 stations present data between 2009 and 2016. Stations provide data at an hourly time step. Similar to prior study, datasets potentially affected by frozen conditions were masked out using an observed temperature threshold of 4ºC (e.g. Albergel et al., 2013a). The second layer of soil of ISBA between 1 and 4 cm depth (the diffusion scheme is used in this study) is compared to in situ measurements at 5 cm depth at a three hourly time step (model output) between April and September in order to avoid as much as possible frozen conditions. The ability of ei_S, e5ei_S and e5_S to reproduce surface soil moisture variability is first assessed using the correlation coefficient (R) and unbiased Root Mean Square Differences (ubRMSD). Climatology differences between model and in-situ observations make a direct comparison difficult (Koster et al., 2009b). ~~Soil moisture time series usually show a strong seasonal pattern possibly artificially increasing the perceived agreement between modeled and observed data sets.~~Soil moisture time series usually show a strong seasonal pattern possibly increasing the skill values between modeled and observed data sets. ~~To avoid seasonal effects, time series of anomalies from a moving monthly averaged are also computed. At each grid and observation points, the difference to the mean is calculated using a sliding window of five week and the difference is scaled by the standard deviation as in Albergel et al., (2013b). Anomaly time series reflect the time-integrated impact of antecedent meteorological forcing.~~To avoid seasonal effects, monthly anomaly time-series are calculated. At each grid and observation point, the difference from the mean is produced for a sliding window of five weeks, and the difference is scaled to the standard deviation as in Albergel et al., (2013b). For each surface soil moisture estimate at day (i), a period F is defined, with F= [i−17,i+17](corresponding to a five-week window). If at least five measurements are available in this period, the average soil moisture value and the standard deviation are calculated. Anomaly time series reflect the time-integrated impact of antecedent meteorological forcing. The latter is mainly reflected in the upper layer of soil. The correlation coefficient is also computed for anomaly time-series (R_{ano}). For correlations, t~~T~~he p-value (a measure of the correlation significance) is also calculated indicating the significance of the test (as in Albergel et al. 2010), and only cases where the p-value is below 0.05 (i.e., the correlation is not a coincidence) are retained. Stations with non-significant R values can be considered suspect and are excluded from the computation of the network average metrics. This process may remove some reliable stations too (e.g., in areas where the model might not realistically represent soil moisture).

Over 2010-2016 river discharge from ei_S, e5ei_S and e5_S are compared to daily streamflow data from the U.S. Geological Survey (USGS; http://nwis.waterdata.usgs.gov/nwis). Data are chosen for sub-basins with large drainage areas (10,000km^2 or greater) and with a long observation time series

(4 years or more). Smaller basins are excluded due to the low resolution of CTRIP (0.5°x0.5°). It is common to express observed and simulated river discharge (Q) data in m^3s^{-1}. Given that the observed drainage areas may differ slightly from the simulated ones, specific discharge in $mm.d^{-1}$ (the ratio of Q to the drainage area) is used in this study, similarly to Albergel et al., 2017. Stations with drainage areas differing by more than 20% from the simulated ones are also discarded. This criterion aims to ensure a meaningful comparison between observed and simulated values. It is necessary for coping with the significant distortions in the model representation of the river network that are caused by the coarse spatial resolution of the CTRIP global river network (0.5°x0.5°). Impact on Q is evaluated using the efficiency score (NSE) (Nash and Sutcliff, 1970). NSE evaluates the model ability to represent the monthly discharge dynamics and is given by:

$$NSE = 1 - \frac{\sum_{t=1}^{T} \left(Q_s^t - Q_o^t\right)^2}{\sum_{t=1}^{T} \left(Q_o^t - \overline{Q_o^t}\right)^2} \qquad (1)$$

where Q_s^t is the simulated river discharge (by either ei_S, e5ei_S or e5_S) at time t and Q_o^t is observed river discharge at time t, T is the total number of days and $\overline{Q_o^t}$ is the average observed discharge. NSE can vary between –∞ and 1. A value of 1 corresponds to identical model predictions and observed data. A value of 0 implies that the model predictions have the same accuracy as the the mean of the observed data. Negative values indicate that the observed mean is a more accurate predictor than the model simulation. Only stations with a~~n~~ NSE greater than -1 for the benchmark experiment, ei_S, are considered, leading to 172 stations over the considered domain. A Normalized Information Contribution (NIC as in Kumar et al. ; 2009) measure is then computed to quantify the improvement or degradation due to the specific atmospheric reanalysis used to force ISBA. The NIC_{NSE} values are computed for both e5_S and e5ei_S with respect to ei_S as:

$$NIC_{NSE(e5;5ei)} = \frac{NSE_{(e5;e5ei)} - NSE_{(ei)}}{1 - NSE_{(ei)}} \qquad (2)$$

The NIC_{NSE} metric provides a normalized measure of the improvement through the use of either NSE_{e5ei} or NSE_{e5} as a fraction of the maximum possible skill improvement ($1-NSE_{ei}$). ~~Positive and negative NIC_{NSE} values indicate improvements and degradations in either e5_S or e5ei_S relative to ei_S river discharges estimates, respectively.~~Positive and negative NIC_{NSE} values indicate improvements and degradations in either e5_S or e5ei_S relative to ei_S river discharges estimates, respectively. NICs along with their 95% confidence interval of the median derived from a 10000 samples bootstrapping are provided for e5_S, e5ei_S. The ratio of simulated and observed river discharges is computed also (Q_s^t / Q_o^t), the closer to one it is, the better the simulated river discharges are.

The Global Historical Climatology Network (GHCN) Daily dataset, developed to meet the needs of climate analysis and monitoring studies that require data at a daily time resolution contains records from over 75000 stations in 179 countries and territories (Menne et al., 2012a, b). Numerous daily variables are provided, including maximum and minimum temperature, total daily precipitation, snowfall, and snow depth. In this study, over North America, stations with daily snow depth data from 2010-2016, with less than 10% missing and at least 15 days of snow presences per year on average (to avoid using stations always reporting zero snow depth) are used, it ~~is~~results in ~~about~~1901 stations out of 20~~00~~56~~ stations~~. The ability of ei_S, e5ei_S and e5_S to reproduce snow depth and its variability is assessed using the bias, correlation coefficient (R) and~~,~~ unbiased Root Mean Square Difference (ubRMSD) ~~and mean absolute error (MAE). In order to provide an easier measure of the added value of e5_S and e5ei_S, statistics are also normalized with respect to ei_S, NIC is used for R:~~

~~(3)~~

~~Normalized MAE,, and ubRMSD,, are respectively computed as follow:~~

~~(4)~~

~~(5)~~

Daily observations of sensible and latent heat fluxes from the FLUXNET-2015 dataset with at least 2-yr of data are used over 2010-2016 to evaluate e5_S, e5ei_S and ei_S ability to reproduce flux variability. The FLUXNET-2015 dataset includes data collected at sites from multiple regional flux networks as well as several improvements to the data quality control protocols and the data processing pipeline (http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/). 37 stations are retained for the evaluations, two metrics are considered: R and ~~ub~~RMSD.~~ Eq.(3) is also applied as well as Eq.(5) on ubRMSD.~~

Performance metrics are applied to each individual station of each network; thereafter, network metrics are computed by providing the median values of the statistics from the individual stations within each network. For each metrics, the 95% confidence interval of the median derived from a 10000 samples bootstrapping is provided.

2.3.2. *satellite derived estimates of surface soil moisture, leaf area index, land evapotranspiration and gross primary production*

In response to the GCOS endorsement of soil moisture as an essential climate variable, the European Space Agency Water Cycle Multimission Observation Strategy (WACMOS) project and Climate Change Initiative (CCI; http://www.esa-soilmoisture-cci.org) have supported the generation of a surface soil moisture product based on multiple microwave sources (ESA-CCI SSM hereafter). The first version of the combined product was released in June 2012 by the Vienna University of Technology (Liu et al. 2011, 2012; Wagner et al., 2012). Several authors (e.g. Albergel et al., 2013a, b; Dorigo et al., 2015, 2017) have highlighted the quality and stability over time of the product. Despite some limitations, this data set has already shown potential in assessing model performance (e.g., Szczypta et al., 2014; van der Schrier et al., 2013). In this study the ESA CCI SM-combined latest version of the product (v4) which merges SSM observations from seven microwave radiometers (SMMR, SSM/I, TMI, ASMR-E, WindSat, AMSR2, SMOS) and four scatterometers (ERS-1 and 2 AMI and MetOp-A and B ASCAT) into a combined data set covering the period November 1978 to December 2016. Data are in volumetric (m^3m^{-3}) units and quality flags (snow coverage or temperature below 0° and dense vegetation) are provided. For a more comprehensive overview of the product see Dorigo et al. (2015, 2017). As topographic relief is known to negatively affect remote sensing estimates of soil moisture (Mätzler and Standley, 2000), the time series for pixels whose average altitude exceeded 1500 m above sea level were discarded. Data on pixels with urban land cover fractions larger than 15% were also discarded, to limit the effects of artificial surfaces. The altitude and urban area thresholds were set according to Draper et al. (2011) and Barbu et al. (2014), who processed satellite-based SSM retrievals for data assimilation ~~exercises~~studies with the ISBA LSM. As for in situ measurements of soil moisture, correlation is applied on both the volumetric and anomaly time series.

The GEOV1 LAI used in this study is produced by the European Copernicus Global Land Service project (http://land.copernicus.eu/global/) as evaluated in Boussetta et al. (2015). The LAI observations are retrieved from the SPOT-VGT and then PROBA-V (from 1999 to present) satellite data according to the methodology proposed by Baret et al. (2013). As in Barbu et al. (2014), the 1 km spatial resolution observations are interpolated by an arithmetic average to the 0.5° model grid points, if at least 50 % of the observation grid points are observed (i.e half the maximum amount). LAI observations have a temporal frequency of 10 days at best (in presence of clouds no observations are available). Correlation and root mean squared differences are used to assess ei_S, e5ei_S and e5_S ability to reproduce LAI variability. ~~Eq.(3) (NIC_R) as well as Normalized RMSD :~~

~~(7)~~

~~are used, also.~~

The GLEAM product uses a set of algorithms to estimate both terrestrial evaporation and RZSM based on satellite data (Miralles et al., 2011). It is a useful validation tool to assess model performance given that such quantities are difficult to measure directly on large scales. Potential evaporation rates are constrained by satellite-derived SSM data while the global evaporation model in GLEAM is mainly driven by various microwave remote-sensing observations. It is now a well-established dataset that has been widely used to study land–atmosphere feedbacks (e.g. Miralles et al., 2014b; Guillod et al., 2015) as well as trends and spatial variability in the hydrological cycle (e.g. Jasechko et al., 2013; Greve et al., 2014; Miralles et al., 2014a; Zhang et al., 2016). This study makes use of the latest version available, v4.0 It is a 37-year dataset spanning from 1980 to 2016 and is derived from a variety of sources, such as vegetation optical depth (VOD) and snow water equivalent~~s~~ (SWEs), satellite-derived SSM estimates, reanalysis air temperature and radiation, as well as a multi source precipitation product (Martens et al., 2017). It is available at a spatial resolution of 0.25°0.25°. A full description of the dataset, including an extensive validation using measurements from 64 eddy-covariance towers worldwide is provided by Martens et al. (2017). As for LAI, correlation and root mean squared differences are the two performance metrics used to evaluate the representation of evapotranspiration from the three datasets.

The final product used in this study is a daily Gross Primary Production (GPP) estimate from, the FLUXCOM project (Jung et al., 2017). It is an upscaled product derived from the FLUXNET network. In FLUXCOM selected machine learning-based regression tools that span the full range of commonly applied algorithms (from model tree ensembles, multiple adaptive regression splines, artificial neural networks, to kernel methods), with several representatives of each family are used to provide a spatial upscaling of GPP at regional to global scales. It is limited to a 0.5°x0.5° spatial resolution and a daily temporal resolution over 1982–2013 (Tramontana et al., 2016). FLUXCOM fluxes can be used as a way of benchmarking LSMs on large scales (Jung et al., 2009, 2010; Beer et al., 2010; Bonan et al., 2011; Jung et al., 2011; Slevin et al., 2017). Product can be found in the Max Planck Institute for Biogeochemistry Data Portal (https://www.bgc-jena.mpg.de/geodb/projects/Home.php). Correlation and root mean squared differences are the two performance metrics used to evaluate the representation of carbon uptake from the three datasets.

