Fairbairn *et. al.*, (2016):
http://www.hydrol-earth-syst-sci-discuss.net/hess-2016-195/
Response to comments from Referee 2

July 26, 2016

Firstly, we would like to thank the reviewer for his/her constructive comments. A point by point response is given below.

Response to major comments:

1

1.1

Referee comment This study describes the implementation of a simple Extended Kalman Filter (SEKF) to assimilate LAI and SSM observations into a hydrological model over France, and its validation against streamflow measurements.

Response:
This is not exactly the objective of this study. It is important to mention that the assimilation is into a land surface model and not into a hydrological model. We will add this sentence to the introduction (Section 1): "In this study, the Simplified Extended Kalman Filter (SEKF) is used to assimilate LAI and SSM observations to update LAI and root-zone soil moisture (WG2) in the ISBA-A-gs land surface model. The drainage and runoff outputs from the land surface model are then used to force the MODCOU hydrogeological model and are validated by comparing the simulated streamflow with observations." It is important to clarify that the land surface model and the assimilated land surface observations are independent of the hydrogeological validation. This is different to other studies such as Thirel *et al.* (2010), where streamflow observations were assimilated and used to update the soil moisture in the ISBA land surface model. We will try to make this more clear throughout the paper, including the abstract, introduction and conclusions.

1.2

The topic is appropriate for the HESS journal, but the paper is not very well written. The technical approach appears sound at places and has some interesting aspects but there are many issues with the results, or at least their explanation which is not clear at all.

Response:
We agree with the reviewer that the experimental setup was not well explained and there were some mistakes in the way we presented the results. We hope that by offering clearer explanations we can resolve these problems.

1.3

The NITm and NITbc simulations use a different minimum LAI (1.2 m2/m2) and a bias-corrected radiative forcing (+5%) respectively, but nothing is said about how these numbers were chosen.

Response:
We admit this was not well explained and we will clarify this in the experimental setup (Section 2.4 of the paper). Fig. R2.1 (at the end of the document) shows a histogram of the observed average annual LAI minimum (GEOV1 satellite-derived observations) for the 133 grid-points over France with predominantly grasslands (the grassland patch fraction exceeding 70%). We chose an augmented grassland LAI minimum value of 1.2 m^2/m^2 for NITm because over 99% of the predominantly grassland points in Fig. R2.1 have an observed average annual LAI minimum above this value. The average annual LAI minimum over France for the original simulation (NIT), the new simulation (NITm) and the GEOV1 data are shown in Fig. R2.2 (note that the original Fig. 4 in the paper did not have the correct scale). Fig. R2.2 emphasizes that the LAI minimum was underestimated (compared to the GEOV1 data) over much of France for the original model simulation. By increasing the grassland LAI minimum to 1.2 m^2/m^2, the model agrees much better with the data over most regions. Szczypta *et al.* (2011) and Le Moigne (2002) demonstrated that the direct short-wave and long-wave radiative forcing respectively are underestimated by approximately 5% averaged over France. We followed Decharme *et al.* (2013) in bias-correcting the direct radiative forcing by +5%.

1.4

Was the new minimum LAI chosen based on the observations? If so, there really is no point in comparing the LAI from the new simulations with the same data.

Response:
The aim of this study was not to carry out an independent validation of LAI for each experiment, for which we would need independent observations. We will add the following in the experimental setup (Section 2): "A system validation was performed by comparing the LAI and WG1 states with the LAI and SSM observations respectively for all the simulations and data assimilation experiments. Note that this is not an independent validation of the performance of the system, for which we would need independent observations. The rationale was to check the effectiveness of the SEKF i.e. to see if it improved the fit between the model simulations and the observations. The fit to the observations was determined by the root mean square difference (RMSD), the correlation coefficient (CC) and the bias. These checks were important because the performance of the SEKF had an important impact on the drainage and runoff fluxes."

1.5

Additionally, the Nash scores of the NITm and NITbc simulations are shown only for the stations where at least one of the simulations had a positive score (p. 9, l. 21-22). Essentially, the average NSEs reported in Table 4 and Fig. 6 are artificially better than what they ought to be since most of the stations in northern and southeast France are excluded from this calculation. No explanation is given at to why this was done, making the discussion of the results rather dubious.