3. Results

Seasonal~~Averaged~~ time-series of the six main land surface variables evaluated in this study over the whole domain for 2010-2016 are illustrated on figure 1, they are (fig1.a) river discharge (although averaging this variable over the whole domain has no real meaning, it is certainly useful to appreciate the differences between the three data set), (fig1.b) snow depth, (fig1.c) leaf area index, (fig1.d) liquid soil moisture in the second layer of soil (1-4 cm depth), (fig1.e) evapotranspiration and (fig1.f) gross primary production. Land surface variables simulated with the ISBA LSM forced by ERA-Interim (ei_S) are in blue, by ERA-5 with precipitation from ERA-Interim (e5ei_S) in green and by ERA-5 (e5_S) in red. From figure 1 one can see that river discharge, snow depth and surface soil moisture are the most impacted by the use of ERA-5, suggesting that precipitation is the main driver of the differences.

3.1. *Evaluations using in situ measurements*

This section presents the results of the comparison versus in situ observations of land surface variables from model simulations using either ei_S, e5ei_S or e5_S starting with soil moisture. The statistical scores for 2010–16 surface soil moisture from ei_S, e5ei_S and e5_S are presented in Table II. Median R values on volumetric time-series (anomaly time series) along with their 95% confidence intervals are 0.66±0.02 (0.53±0.02), 0.69±0.02 (0.54±0.04) and 0.71±0.02 (0.58±0.03) while median ubRMSD are 0.052±0.003, 0.052±0.002 and 0.050±0.003 for ei_S, e5ei_S and and e5_S, respectively. These results underline the better capability of the ISBA LSM to represent surface soil moisture variability when forced by ERA-5 reanalysis. Also the latest configuration (e5_S) presents more stations with better R values on volumetric time-series (anomaly time series) than both ei_S and e5ei; respectively 60% and 75% (out of 110 and 107 stations, respectively). This is also reflected on figure 2 illustrating correlations values on volumetric time-series (fig.2a) and anomaly time-series (fig.2b) on maps. Stars symbols represent stations for which ISBA LSM performs best when forced ERA-Interim, circles when it is forced by ERA-5 with ERA-Interim precipitations and downward pointing triangles when it is forced by all ERA-5 atmospheric variables. Both maps on figure 2 are dominated by downward pointing triangles. Fig.2c(d) shows histograms of R differences on volumetric (anomaly) time-series, for soil moisture from e5_S (in red) e5ei_S (in green) with respect to ei_S, median values of the differences are reported, also.

172 out of 344 gauging stations retained for the evaluation according to the criteria described in the methodology section presents ~~efficiency~~NSE scores in the [-1, 1] interval. Figure 3 represents performance of each dataset for this pool of stations. Fig3.a is a scatterplot of ~~efficiency~~NSE scores between in situ and simulated river discharges *Q*; ~~efficiency~~NSE scores for *Q* simulated with either ERA-5 but ERA-Interim precipitations (e5ei_S, green crosses) or ERA-5 (e5_S, red dots) function of ~~efficiency~~NSE scores for *Q* simulated using ERA-Interim (ei_S). When considering e5_S, almost all the red dots are above the 1:1 diagonal suggesting a general improvement from the use of e5_S. For a large part, e5ei_S green crosses are above this diagonal, suggesting that the improvement in e5_S does not only comes from precipitation but from other variables, also. Median ~~efficiency~~NSE values are 0.06±0.06, 0.12±0.07 and 0.24±0.05 for ei_S, e5ei_S and e5_S, respectively. Fig.3b shows an histogram of river discharges ratio for ei_S (Qr_ei in blues), e5ei_S (Qr_e5ei in green)

and e5_S (Qr_e5 in red), median values are 0.67, 075 and 0.77, respectively. While all three experiments underestimate Q (a value of 1 being a perfect match), the use of e5ei_S and e5_S leads to better results. Finally, figure 3c illustrates hydrographs for a river station in Lousiana (33.08°N, -93.85°W) representing scaled Q (using either observed or simulated drainage areas), in situ data (black crosses), simulated river discharges from ei_S (blue solid line), e5ei_S (green solid line) and e5_S (red solid line). From this hydrograph, the added value of e5_S is clear, particularly for the 2011 and 2015 main events. ~~Efficiency~~NSE scores are 0.47, 0.61 and 0.76 for ei_S, e5ei_S and e5_S, respectively. Figure 4 illustrates the added value of using e5_S (a) or e5ei_S (b) with respect to ei_S. For 156 out of the pool of 172 stations NIC_{NSE} values computed using e5_S with respect to ei_S are positive (large blue circles) showing an general improvement from the use of e5_S (representing 91% of the stations) with a median NIC_{NSE} value of 14%±0.05. When considering e5ei_S versus ei_S, they are still 118 (69%) with a median NIC_{NSE} value of 4%±0.02 suggesting that the improvement in e5_S does not only comes from precipitation but from other variables, also. It is also worth-noticing that stations where a score degradation is observed (large red circles) are located in areas known for irrigation which is not represented in ISBA. All scores computed for seasons (December-January-Februray, March-April-May, Jun-July-August, September-October-November) suggest the same ranking (not shown).

The mean snow depth bias of ei_S (see Figure 5) highlights a clear underestimation of winter snow depth accumulation mainly over the Rocky Mountains. This is likely a result of the underestimation of snowfall by ei_S associated with an overestimation of snow melt due to the coarse resolution of the ei_S reflected in a smooth topography. The replacement of all forcing variables by e5_S but keeping ei_S precipitation (e5ei_S, Fig.5b) shows a slight increase in snow depth. This result justifies the above hypothesis that part of the snow underestimation is also due to temperature issues linked with a coarse model orography. Moving to the full e5_S forcing there is a clear increase of snow depth, when compared with both ei_S and e5ei_S forced simulations resulting from an increase in snowfall in e5_S. Figure 6 presents the mean seasonal cycle of bias and ubRMSD (fig.6a) and correlations (fig.6b) over 2010-2016. In addition to the added values of e5_S in terms of the mean snow depth already presented in figure 5, the temporal variability and random errors are also improved ~~(see Figure 6)~~. Comparably with what was discussed for the mean bias, e5ei_S shows some benefits when compared with ei_S in terms ~~;of mae (N_{MAE}~5%)~~ ubRMSD ~~(N_{ubRMSD}~4%)~~ and correlation ~~(NIC_R of 0.1)~~ (median bias, ubRMSD and R values of e5_ei over the whole period are; -1.70±0.33 cm., 7.40±0.65 cm. and 0.77± 0.01, respectively, for ei_S they are; -2.11±0.33 cm., 7.58±0.65 cm. and 0.75± 0.01, respectively) while e5_S has a clear improvement ~~in mae (N_{MAE}~16%)~~, ubRMSD ~~(N_{ubRMSD}~14%)~~ and correlation ~~(NIC_R of 0.25)~~(median bias, ubRMSD and R values of e5_ei over the whole period are; -0.64±0.19 cm., 7.00±0.65 cm. and 0.82± 0.01, respectively). The improvements on the snow depth simulations are consistent throughout the entire snow covered season (see Fig.6a and b~~d~~) with a maximum improvement from January to March. These results highlight the cumulative effect of the forcing quality on the snow depth simulation. Finally Table III presents scores from the comparison of snow depth with in situ measurements, median Bias, ubRMSD and R values are given for the three seasons affected by snow (September-October-November, December-January-February and Mars-April-May) and for the whole period. e5_S always presents better scores when compared to ei_S and it is always the configuration presenting the highest percentage of stations with the best scores. Looking at the 95% confidence interval, for the correlation and bias it is clear that the changes are significant.

Results from the comparisons between ei_S, e5ei_S, e5_S and in situ sensible and latent flux measurements are presented in table IV~~II~~ and illustrated by figure 7. 37 stations present significant correlation values (at p-value < 0.05). For sensible heat flux, median correlation and **~~ub~~RMSD values are 0.62**±0.11**, 0.62**±0.11 **and 0.65**±0.11**,** ~~34.85~~39.58±3.71 **W.m^{-2},** ~~30.66~~32.89±3.86 **W.m^{-2} and** ~~30.38~~32.73±2.61 **W.m^{-2} for ei_S, e5ei_S and e5_S, respectively. For latent heat flux, they are 0.63**±0.05**, 0.62**±0.07 **and 0.70**±0.04**,** 39.00±5.38~~33.93~~ **W.m^{-2},** ~~31.66~~37.12±4.37 **W.m^{-2} and** ~~30.98~~36.66±4.94 **W.m^{-2}.** As for surface soil moisture, river discharge and snow depth, e5_S presents better results than e5ei_S and ei_S. At the station level, figure 7 illustrates scatter plots of correlations and **~~ub~~RMSD** for sensible and latent heat flux from ei_S, e5ei_S, e5_S against in situ measurements of sensible (fig.7a for correlation, fig.7c for **~~ub~~RMSD**) and latent (fig.7b for correlation, fig.7d for **~~ub~~RMSD**) heat flux. Scores for either e5ei_S (green dots) or e5_S (in red) are presented function of those for ei_S. When looking at the correlations, almost all of e5_S and e5ei_S symbols (in red and green, respectively on fig.7a, fig.7c) are above the 1:1 diagonal indicating that e5_S and e5ei_S better represent sensible and latent heat flux than ei_S. Same tendency is observed for **~~ub~~RMSD** with most of the symbol below the 1:1 diagonal. If **~~ub~~RMSD** values are comparable for e5_S and e5ei_S, R values are clearly higher for e5_S. ~~Finally figure 8 presents NIC scores based on correlations values between in situ measurements from the fluxnet sites data and (fig8.a) e5_S with respect to ei_S for sensible heat flux and (fig8.b) e5_S with respect to ei_S for latent heat flux. Normalised ubRMSD values between in situ measurements from the fluxnet sites data and (fig.8c) e5_S with respect to ei_S for sensible heat flux and (fig8.d) e5_S with respect to ei_S for latent heat flux. Blue circles indicate improvement compared to ei_S (positive values of NIC_R in fig.8a,b, and negative values of N_{ubRMSD} in fig.8c,d) while red circles correspond to a degradation (negative values of NIC_R in fig.8a,b, and positive values of N_{ubRMSD} in fig.8c,d). The~~

~~four maps of figure 8 are largely dominated by large blue circles, therefore with dominant improvements.~~

3.2. *Evaluations using satellite derived estimates*

Figure ~~9~~8 illustrates the comparison between ESA CCI SSM_v4 and soil moisture from ISBA second layer of soil over 2010-2016. Fig.8~~9~~a shows seasonal correlations on volumetric time-series and fig.8~~9.~~b on anomaly time-series. Scores for ISBA LSM forced by ERA-Interim (ei_S) are in blue, ERA-5 but with precipitation from ERA-Interim (e5ei_S) green and ERA-5 (e5_S) in red. From fig8~~9~~.a one can appreciate the added value of using ERA-5 atmospheric forcing particularly from April~~s~~ to September. It is also interesting to notice that when using all ERA-5 atmospheric fields except for the precipitations, a similar added value is noticeable suggesting that all improvements from ERA-5 do not only come from precipitation. However when evaluating the short-term variability of soil moisture (i.e. removing the seasonal effect), it is really ERA-5 that provides the best results. Correlation on volumetric (anomaly) time-series for all grid points put together over 2010-2016 are 0.668 (0.464), 0.682 (0.468) and 0.689 (0.490) for ei_S, e5ei_S and e5_S, respectively. Additionally to the global seasonal scores, fig.~~9~~8c and fig.~~9~~8d present maps of correlations differences between soil moisture from e5_S and ei_S on volumetric time-series and anomaly time-series, respectively. Grey areas represent areas that were flagged out for elevation greater than 1500 m above sea level. As visible on fig.~~9~~8c and fig.8~~9~~d the use of ERA-5, mainly leads to improvements all over the considered domain. Focusing on correlation differences, (R_{e5}-R_{ei}) on volumetric (anomaly) time-series, 68% (77%) of the values are positives (indicating an improvement from e5) with median values of 4.5% (4.11%) and include values up to 40% (45%). It shows the added value of using ERA-5 to force ISBA LSM compared to ERA-Interim.

~~Figure 10 illustrates seasonal scores between ISBA LSM forced by either ERA-Interim (ei_S in blue) ERA-5 but ERA-Interim precipitation (e5ei in green) or ERA-5 (e5_S in red) and the three lasting dataset; (fig10.a, fig10.b) evapotranspiration estimates from the GLEAM project over 2010-2016, (fig10.c, fig10.d) upscaled GPP from the FLUXCOM project over 2010-2013 and (fig10.e, fig10.f) LAI estimates from the Copernicus GLS project over 2010-2016.~~Figure 9 illustrates seasonal scores between ISBA LSM forced by either ERA-Interim (ei_S in blue) ERA-5 but ERA-Interim precipitation (e5ei in green) or ERA-5 (e5_S in red) for; (fig9.a, fig9.b) evapotranspiration estimates from the GLEAM project over 2010-2016, (fig9.c, fig9.d) upscaled GPP from the FLUXCOM project over 2010-2013 and (fig9.e, fig9.f) LAI estimates from the Copernicus GLS project over 2010-2016. Left column (fig9.a, c and e) are for RMSDs and right column (fig9.b, d, e) for correlations. ~~Left column (fig10.a, c and e) are for RMSD and right column (fig8.b, d, e) for correlations.~~ For evapotranspiration and to a lesser extend GPP, one can notice a decrease in RMSD when using ERA-5 atmospheric reanalysis compared to ERA-Interim atmospheric reanalysis. However it fails at improving LAI. Considering evapotranspiration, correlation (RMSD) values for all grid points put together over 2010-2016 are 0.786 (0.927 $kg.m^{-2}.d^{-1}$), 0.778 (0.917 $kg.m^{-2}.d^{-1}$) and 0.795 (0.889 $kg.m^{-2}.d^{-1}$) for ei_S, e5ei_S and e5_S, respectively. They are 0.726 (2.429 $kg.m^{-2}.d^{-1}$), 0.733 (2.167 $kg.m^{-2}.d^{-1}$) and 0.734 (2.227 $kg.m^{-2}.d^{-1}$) for GPP and 0.715 (1.050 $m^2.m^{-2}$), 0.710 (1.026 $m^2.m^{-2}$) and 0.697 (1.079 $m^2.m^{-2}$) for LAI.

Improvements (in red) and degradations (in blue) from the use of ERA-5 in the ISBA LSM with respect to ERA-Interim for evapotranspiration, Gross Primary Production and Leaf Area Index are illustrated by figure 1~~1~~0 (respectively from top to bottom). Fig.10~~1~~a, c and e ~~follow eq.(7) (N_{RMSD})~~show RMSD differences while Fig.1~~1~~0b, d and f ~~follow Eq.(3) (NIC_R)~~show R differences. Both ~~N_{RMSD} and NIC_R~~ differences on RMSDs and R values suggest an improvement from the use of ERA-5 as the two figures are mainly dominated by red colors, they represent 56% and 53% of the domain, respectively for evapotranspiration (fig.10~~1~~a, b), 60% and 69% for GPP (fig.1~~1~~0c, d) but only 47% and 44% for LAI (fig.1~~1~~0e, f). ~~of clarity.sackOn figure 11, values between -12.5 and 12.5 for N_{RMSD}, -25 and 25 NIC_R are not represented for~~

4. Discussion and conclusions

This study assesses the ability of ECMWF recently released ERA-5 atmospheric reanalysis to force the ISBA LSM with respect to ERA-Interim reanalysis over North America for 2010-2016. The results presented above using the three atmospheric reanalysis data set (ERA-Interim -ei_S-, ERA-5 but with precipitation from ERA-Interim -e5ei_S- and ERA-5 -e5_S- all meteorological variables) to force the ISBA LSM provide two important insights ; (i) firstly the use of ERA-5 leads to significant improvements in the representation of the LSVs linked to the terrestrial water cycle assessed in this study (surface soil moisture, river discharges, snow depth and turbulent fluxes) but failed impacting LSVs linked to the vegetation cycle (carbon uptake and LAI). Even when they are are small, improvements are systematic when using ERA-5. (ii) Secondly if most of the improvements seems to come from a better representation of the precipitation in ERA-5, the e5ei_S experiment also present improvements with respect to the ei_S experiment suggests that the other meteorological forcing from ERA-5 are better represented too. It is however acknowledged that the use of 3-hourly ERA-Interim liquid and solid precipitations re-scaled at an hourly time step in ERA-5 might have sometimes led to inconsistent configurations (e.g., precipitations while having a very strong net radiation).

~~able to sequentially assimilate satellite derived estimates of surface soil moisture and LAI.)LDAS-Mondea Land Data Assimilation System (Albergel et al., 2017, 2018 (in prep.) recently presented~~ ERA-5 has a great potential to further improve the representation of land surface variables if used to force offline LDAS. In the past recent years, several LDAS have emerged at different spatial scales, (i) regional like the Coupled Land Vegetation LDAS (CLVLDAS, Sawada and Koike, 2014, Sawada et al., 2015), the Famine Early Warning Systems Network (FEWSNET) LDAS (FLDAS, McNally et al., 2017), (ii) continental like the North American LDAS (NLDAS, Mitchell et al., 2004; Xia etal., 2012), the National Climate Assessment LDAS (NCA-LDAS Kumar et al., 2018) as well as at (iii) global scale like the Global Land Data assimilation (GLDAS, Rodell et al., 2004) and more recently LDAS-Monde (Albergel et al., 2017, 2018 in prep). LDAS-Monde is a global capacity system able to sequentially assimilate satellite derived estimates of surface soil moisture and LAI. ~~They~~Albergel et al. (2017) found that the main improvements of their analysis (i.e. with assimilation) when compared to an open-loop experiment (simple model run) were linked to vegetation variables and the assimilation of vegetation estimates. They have also proposed further advances on a better use of satellite-based microwave data in the assimilation system. Having LDAS-Monde analysis forced by ERA-5 atmospheric forcing should both combined the strengths of an improved atmospheric reanalysis on the terrestrial water cycle and of the assimilation of satellite derived products on the vegetation cycle. Effort will now be concentrate on the use of ERA-5 and strengthening LDAS-Monde through the direct assimilation of satellite-based soil moisture and vegetation properties from microwave remote sensing. It will enable fostering links with potential applications like climate reanalysis of the LSVs as well as going from a monitoring system of the LSVs and extreme events (like agricultural drought) to a forecasting system. Preliminary results suggest that a LSVs forecast initialized by an analysis is more robust than one initialized by a simple model run (Albergel et al., 2018, in prep). Preliminary tests over Europe also indicate similar benefits from the use of ERA-5 (not shown). When the whole ERA-5 period will be available (1979-present), in addition with the availability of the ERA-5 10-member Ensemble of Data Assimilation (at lower spatial and temporal resolution though), it will be possible to develop a global long term ensemble of land surface variables reanalysis forced by high quality atmospheric data. It will make it possible providing uncertainties in the representation of the atmospheric forcing, while land surface variables may require special considerations and perturbation methods. Capturing those uncertainties coming from the simplifications and assumptions in the LSM is of paramount interest for many applications from monitoring to forecasting.

Acknowledgments-*Results where Generated using Copernicus Climate Change Service Information 2017. E. Dutra work was supported by the Portuguese Science Foundation (FCT) under project IF/00817/2015.*

References

Albergel, C., Calvet, J.-C., de Rosnay, P., Balsamo, G., Wagner, W., Hasenauer, S., Naeimi, V., Martin, E., Bazile, E., Bouyssel, F., and Mahfouf, J.-F.: Cross-evaluation of modelled and remotely sensed surface soil moisture with in situ data in southwestern France, Hydrol. Earth Syst. Sci., 14, 2177-2191, https://doi.org/10.5194/hess-14-2177-2010, 2010.

Albergel, C., Dorigo, W., Balsamo, G., Muñoz-Sabater, J., de Rosnay, P., L. Isaksen, Brocca, L., R. de Jeu and Wagner, W.: Monitoring multi-decadal satellite earth observation of soil moisture products through land surface reanalyses, Remote Sens. Environ., 138, 77–89, https://doi.org/10.1016/j.rse.2013.07.009, 2013a

Albergel, C., Dorigo, W., Reichle, R. H., Balsamo, G., de Rosnay, P., Munoz-Sabater, J., Isaksen, L., de Jeu, R., and Wagner, W.: Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing, J. Hydrometeorol., 14, 1259–1277, https://doi.org/10.1175/JHM-D-12-0161.1, 2013b

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.

Albergel, C., S. Munier, A. Bocher, C. Draper, D. J. Leroux, A. L. Barbu, J.-C. Calvet: LDAS-Monde global capacity integration of satellite derived observations applied over North America: assessment, limitations and perspectives. to be sumitted to Remote Sensing, Special Issue "Assimilation of Remote Sensing Data into Earth System Models", 2018.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19, 389-407, https://doi.org/10.5194/hess-19-389-2015, 2015.

Barbu, A. L., Calvet, J.-C., Mahfouf, J.-F., and Lafont, S.: Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France, Hydrology and Earth System Sciences, 18, 173–192, doi:10.5194/hess-18-173-2014, 2014.

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmared, H., Pacholczyk, P., and Smetse, B.: GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products, Part 1: Principles of development and production, Remote Sens. Environ., 137, 299–309, 2013.

Berrisford, P., D. P. Dee, K. Fielding, M. Fuentes, P. Kallberg, S. Kobayashi, and S. M. Uppala, 2009: The ERA-Interim archive. ERA Rep. 1, 16 pp. [Available online at http://www.ecmwf.int/publications/library/ecpublications/_pdf/era/era_report_series/RS_1.pdf.]

Bell, J.E., M.A. Palecki, W.G. Collins, J.H. Lawrimore, R.D. Leeper, M.E. Hall, J. Kochendorfer, T.P. Meyers, T. Wilson, B. Baker, and H.J. Diamond. 2013. U.S. Climate Reference Network soil moisture and temperature observatons. J. Hydrometeorol. doi:10.1175/JHM-D-12-0146.1.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate, Science, 329, 834–838, https://doi.org/10.1126/science.1184984, 2010.

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M., and Swenson, S. C.: Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, J. Geophys. Res., 116, G02014,

https://doi.org/10.1029/2010JG001593, 2011.

Boone, A. and Etchevers, P.: An intercomparison of three snow schemes of varying complexity coupled to the same land-surface model: local scale evaluation at an Alpine site, J. Hydrometeorol., 2, 374–394, 2001.

Boone, A., Masson, V., Meyers, T., and Noilhan, J.: The influence of the inclusion of soil freezing on simulations by a soil vegetation-atmosphere transfer scheme, J. Appl. Meteorol., 39, 1544–1569, 2000.

Boussetta S., G. Balsamo, E. Dutra, A. Beljaars and C. Albergel: Assimilation of surface albedo and vegetation states from satellite observations and their impact on numerical weather prediction, Remote Sensing of Environment, Vol 163, pp 111-126, 2015 ; doi: 10.1016/j.rse..2015.03.009.

Brooks, R. H. and Corey, A. T.: Properties of porous media affecting fluid flow, J. Irrig. Drain. Div. Am. Soc. Civ. Eng., 17, 187–208, 1966.

 20

Calvet, J.-C., Noilhan, J., Roujean, J.-L., Bessemoulin, P., Cabelguenne, M., Olioso, A., and Wigneron, J.-P.: An interactive vegetation SVAT model tested against data from six contrasting sites, Agr. Forest Meteorol., 92, 73–95, 1998.

Calvet, J.-C., Rivalland, V., Picon-Cochard, C., and Guehl, J.-M.: Modelling forest transpiration and CO2 fluxes – response to soil moisture stress, Agr. Forest Meteorol., 124, 143–156, https://doi.org/10.1016/j.agrformet.2004.01.007, 2004.

Canal, N., Calvet, J.-C., Decharme, B., Carrer, D., Lafont, S., and Pigeon, G.: Evaluation of root water uptake in the ISBA-A-gs land surface model using agricultural yield statistics over France, Hydrol. Earth Syst. Sci., 18, 4979–4999, https://doi.org/10.5194/hess-18-4979-2014, 2014.