Response:
We agree with the reviewer that we did not present these results correctly and have therefore shown the results with all the stations (including the negative scores) included in the calculations. Fig. R2.3(a) shows a scatter plot of the Nash efficiency for the NIT simulation against the NITbc simulation. The density of the points is derived from the kernel density estimation of Scott (1992). The NIT simulation is the original simulation. The NITbc simulation is the new simulation with the augmented LAI minimum and radiative forcing. The results are improved for about 80% of the stations, including most of the stations with negative Nash scores. Fig. R2.3(b) shows a scatter plot of the NIT against LDAS2 (NIT with the assimilation of SSM and LAI). The assimilation degrades the SIM discharge scores for about 70% of the stations. These results are consistent with the original conclusions of the study. Note that the LDAS2 experiment

was performed with a more appropriate estimate of the observation error standard deviation than the original experiment. We used a value of 0.65(w_{fc}-w_{wilt}) instead of the original value of 0.4(w_{fc}-w_{wilt}), where w_{fc} is the field capacity and w_{wilt} is the wilting point. This averages to 0.055 m^3/m^3 over France. We used a slightly larger error than the estimated error of 0.05 m^3/m^3 by Draper *et al.* (2011) in order to approximate the oversampling issue i.e. the same ASCAT observation covers several gridpoints. This is comparable with observation errors expected for remotely sensed SSM observations (de Jeu *et al.*, 2008; Draper *et al.*, 2013). Note that this larger observation error slightly reduces the impact of the assimilation of SSM relative to the original experiment, but the conclusions of the study remain unchanged. Table R2.1 (at the end of the document) shows the new WG1 scores (model state compared with SSM observations) and will replace Table 3 in the paper.

Table R2.2 will replace Table 4 in the paper and shows the median Nash efficiency scores for the all the experiments. Following a comment from Jean-Philippe Vidal (see 1.6 below), the scores are also shown for the 67 stations with low anthropogenic influence. Note that the median is calculated rather than the mean because the majority of stations ($> 80\%$) have positive Nash efficiency scores, but a few outliers have scores near to -100. The median is a more appropriate metric as it is less sensitive to extreme outliers and is a better indicator for highly skewed distributions (Moriasi *et al.*, 2007). The results in Table R2.2 are very similar to Table 4 in the paper except that the LDAS2 experiment has less impact (due to the larger observation error). Therefore, including the stations with negative scores does not change the conclusions of this study.

In the revised version of the paper, we will replace Figure 6 with the Median Nash efficiency scores for all the stations. Note that Figure 5 in the paper was actually correctly presented and there was a mistake in the caption - all the stations were considered in the calculations, not just the stations with positive scores. Fig. R2.3 will be included in a supplement.

1.6

Dr Jean-Phillipe Vidal from IRSTEA posted a short comment, which is related to comment 1.5. He was right to point out that many of the stations included in the calculations are influenced by anthropogenic water management, which is not simulated by the MODCOU hydrogeological model. He was concerned that the results might be interpreted as being closer to anthropogenically influenced streamflow. He suggested the following:

1. To consider only catchments with low anthropogenic influence in order not to compare apples and oranges and avoid drawing conclusions on the ability of

SIM (with or without data assimilation) to simulate anthropogenically influenced streamflow,

2. To show scatter plots of NSEs instead of distributions (possibly with marginal distributions) to reduce the potential spatial bias effect mentioned above.

Response:
We have decided to follow Dr Jean-Phillipe Vidal's suggestions by showing the results for the stations with low-anthropogenic influence. We have used the suggested reference networks of Giuntoli *et al.* (2012, 2013) to extract a subset of 67 river gauges with low-anthropogenic influence from the original 546 stations, valid for both low and high flows. A map of these stations is shown in Fig. R2.4. A scatter plot is shown of the Nash efficiency of these stations (labeled as "Low anth. influence") and all the other stations (labeled as "High anth. influence") in Fig. R2.5. The same results are shown as in Fig. R2.3, but for the sake of clarity, in Fig. R2.5 only the stations are shown in the range of Nash scores -1.0 to 1.0. The 'low anth. influence' stations follow a similar pattern to the 'high anth. influence' stations. Furthermore, we calculated the Median Nash efficiency scores for the 67 stations in Table R2.2. The scores for this subset are improved relative to the 546 stations in Table R2.2, as expected. In particular, the percentage of stations with good scores (Nash efficiency > 0.6) is increased significantly. The discharge bias is also slightly less for the stations with low anthropogenic influence relative to the 546 stations. This supports Jean-Phillipe Vidal's suggestion that part of the positive bias in the discharge ratio of the NIT simulation for the 546 stations could be attributed to abstractions not being accounted for. However, the majority of the discharge bias in the NIT simulation is still present with the 67 stations with low anthropogenic influence. Moreover, the relative performance of the experiments is very similar to the original 546 stations. Therefore, the conclusions of the experiments are not affected by the ability of SIM (with or without data assimilation) to simulate anthropogenically influenced streamflow.

We will explain these results in Section 3.4 of the paper. Figures R2.4 and R2.5 will be included in a supplement.