Decharme, B., Boone, A., Delire, C., and Noilhan, J.: Local evaluation of the Interaction between soil biosphere atmosphere soil multilayer diffusion scheme using four pedotransfer functions, J. Geophys. Res., 116, D20126, https://doi.org/10.1029/2011JD016002, 2011.

Decharme, B., Alkama, R., Douville, H., Becker, M., and Cazenave, A.: Global evaluation of the ISBA-TRIP continental hydrologic system, Part 2: Uncertainties in river routing simulation related to flow velocity and groundwater storage, J. Hydrometeorol., 11, 601–617, 2010.

Decharme, B., Alkama, R., Papa, F., Faroux, S., Douville, H., and Prigent, C.: Global offline evaluation of the ISBA-TRIP flood model, Clim. Dynam., 38, 1389–1412, https://doi.org/10.1007/s00382-011-1054-9, 2012.

Decharme, B., Martin, E., and Faroux, S.: Reconciling soil thermal and hydrological lower boundary conditions in land surface models, J. Geophys. Res.-Atmos., 118, 7819–7834, https://doi.org/10.1002/jgrd.50631, 2013.

Decharme, B., Brun, E., Boone, A., Delire, C., Le Moigne, P., and Morin, S.: Impacts of snow and organic soils parameteriation on northern Eurasian soil temperature profiles simulated by the ISBA land surface model, The Cryosphere, 10, 853–877, https://doi.org/10.5194/tc-10-853-2016, 2016.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P.. Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Dirmeyer, P. A., Gao, X., and Oki, T.: The second Global Soil Wetness Project – Science and implementation plan, IGPO Int. GEWEX Project Office Publ. Series 37, Global Energy and Water Cycle Exp. (GEWEX) Proj. Off., Silver Spring, MD, 65 pp., 2002.

Dirmeyer, P. A.: A history and review of the Global SoilWetness Project (GSWP), J. Hydrometeorol., 12, 729–749,doi:10.1175/JHM-D-10-05010.1, 2011.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki 2006: The Second Global Soil Wetness Project (GSWP-2): Multi-model analysis and implications for our perception of the land surface, Bull. Am. Meteorol. Soc., 87, 1381–1397, doi:10.1175/BAMS-87-10-1381, 2006.

Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., C. Albergel, Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sens. Environ., 162, 380–395, https://doi.org/10.1016/j.rse.2014.07.023, 2015.

Dorigo, W., Wagner, W., Albergel, C. Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., William Lahoz g, Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.:ESA CCI soil moisture for improved Earth system understanding: state-of-the art and future directions, Remote Sens. Environ., 201, RSE-10331, https://doi.org/10.1016/j.rse.2017.07.001, 2017.

Draper, C., Mahfouf, J.-F., Calvet, J.-C., Martin, E., and Wagner, W.: Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France, Hydrology and Earth System Sciences, 15, 3829–3841, doi:10.5194/hess-15-3829-2011, 2011.

Gelaro, R., W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A.M. da Silva, W. Gu, G. Kim, R. Koster, R. Lucchesi, D. Merkova, J.E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, and B. Zhao, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate,* 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1

Gibelin, A.-L., Calvet, J.-C., Roujean, J.-L., Jarlan, L., and Los, S. O.: Ability of the land surface model ISBA-A-gs to simulate leaf area index at global scale: comparison with satellite products, J. Geophys. Res., 111, 1–16, https://doi.org/10.1029/2005JD006691, 2006.

Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., and Seneviratne, S. I.: Global assessment of trends in wetting and drying over land, Nat. Geosci., 7, 716–721, https://doi.org/10.1038/ngeo2247, 2014.

Guillod, B. P., Orlowsky, B., Miralles, D. G., Teuling, A. J., and Seneviratne, S. I.: Reconciling spatial and temporal soil moisture effects on afternoon rainfall, Nat. Commun., 6, 6443, https://doi.org/10.1038/ncomms7443, 2015.

Entekhabi, D., R. H. Reichle, R. D. Koster, and W. T. Crow, 2010: Performance metrics for soil moisture retrieval and application requirements. J. Hydrometeor., 11, 832–840.

Hersbach H. et al., 2016: “ERA-5 reanalysis is in production”. ECMWF newsletter, number 147, Spring 2016, p.7.

Jackson, R. B., Canadell, J., Ehleringer, J. R., Mooney, H. A., Sala, O. E., and Schulze, E. D.: A global analysis of root distributions for terrestrial biomes, Oecologia, 108, 389–411, https://doi.org/10.1007/BF00333714, 1996.

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., and Fawcett, P. J.: Terrestrial water fluxes dominated by transpiration, Nature, 496, 347–350, https://doi.org/10.1038/nature11983, 2013.

Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences, 6, 2001–2013, https://doi.org/10.5194/bg-6-2001-2009, 2009.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., de Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, Nature, 467, 951–954, https://doi.org/10.1038/nature09396, 2010.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M.,Vaccari, F., and Williams, C.: Global patterns of land–atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res., 116, G00J07, https://doi.org/10.1029/2010JG001566, 2011.

Jung, M., M. Reichstein, Christopher R. Schwalm, Chris Huntingford, Stephen Sitch, Anders Ahlström, Almut Arneth, Gustau Camps-Valls, Philippe Ciais, Pierre Friedlingstein, F. Gans, Kazuhito Ichii, Atul K. Jain, Etsushi Kato, Dario Papale, Ben Poulter, Botond Raduly, C. Rödenbeck, Gianluca Tramontana, Nicolas Viovy, Ying-Ping Wang, U. Weber, S. Zaehle and Ning Zeng, Compensatory water effects link yearly global land CO2 sink changes to temperature, Nature, 541, 516-520, doi:10.1038/nature20780, 2017

Koster, R., Mahanama, S., Yamada, T., Balsamo, G., Boisserie, M., Dirmeyer, P., Doblas-Reyes, F., Gordon, T., Guo, Z., Jeong, J.-H., Lawrence, D., Li, Z., Luo, L., Malyshev, S., Merryfield, W., Seneviratne, S. I., Stanelle, T., van den Hurk, B., Vitart, F., and Wood, E. F.: The contribution of land surface initialization to sub-seasonal forecast skill: First resultsfrom the GLACE-2 Project, Geophys. Res. Lett., 37, L02402, doi:10.1029/2009GL041677, 2009a.

Koster, R., Z. Guo, R. Yang, P. Dirmeyer, K. Mitchell, and M. Puma: On the nature of soil moisture in land surface models. J. Climate, 22 (16), 4322{4335, doi:10.1175/ 2009JCLI2832.1., 2009b.

Koster, R., Mahanama, S. P. P., Yamada, T. J., Balsamo, G., Berg, A. A., Boisserie, M., Dirmeyer, P. A., Doblas-Reyes, F. J., Drewitt, G., Gordon, C. T., Guo, Z., Jeong, J.-H., Lee, W.-S., Li, Z., Luo, L., Malyshev, S., Merryfield, W. J., Seneviratne, S. I., Stanelle, T., van den Hurk, B. J. J. M., Vitart, F., and Wood, E. F.: The second phase of the global land-atmosphere coupling experiment: soil moisture contributions to sub-seasonal forecast skill, J. Hydrometeorol., 12, 805–822, doi:10.1175/2011JHM1365.1, 2011.

Kumar, S., R. H. Reichle, R. D. Koster, W. T. Crow, and C. Peters-Lidard. 2009. "Role of Subsurface Physics in the Assimilation of Surface Soil Moisture Observations." *J. Hydrometeor*, 10 (6): 1534-1547 [10.1175/2009JHM1134.1]

Kumar, S.V., M. Jasinski, D. Mocko, M. Rodell, J. Borak, B. Li, H. Kato Beaudoing, and C.D. Peters-Lidard: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment. J. Hydrometeor., 0, https://doi.org/10.1175/JHM-D-17-0125.1

Liu, Y. Y., R. M. Parinussa, W. A. Dorigo, R. A. M. De Jeu, W. Wagner, A. I. J. M. van Dijk, M. F. McCabe, and J. P. Evans, 2011: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrol. Earth Syst. Sci., 15, 425–436, doi:10.5194/hess-15-425-2011.

Liu, Y. Y., W. A. Dorigo, R. M. Parinussa, R. A. M. De Jeu, W. Wagner, M. F. McCabe, J. P. Evans, and A. I. J. M. van Dijk, 2012: Trend-preserving blending of passive and active microwave soil moisture retrievals. Remote Sens. Environ., 123, 280–297, doi:10.1016/j.rse.2012.03.014.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Mätzler, C., and Standley, A.: Relief effects for passive microwave remote sensing. International Journal of Remote Sensing, 21, 12, 2403–2412, doi:10.1080/01431160050030538, 2000.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geosci. Model Dev., 6, 929–960, https://doi.org/10.5194/gmd-6-929-2013, 2013.

McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., Funk, C., Peters-Lidard, C. D. and Verdin, J. P.: A land data assimilation system for sub-Saharan Africa food and water security applications. Scientific Data, 4, 170012, :10.1038/sdata.2017.12, 2017.

Menne, M.J., I. Durre, R.S. Vose, B.E. Gleason, and T.G. Houston: An overview of the Global Historical Climatology Network-Daily Database. Journal of Atmospheric and Oceanic Technology, 29, 897-910, doi.10.1175/JTECH-D-11-00103.1, 2012a.

Menne, Matthew J., Imke Durre, Bryant Korzeniewski, Shelley McNeal, Kristy Thomas, Xungang Yin, Steven Anthony, Ron Ray, Russell S. Vose, Byron E.Gleason, and Tamara G. Houston: Global Historical Climatology Network - Daily (GHCN-Daily), Version 3. [snow depth]. NOAA National Climatic Data Center. doi:10.7289/V5D21VHZ, 2012b [accessed November 2017].

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci., 15, 453–469, https://doi.org/10.5194/hess-15-453-2011, 2011.

 25

Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C., and de Arellano, J. V.-G.: Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation, Nat. Geosci., 7, 345–349, https://doi.org/10.1038/NGEO2141, 2014a.

Miralles, D. G., van den Berg, M. J., Gash, J. H., Parinussa, R. M., de Jeu, R. A. M., Beck, H. E., Holmes, T. R. H., Jiménez, C., Verhoest, N. E. C., Dorigo, W. A., Teuling, A. J., and Dolman, A. J.: El Niño-La Niña cycle and recent trends in continental evaporation, Nat. Clim. Change, 4, 122–126, https://doi.org/10.1038/NCLIMATE2068, 2014b.

Mitchell, K. E., et al. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, J. Geophys. Res.,109, D07S90, 2004. doi:10.1029/2003JD003823

Noilhan, J. and Mahfouf, J.-F.: The ISBA land surface parameterisation scheme, Global Planet. Change, 13, 145–159, 1996.

Oki, T. and Sud, Y. C.: Design of Total Runoff Integrating Pathways (TRIP), a global river chanel network, Earth Interact., 2, 1–36, 1998.

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Toure, A.: Assessment and enhancement of MERRA land surface hydrology estimates, J. Climate, 24, 6322–6338, doi:10.1175/JCLI-D-10-05033.1, 2011.

Reichle, R.H., C.S. Draper, Q. Liu, M. Girotto, S.P. Mahanama, R.D. Koster, and G.J. De Lannoy, 2017: Assessment of MERRA-2 Land Surface Hydrology Estimates. *J. Climate,* 30, 2937–2960, https://doi.org/10.1175/JCLI-D-16-0720.1

Richards, L. A.: Capillary conduction of liquids in porous mediums, Physics, 1, 318–333, 1931.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Julio, B., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: **MERRA– NASA's modern-era retrospective analysis for research and applications, J. Climate, 24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.**

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, The Global Land Data Assimilation System, Bull. Amer. Meteor. Soc., 85(3), 381–394, 2004.

Sawada, Y., T. Koike, and J. P. Walker, A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics, J. Geophys. Res.Atmos., 120, doi: 10.1002/2014JD022895, 2015.

Sawada, Y., and T. Koike, Simultaneous estimation of both hydrological and ecological parameters in an ecohydrological model by assimilating microwave signal, J. Geophys. Res. Atmos., 119, doi:10.1002/2014JD021536, 2014.

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., van Dijk, A., Sperna Weiland, F., Minvielle, M., Calvet, J.-C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W., and Weedon, G. P.: A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, Earth Syst. Sci. Data, 9, 389-413, https://doi.org/10.5194/essd-9-389-2017, 2017.

Slevin, D., Tett, S. F. B., Exbrayat, J.-F., Bloom, A. A., and Williams, M.: Global evaluation of gross primary productivity in the JULES land surface model v3.4.1, Geosci. Model Dev., 10, 2651–2670, https://doi.org/10.5194/gmd-10-2651-2017, 2017.

Stanski, H. R., L. J. Wilson, and W. R. Burrows, 1989: Survey of common verifications methods in meteorology. WMO World Weather Watch Tech. Rep. 8/WMO/TD-358, 114 pp.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., Kiely, G., Merbold, L., Serrano-Ortiz, P., Sickert, S., Wolf, S., and Papale, D.: Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, Biogeosciences, 13, 4291-4313, https://doi.org/10.5194/bg-13-4291-2016, 2016.

Szczypta, C., Calvet, J.-C., Maignan, F., Dorigo, W., Baret, F., and Ciais, P.: Suitability of modelled and remotely sensed essential climate variables for monitoring Euro-Mediterranean droughts, Geosci. Model Dev., 7, 931–946, https://doi.org/10.5194/gmd-7-931-2014, 2014.

van der Schrier, G., Barichivich, J., Briffa, K. R., and Jones, P. D.: A scPDSI-based global data set of dry and wet spells for 1901–2009, J. Geophys. Res.-Atmos., 118, 4025–4048, 2013.

Vergnes, J.-P. and Decharme, B.: A simple groundwater scheme in the TRIP river routing model: global off-line evaluation against GRACE terrestrial water storage estimates and observed river discharges, Hydrol. Earth Syst. Sci., 16, 3889–3908, https://doi.org/10.5194/hess-16-3889-2012, 2012.

Vergnes, J.-P., Decharme, B., and Habets, F.: Introduction of groundwater capillary rises using subgrid spatial variability of topography into the ISBA land surface model, J. Geophys. Res.-Atmos., 119, 11065–11086, https://doi.org/10.1002/2014JD021573, 2014.

Wagner, W., W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, and M. Ertl, 2012: Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture. Proc. XXII ISPRS Congress, Melbourne, Australia, ISPRS, 315–321.

Voldoire, A., Decharme, B., Pianezze, J., Lebeaupin Brossier, C., Sevault, F., Seyfried, L., Garnier, V., Bielli, S., Valcke, S., Alias, A., Accensi, M., Ardhuin, F., Bouin, M.-N., Ducrocq, V., Faroux, S., Giordani, H., Léger, F., Marsaleix, P., Rainaud, R., Redelsperger, J.-L., Richard, E., and Riette, S.: SURFEX v8.0 interface with OASIS3-MCT to couple atmosphere with hydrology, ocean, waves and sea-ice models, from coastal to global scales, Geosci. Model Dev., 10, 4207-4227, https://doi.org/10.5194/gmd-10-4207-2017, 2017.

Xia, Y., et al. 2012, Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, J. Geophys. Res., 117, D03109, doi:*10.1029/2011JD016048*, 2012.

Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H. S., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang, Y., Liu, Y. Y., Miralles, D. G., and Pan, M.: Multi-decadal trends in global terrestrial evapotranspiration and its components, Sci. Rep.-UK, 6, 19124, https://doi.org/10.1038/srep19124, 2016.

Tables and Illustrations

Table I : Evaluation datasets and associated metrics used in this study.

Datasets used for the evaluation	Source	Metrics associated
In situ measurements of soil moisture (USCRN, Bell et al., 2013)	https://www.ncdc.noaa.gov/crn	R (on both volumetric and anomaly time-series) ubRMSD
In situ measurements of streamflow (USGS)	https://nwis.waterdata.usgs.gov/nwis	Nash Efficiency (NSE), Normalized Information Contribution (NIC) based on NSE, Ratio of simulated and observed streamflow (Q)
In situ measurements of snow depth (GHCN, Menne et al., 2012a, b)	https://www.ncdc.noaa.gov/climate-monitoring/	R, ~~stde and MAE, NIC based on R, Normalized stde and MAE~~bias and ubRMSD
In situ measurements of sensible and latent heat fluxes (FLUXNET-2015)	http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/	R, ~~ub~~RMSD~~, NIC based on R, Normalized ubRMSD~~
Satellite derived surface soil moisture (ESA CCI SSM v4, Dorigo et a., 2015, 2017)	http://www.esa-soilmoisture-cci.org	R (on both volumetric and anomaly time-series)
Satellite derived Leaf Area Index (GEOV1, Baret et al., 2013)	http://land.copernicus.eu/global/	R and RMSD~~, NIC based on R, Normalized RMSD~~
Satellite-driven model estimates of land evapotranspiration (GLEAM, Martns et al., 2017)	http://www.gleam.eu	R and RMSD~~, NIC based on R, Normalized RMSD~~
Upscaled estimates of Gross Primary Production (GPP, Jung et al., 2017)	https://www.bgc-jenna.mpg.de/geodb/projects/Home.php	R and RMSD~~, NIC based on R, Normalized RMSD~~

Table II: Comparison of surface soil moisture with in situ observations for ei_S, e5ei_S and e5_S over 2010-2016 (April to September months are considered). Median correlations R (on volumetric and anomaly time series) and ubRMSD are given for the USCRN. Scores are given for significant correlations with p-values <0.05.

	Median R* on volumetric time series, 95 % Confidence Interval** (% of stations for which this configuration is the best)	Median R*** on anomalies time series, 95 % Confidence Interval** (% of stations for which this configuration is the best)	Median ubRMSD* (m^3m^{-3}), 95 % Confidence Interval** (% of stations for which this configuration is the best)
ei_S	0.66±0.02 (20 %)	0.53±0.02 (15 %)	0.052±0.003 (19 %)
e5ei_S	0.69±0.02 (20 %)	0.54±0.04 (10 %)	0.052±0.002 (24 %)
e5_S	0.71±0.02 (60 %)	0.58±0.03 (75 %)	0.050±0.003 (57 %)

*** only for stations presenting significant R values on volumetric time series (p-value<0.05): 110 stations**
**** 95% confidence interval of the median derived from a 10000 samples bootstrapping**
***** only for stations presenting significant R values on anomaly time series (p-value<0.05): 107 stations**

Table III: Comparison of snow depth with in situ measurements, median Bias, ubRMSD and R values are given for the three seasons affected by snow (SON, DJF, MAM) and for the whole period (All). SON, DJF and MAM stand for September-October-November, December-January-February and Mars-April-May, respectively.

		Median bias (cm)*, 95 % Confidence Interval** (% of stations for which this configuration is the best)	Median ubRMSD (cm)*, 95 % Confidence Interval** (% of stations for which this configuration is the best)	Median R*, 95 % Confidence Interval** (% of stations for which this configuration is the best)
ei_S	SON	-0.27±0.04 (13 %)	2.05±0.17 (13 %)	0.70±0.01 (21 %)
	DJF	-6.28±0.86 (11 %)	10.34±0.63 (17 %)	0.72± 0.01 (20 %)
	MAM	-1.90±0.33 (15 %)	7.82±0.79 (17 %)	0.65± 0.01 (18 %)
	All	-2.11±0.33 (11 %)	7.58±0.65 (14 %)	0.75± 0.01 (19 %)
e5ei_S	SON	-0.25±0.04 (12 %)	2.03±0.15 (10 %)	0.74± 0.01 (23 %)
	DJF	-4.84±0.80 (14 %)	9.98±0.50 (14 %)	0.75± 0.01 (21 %)
	MAM	-1.49±0.33 (14 %)	7.61±0.76 (13 %)	0.69±0.02 (22 %)
	All	-1.70±0.33 (14 %)	7.40±0.65 (14 %)	0.77± 0.01 (20 %)
e5_S	SON	-014±0.03(76 %)	1.83±0.14 (77 %)	0.79± 0.01 (56 %)
	DJF	-1.70±0.44 (75 %)	9.64±0.46 (69 %)	0.80± 0.01 (59 %)
	MAM	-0.57±0.22 (71 %)	7.43±0.79 (70 %)	0.76± 0.01 (60 %)
	All	-0.64±0.19 (75 %)	7.00±0.65 (72 %)	0.82± 0.01 (61 %)

*** only for stations presenting more than 80% of (daily) data; 1901 out of 2056 stations.**
**** 95% confidence interval of the median derived from a 10000 samples bootstrapping**

Table IV~~II~~: Comparison of sensible (H) and latent (LE) heat flux with in situ observations for ei_S, e5ei_S and e5_S. Median correlations (R) and median ~~ub~~RMSD are given for the fluxnet stations. Scores are given for significant correlations with p-values <0.05.

	H Median R*, 95 % Confidence Interval** (% of stations for which this configuration is the best)~~H Median R* (% of stations for which this configuration is the best)~~	H Median RMSD* W.m^{-2}, 95 % Confidence Interval** (% of stations for which this configuration is the best)~~H Median ubRMSD* W.m^{-2} (% of stations for which this configuration is the best)~~	LE Median R*, 95 % Confidence Interval** (% of stations for which this configuration is the best)~~LE Median R* (% of stations for which this configuration is the best)~~	LE Median RMSD* W.m^{-2}, 95 % Confidence Interval** (% of stations for which this configuration is the best)~~LE Median ubRMSD* W.m^{-2} (% of stations for which this configuration is the best)~~
ei_S~~ei_S~~	0.62±0.11 (8 %)~~0.62 (8 %)~~	39.58±3.71 (5 %)~~34.85 (0 %)~~	0.63±0.05 (8 %)~~0.63 (8 %)~~	39.00±5.38 (16 %)~~33.93 (14 %)~~
e5ei_S~~e5ei_S~~	0.62±0.11(27 %)~~0.62(27 %)~~	32.89±3.86 (27%)~~30.66 (22 %)~~	0.62±0.07 (11 %)~~0.62 (11 %)~~	37.12±4.37 (22 %)~~31.66 (24 %)~~
e5_S~~e5_S~~	0.65±0.11 (65 %)~~0.65 (65 %)~~	32.73±2.61 (68 %)~~30.38 (78 %)~~	0.70±0.04 (81 %)~~0.70 (81 %)~~	36.66±4.94 (62 %)~~30.98 (62 %)~~

*** only for stations presenting significant R values (p-value<0.05): 37 stations**
**** 95% confidence interval of the median derived from a 10000 samples bootstrapping**