1.7

Furthermore, the assimilation doesnt appear to have much of an impact on the streamflow simulations and actually decreases the skill (even when excluding the stations that had the negative NSE). I wonder what the rationale was of not using a more sophisticated data assimilation algorithm that could overcome some of the limitations in the SEKF. There are many limitations with this approach that I dont see any worthwhile scientific contribution added by this study, although there are some interesting aspects in this work.

Response:
The reviewer is right that in many data assimilation applications with hydrological models, more sophisticated algorithms are commonly used that take into account the "errors of the day". We will clarify in the introduction the differences between hydrological and land surface data assimilation, and why we chose the SEKF for our experiments.

Many studies have investigated the assimilation of SSM and streamflow observations into hydrological models in order to improve streamflow predictions and hydrological parameters (Thirel *et al.*, 2010; Aubert *et al.*, 2003; Clark *et al.*, 2008; Moradkhani *et al.*, 2005, 2012). In large-scale streamflow assimilation, a DA method is typically chosen that can take into account lateral background-error covariances and flow-dependence. These features are important because streamflow has important horizontal interactions. For example, Thirel *et al.* (2010) used the Best Linear Unbiased Estimate (BLUE) method to assimilate steamflow observations into the MODCOU hydrogeological model, which they used to update soil moisture in the ISBA land surface model (LSM). Although they used a fixed diagonal background-error covariance at the start of each window, they generated implicit background-error covariances between the river subbasins using finite differences in the observation operator Jacobian calculation. This led to improved streamflow predictions.

LSMs concern water and energy fluxes between the soil and atmosphere. Unlike hydrological models, layer-based LSMs such as the ISBA model are typically pointwise (there is no horizontal interaction between the gridpoints), since this greatly reduces the computational expense. It is also common to use a DA method with 1D Kalman filtering (where observations are used to update colocated gridpoints) as opposed to 2D Kalman filtering (where observations are used to update colocated gridpoints and neighbouring gridpoints). Moreover, a study by Gruber *et al.* (2015) over the contiguous US found that there was no advantage of 2D Kalman filtering over 1D Kalman filtering when assimilating ASCAT SSM data into a soil moisture model. They explained these results using an analytical evaluation of the impact of spatial-error autocorrelations on the steady-state Kalman gain.

In large-scale land surface DA, it is common to assimilate satellite-derived surface soil moisture (SSM) observations and screen-level temperature and humidity observations into a LSM, in order to improve soil moisture and screen-level variables. Typically WG2 is of more interest than SSM as it has a much larger water capacity and therefore has a greater influence on vegetation and water fluxes. Land surface DA is commonly performed using a 1D ensemble Kalman filter (EnKF) or a simplified extended Kalman filter (SEKF). The SEKF simplifies the EKF by assuming that the errors in the background state are fixed and uncorrelated between gridpoints. It uses finite differences for computing Jacobians necessary to extract information from the observations to the prognostic variables.

There has been increasing interest in ensemble DA for LSMs over the last two decades (Reichle *et al.*, 2002, 2008; Zhou *et al.*, 2006; Muñoz Sabater *et al.*, 2007; Draper *et al.*, 2012; Carrera *et al.*, 2015), partly because these methods can estimate the "errors of the day" in the background-error covariance. At Environment Canada, the development of an operational EnKF using the same ISBA 3-layer model we employed is also motivated by the requirements of coupling land surface DA with NWP ensemble prediction (Carrera *et al.*, 2015). However, the correct representation of the "errors of the day" is challenging in land surface DA. A large proportion of the errors in LSMs come from the model and the atmospheric forcing, rather than the initial conditions. The integrating nature and the nonlinear interactions in LSMs mean that short-term errors dissipate over time, including random errors in the precipitation. For example, a study by Maggioni *et al.* (2011) found that errors in WG2 are not very sensitive to the rainfall error modelling approach. Indeed, experiments assimilating in situ SSM observations with the ISBA-A-gs model have demonstrated that the EnKF and the SEKF produce a WG2 analysis with comparable accuracy and both methods improve on the model simulation (Muñoz Sabater *et al.*, 2007; Fairbairn *et al.*, 2015).

Due to its efficacity, simplicity and low computational cost, the SEKF is the preferred method at several meteorological operational centres for analyzing soil moisture and screen level variables. Hess (2001) developed a simplified 2D-Var (theoretically equivalent to an SEKF) scheme for the assimilation of screen-level temperature and humidity at the German Weather service (DWD). The European Centre for Medium Range Weather Forecasts (ECMWF) assimilate screen-level temperature and humidity operationally with an SEKF (de Rosnay *et al.*, 2013). An SEKF was developed for research purposes to assimilate ASCAT satellite derived soil moisture at Meteo-France (Draper *et al.*, 2009; Mahfouf, 2010). Recently, ECMWF have also modified their SEKF to assimilate ASCAT derived SSM observations (ECMWF, 2016). At the UK Met Office, an SEKF has been developed for research purposes for the assimilation of a wide variety of observation types, including screen-level variables and satellite derived SSM observations (Candy *et al.*, 2012).