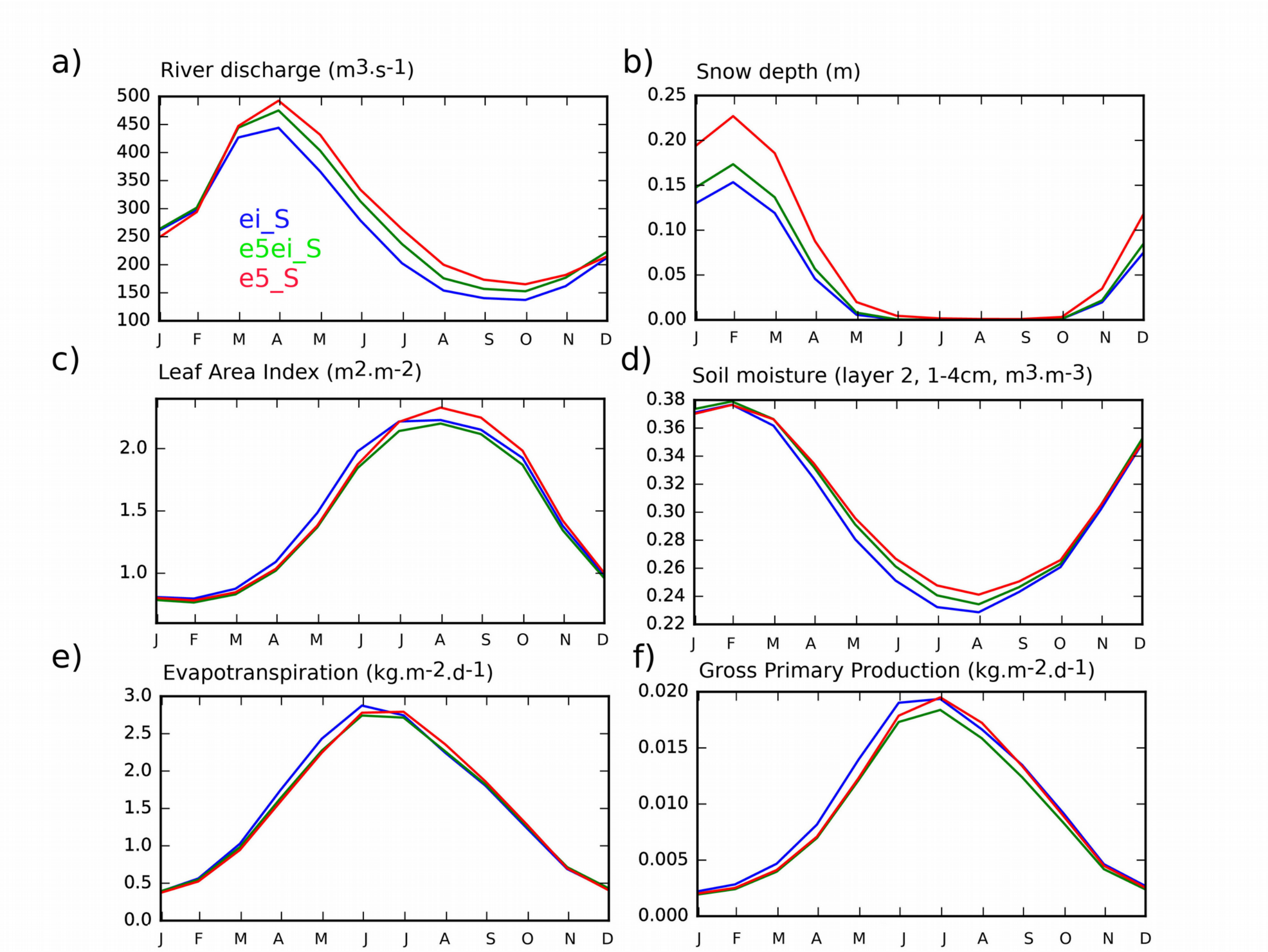

Illustration 1: Seasonal ~~time-series~~cycle of the 6 main land surface variables evaluated in this study over the whole domain for 2010-2016 :~~.~~ (a) river discharge, (b) snow depth, (c) leaf area index , (d) liquid soil moisture in the second layer of soil (1-4 cm depth), (e) evapotranspiration and (f) gross primary production. Land surface variables simulated with SURFEX forced by ERA-Interim (ei_S) are in blue, by ERA-5 (e5_S) with precipitation from ERA-Interim (e5ei_S) in blue and by ERA-5 in red.

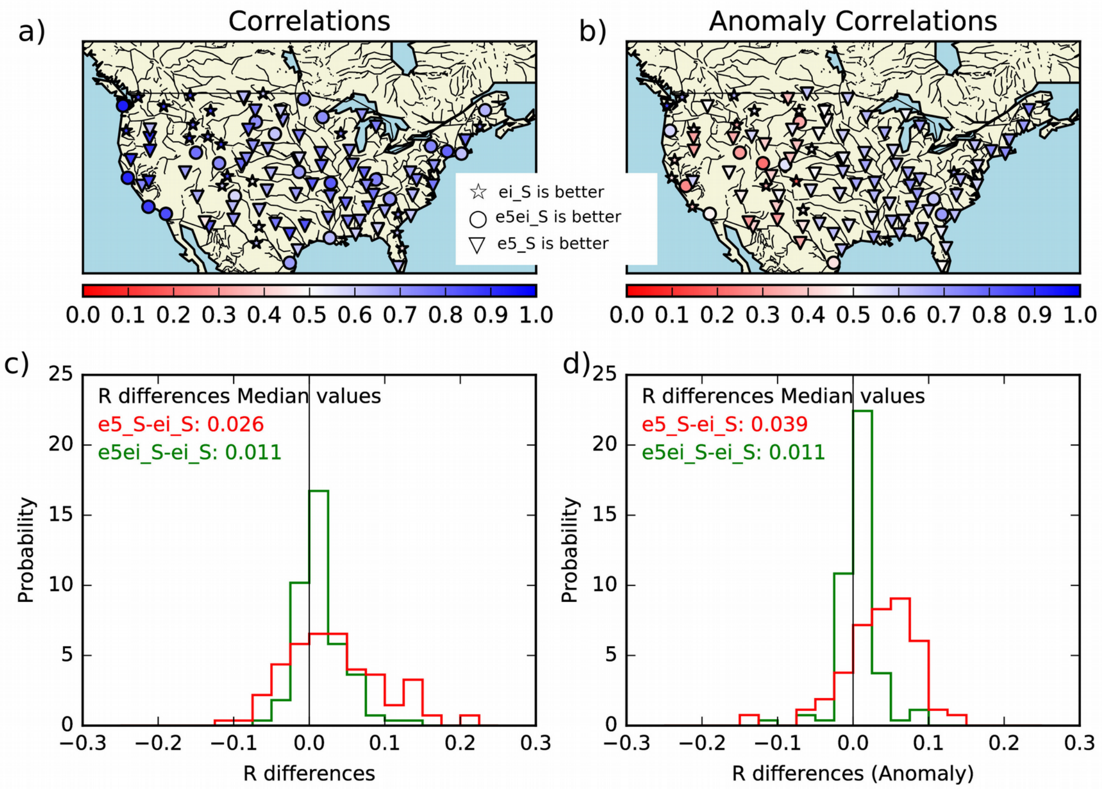

Figure 2 : Maps of correlation (R) on volumetric time-series (a) and anomaly time-series (b) between in situ measurements at 5 cm depth from the USCRN network and the ISBA Land Surface Model within the SURFEX modeling platform forced by either ERA-Interim (ei_S), ERA-5 with ERA-Interim precipitations (e5ei_S) and ERA-5 (e5_S). For each stations presenting significant R (p-values < 0.05) simulation that presents the better R values is represented. Stars symbols are when ei_S, presents the best value, circles when it e5ei_S and downward pointing triangles when it is e5_S. (c) Shows histograms of R differences on volumetric time-series, R(e5_S)-R(ei_S) in red and R(e5ei_S)-R(ei_S) in green, median values of the differences are reported, also. (d) Same as (c) for R values on anomaly time-series.

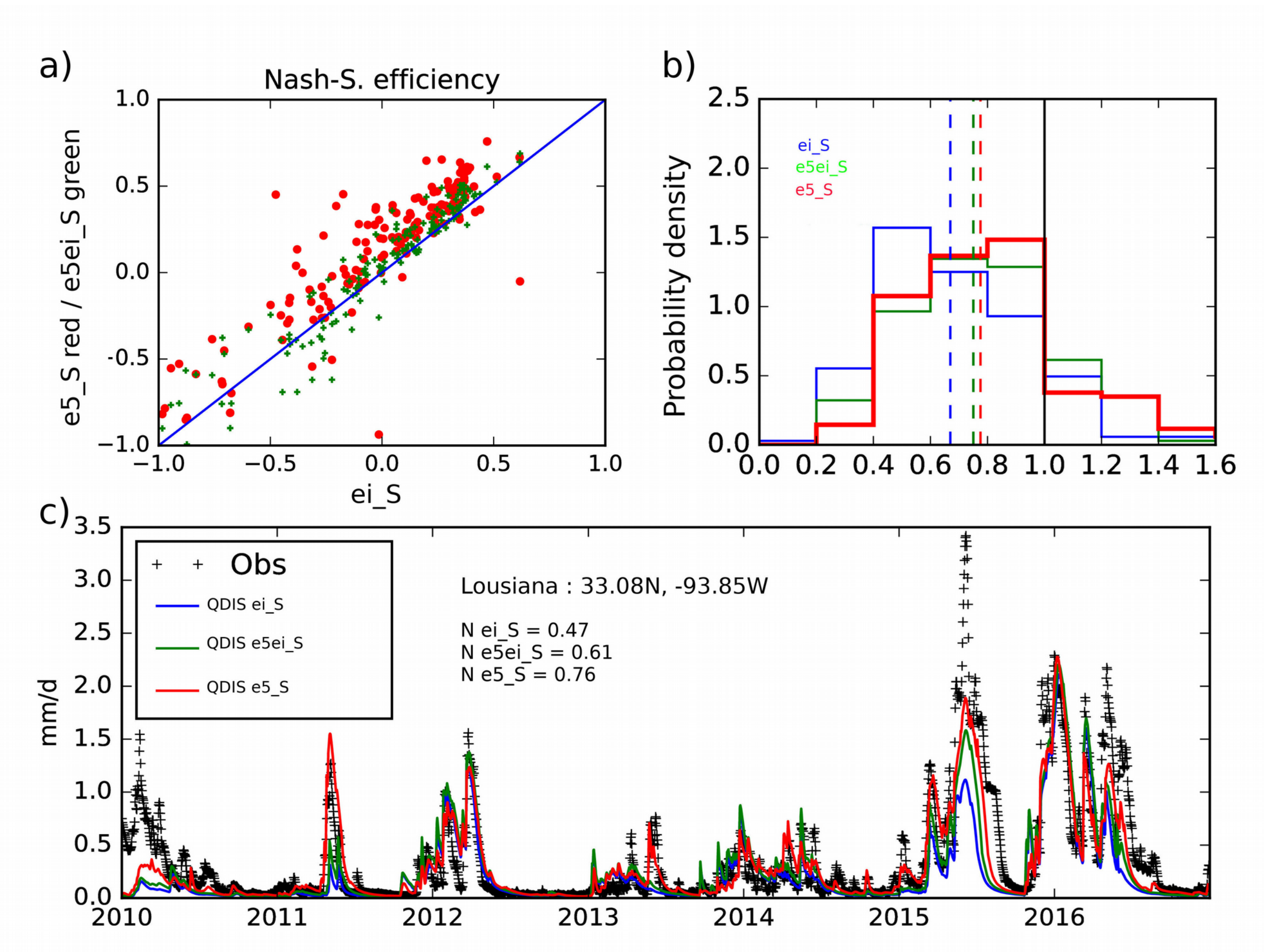

Figure 3: (a) Scatterplots of efficiency scores between in situ and simulated river discharges *Q*; efficiency scores for *Q* simulated with SURFEX forced either by ERA-5 but ERA-Interim precipitations (e5ei_S, green crosses) or ERA-5 (e5_S, red dots) function of efficiency scores for *Q* simulated using ERA-Interim (ei_S). (b) Histograms of river discharges ratio for ei_S (Qr_ei in blues), e5ei_S (Qr_e5ei in green) and e5_S (Qr_e5 in red). (c) Hydrograph for a river station in Lousiana (33.08°N, 1.52°W) representing scaled *Q* (using either observed or simulated drainage areas), in situ data (black crosses), simulated river discharges from ei_S (blue solid line), e5ei_S (green solid line) and e5_S (red solid line).

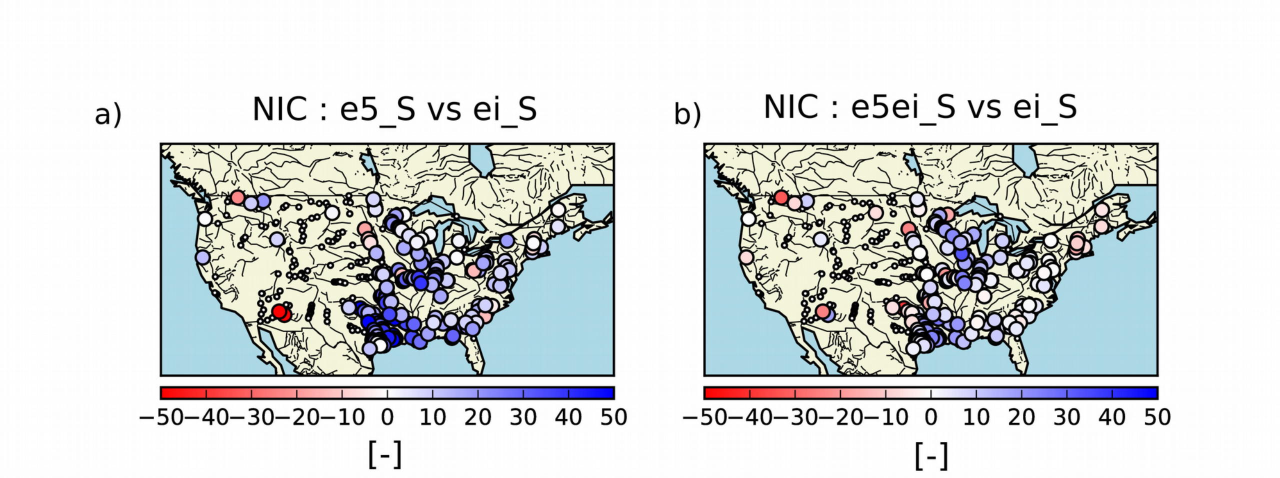

Figure 4: Normalized Information Contribution scores based on efficiency scores (NIC$_{NSE}$) (a) e5_S with respect to ei_S and (b) e5ei_S with respect to ei_S. Small dots represent station for which the benchmark experiment (ei_S) present efficiency scores smaller than -1, large circles when it presents values higher than -1. Positive values (blue large circles) suggest an improvement over ei_S, negative values (red large circles) a degradation. For sack of clarity, a factor 100 has been applied on NIC.

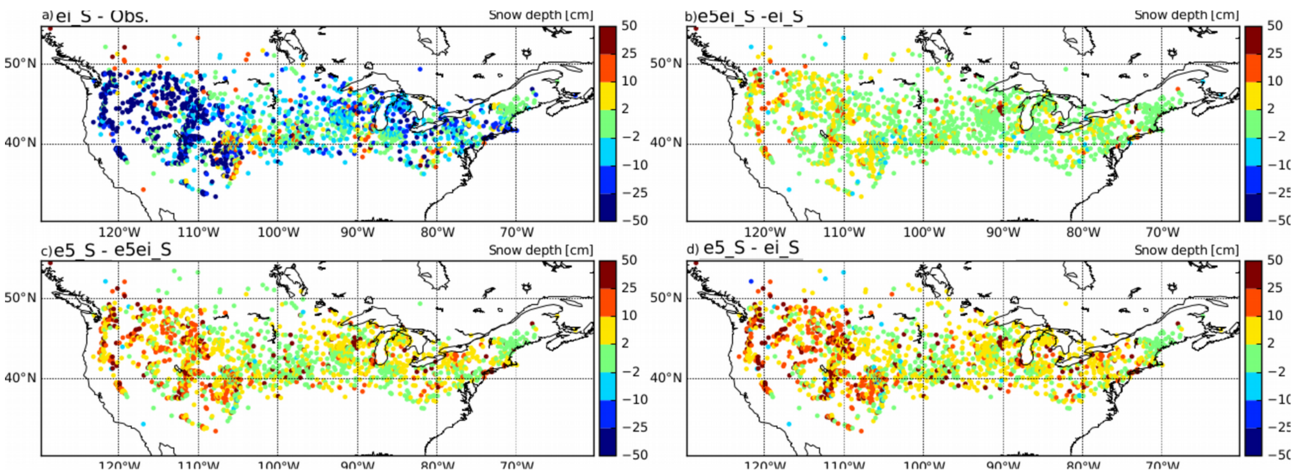

Figure 5: Mean snow depth bias for December-January-February in ei_S (a) and differences between e5ei_S and ei_S (b), e5_S and e5ei_S (c), e5_S and ei_5 (d).

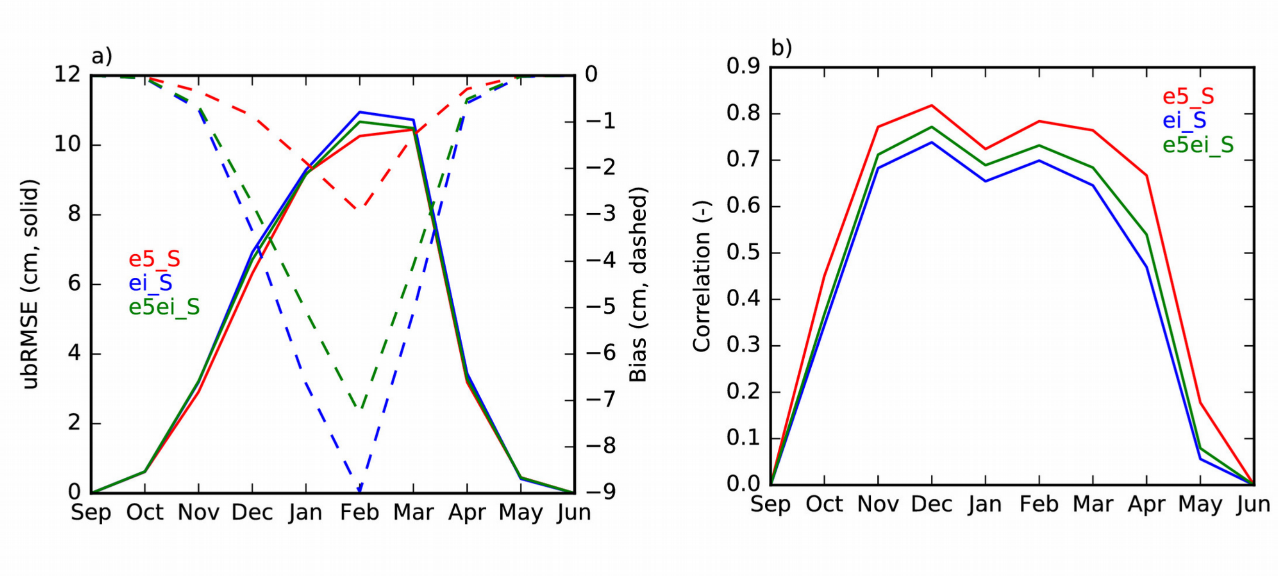

Figure 6: ~~Snow depth Normalized Information Contribution scores (eq. 3) in respect to ei_S for (a) the mean absolute error (MAE), (b) unbiased Root Mean Square Differences (ubRMSD) and (c) correlation. The error bars denote a 95% confidence interval of the mean derived from a 10000 samples bootstrapping.~~ **(a~~b~~) M~~hows the ms~~ean ~~annual~~seasonal cycle of the bias (~~solid~~dashed lines) and ubRMSD (~~dashed~~solid lines) averaged over all stations and (b)~~.~~ ~~–~~the mean seasonal cycle of the correlations for ei_S (in blue), e5ei_S (in green) and e5_S (in red).**

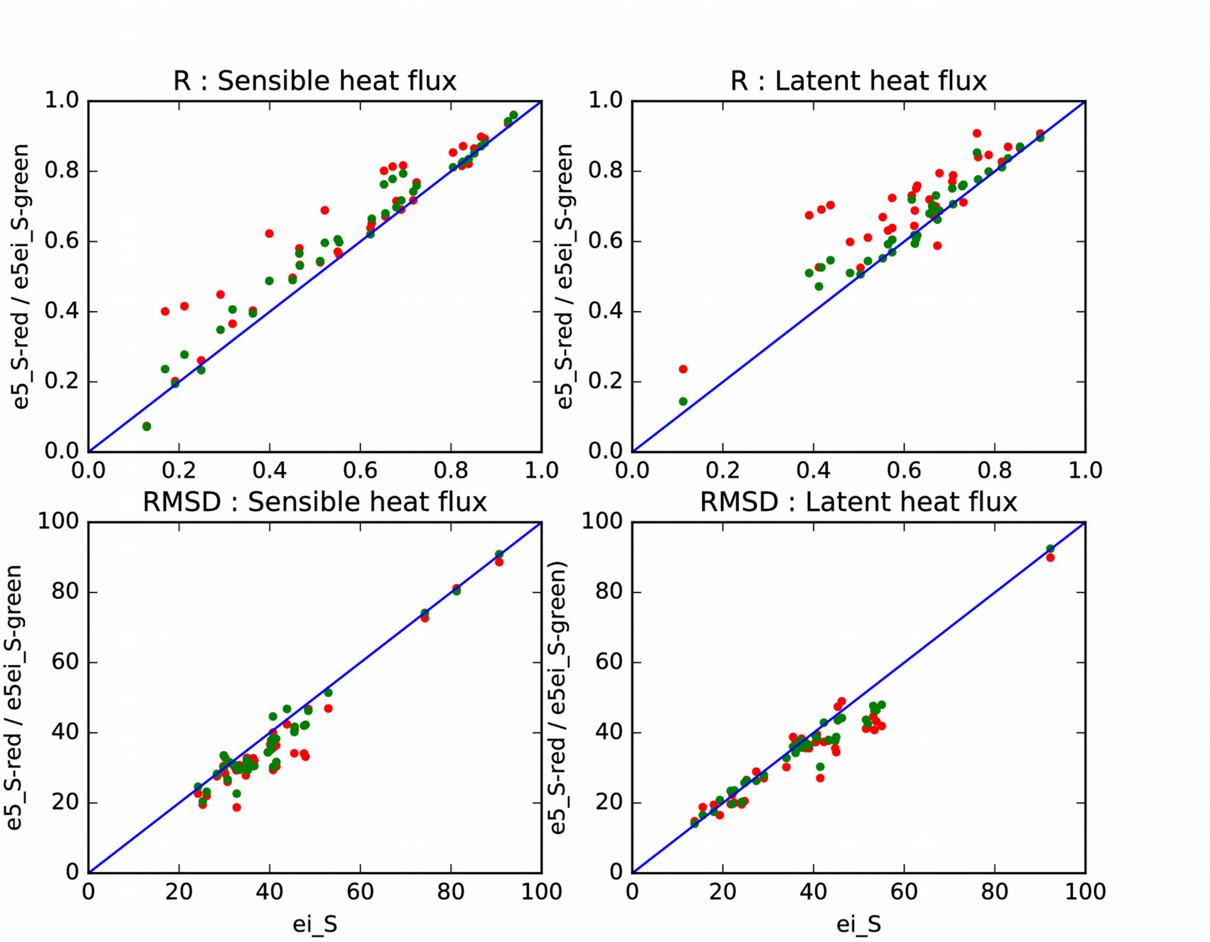

Figure 7: S~~s~~catterplots illustrating evaluation of ei_S, e5ei_S, e5_S against in situ measurements of sensible (a for correlation, c for ~~ub~~RMSD) and latent (b for correlation, d for ~~ub~~RMSD) heat flux. Scores for either e5ei_S (green dots) or e5_S (in red) are presented as function of those for ei_S.

~~**Figure 8: Normalized Information Contribution scores based on correlations values between in situ measurements from the fluxnet sites data and (a) e5_S with respect to ei_S for sensible heat flux and (b) e5_S with respect to ei_S for latent heat flux. Normalised ubRMSD values between in situ measurements from the fluxnet sites data and (c) e5_S with respect to ei_S for sensible heat flux and (d) e5_S with respect to ei_S for latent heat flux. Blue circles indicate improvement compared to ei_S (positive values of NIC_R in a,b, and negative values of N_{ubRMSD} in c,d) whereas red circles correspond to a degradation (negative values of NIC_R in a,b, and positive values of N_{ubRMSD} in c,d). For sack of clarity a factor 100 has been applied on NIC_R.**~~

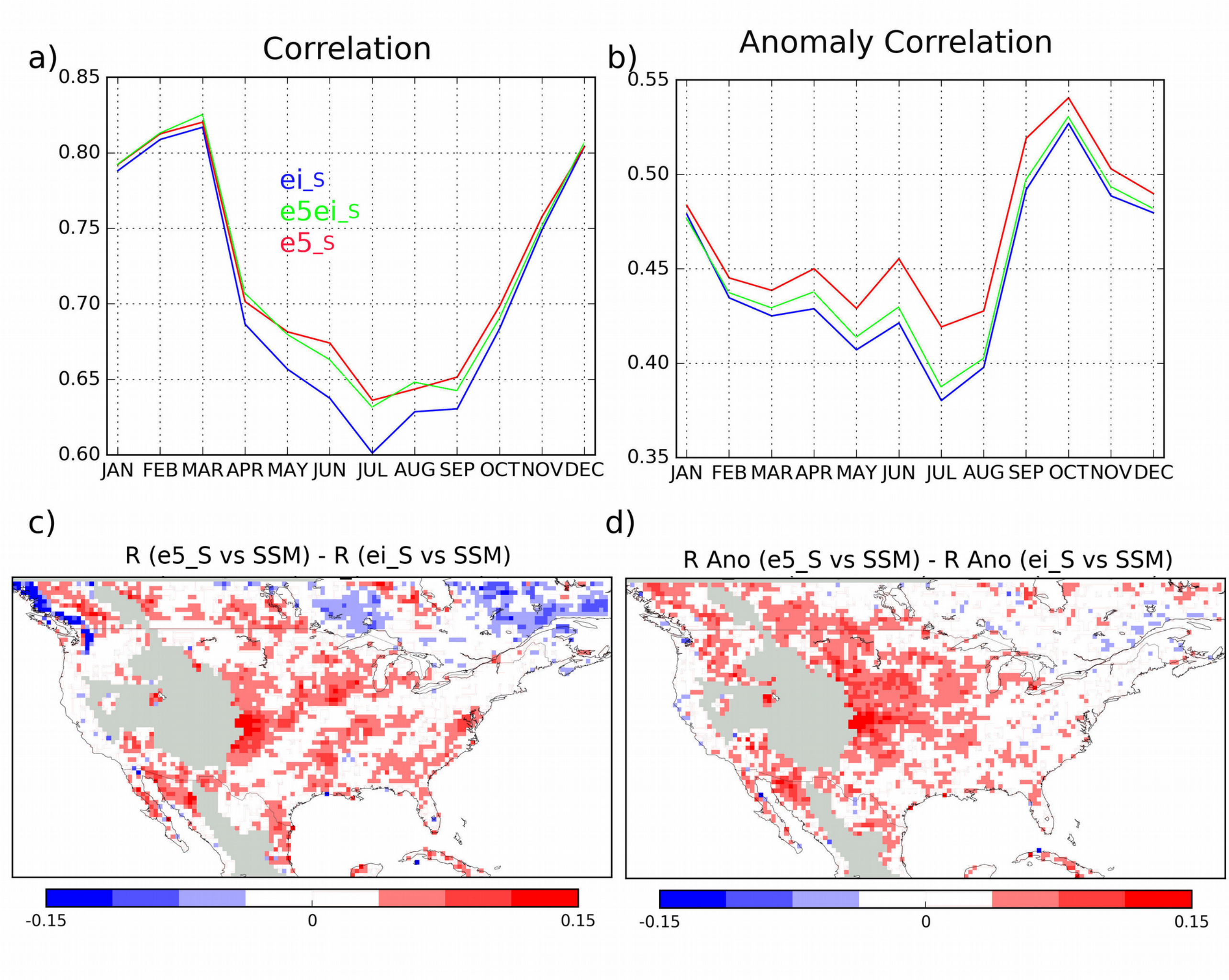

Figure 89 : Seasonal correlations on (a) volumetric time-series and (b) anomaly time-series between surface soil moisture estimates from the ESA CCI project (ESA-CCI SSM v4) and soil moisture from the second layer of soil of the ISBA LSM forced by ERA-Interim (ei_S in blue), ERA-5 but with precipitation from ERA-Interim (e5ei_S in green) and ERA-5 (e5_S in red) over 2010-2016. Maps of correlations differences between soil moisture from e5_S and ei_S on volumetric time-series (c) and anomaly time-series (d), areas in red represent an improvement from the use of ERA-5. Grey areas represents areas that where flagged out for elevation greater than 1500 m above sea level.

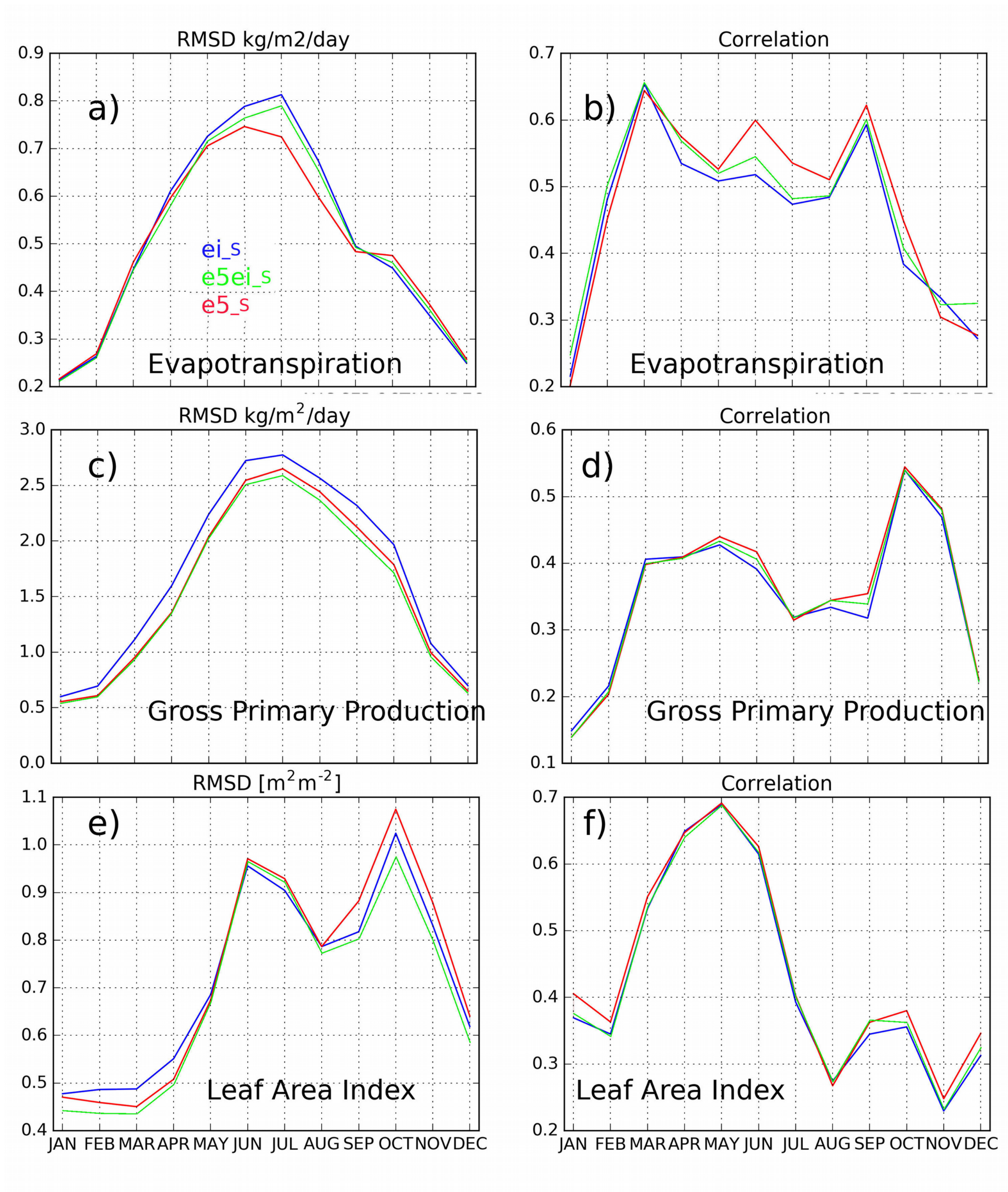

Figure 9~~10~~: Seasonal scores between ISBA LSM within SURFEX forced by either ERA-Interim (ei_S in blue) ERA-5 but ERA-Interim precipitation (e5ei_S in green) or ERA-5 (e5_S in red) and (a, b) evapotranspiration estimates from the GLEAM project over 2010-2016, (c, d) upscaled GPP from the FLUXCOM project over 2010-2013 and (e, f) LAI estimates from the Copernicus GLS project over 2010-2016. Left column (a, c and e) are for RMSD and right column (b, d, e) for correlations.

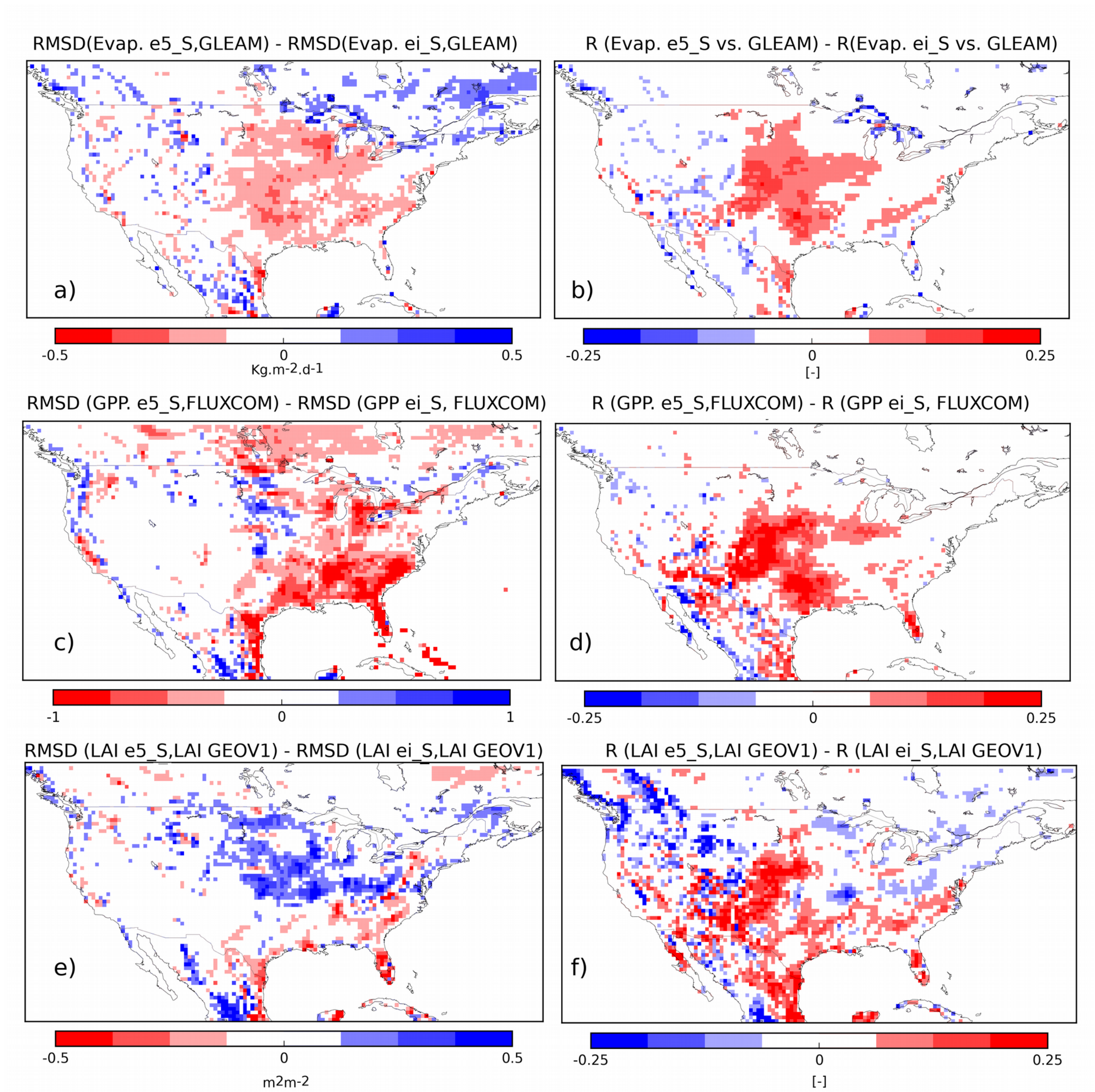

Figure 10~~1~~: ~~Normalised~~ RMSD differences (a, c, e) and ~~Normalised Information Contribution~~Correlation differences (~~NIC,~~ b, d, f) ~~based on correlation values~~ for e5_S simulations with respect to ei_S simulations for three land surface variables:~~;~~ evapotranspiration, Gross Primary Production and Leaf Area Index from top to bottom. Areas in red represent an improvement from the use of ERA-5. ~~For sack of clarity a factor 100 has been applied on NIC$_R$.~~