In our study, we use an SEKF to assimilate LAI and SSM observations to update LAI and WG2 in the ISBA-A-gs LSM in the SAFRAN-ISBA-MODCOU (SIM) hydrological suite. SIM consists of three stages: (1) An atmospheric reanalysis (SAFRAN) over France, which forces (2) the ISBA-A-gs land surface model, which then provides drainage and runoff inputs to (3) the MODCOU hydrogeological model. The drainage and runoff outputs from ISBA-A-gs are validated by comparing the simulated streamflow from MODCOU with observations. Our study is different to the hydrological studies mentioned earlier because the LSM and the DA are independent of the hydrogeological model. This study is relevant to the land surface DA community because several operational centres assimilate SSM observations using an SEKF to update WG2. Many studies have demonstrated that the force-restore dynamics of the ISBA 3-layer model can skillfully simulate soil moisture and propagate the increments downwards from the surface to the root-zone (Muñoz Sabater *et al.*, 2007; Draper *et al.*, 2009; Mahfouf *et al.*, 2009; Barbu *et al.*, 2011, 2014). An integrated validation using SIM has also demonstrated that the ISBA 3-layer model can skillfully simulate drainage and runoff fluxes over France (Habets *et al.*, 2008). But relatively few studies have assessed the SEKF performance using an integrated validation of the soil moisture fluxes. To our knowledge, this is the first article to perform this type of validatation for LAI assimilation. Moreover, the validation is robust because it is performed using more than 500 river gauges over France and the length of the analysis period spans several years.

In the discussion section we mentioned that in future studies we would like to test the EnKF over France with a stochastic representation of precipitation and model errors using a similar hydrogeological evaluation to this study. The EnKF would not be affected by some of the issues we encountered with the SEKF, including the collapse of the observation operator Jacobians during wet conditions. However, the choice of DA method is not the only problem. In the discussion section we also mentioned important deficiencies in the 3-layer ISBA-A-gs land surface model, including no vertical variability in WG2. These deficiencies inhibit the SEKF performance. We expect the SEKF to perform significantly better with a new multi-layer diffusion based model (ISBA-DIF). For example, with the 3-layer model we assimilate SSM observations into a very shallow layer (0-1 cm), which is very sensitive to the atmospheric forcing over the 24 hour assimilation window. It is possible with ISBA-DIF to assimilate them into a slightly deeper layer (1-5 cm), which is less sensitive to the atmospheric forcing.

1.8 Minor comments

1. p. 2, l. 10: I would replace the term network, which usually refers to in-situ measure- ments.

   Response: Agreed, we will replace "network" by "coverage".

2. p. 2, l. 10 "a short forecast from the past": it doesnt have to be from the past, it can be a prediction of the current time (i.e. observation time).

Response: Agreed, we will replace "a short forecast from the past" with "a short forecast from the previous analysis".

3. p. 4, l. 22: can the authors add a sentence on what the delayed cut-off version of SAFRAN is?

   Response: We will add: "The delayed cut-off version of SAFRAN includes additional observations obtained after the real-time cut-off, which makes it more accurate. The delayed cut-off version of SAFRAN uses additional observations from over 3000 climatological observing stations, which report once monthly".

4. p. 5, l. 15: why were only ASCAT observations used and not SMOS for example? Is it because of the study period?

   Response: Yes, ASCAT has the advantage of being available over the study period and ASCAT-like data will be available for decades to come. Also, the SEKF at Meteo-France is calibrated to assimilate ASCAT observations and the assimilation has already been performed in a number of studies (Draper *et al.*, 2009; Barbu *et al.*, 2014). The aim of the study was not to test new soil moisture data sources, but to validate the soil water fluxes of the existing system using a hydrogeological model. The assimilation of multiple satellite products is being explored in a different study.

5. p. 5, l. 21: why do the soil water index data need to be interpolated to the model resolution? Can the SEKF not handle different spatial resolutions between the model and the observations?

   Response: The SEKF assimilates observations in model space (i.e. the same grid as the model), so it is necessary to perform this interpolation.

6. p. 5, l. 21: p. 5, l. 25: has the WG1 soil moisture climatology been validated?

   Response: The ISBA model soil moisture states have been compared with satellite or in situ observations in several studies and generally show a good level of skill (e.g. Draper *et al.* (2009); Muñoz Sabater *et al.* (2007); Albergel *et al.*

(2008, 2010b); Barbu *et al.* (2014); Fairbairn *et al.* (2015)). Most of these studies have also demonstrated that the SEKF can significantly improve the soil moisture scores.

7. p. 5, l. 32: were additional LAI products considered (e.g. MODIS)?

   Response: Extensive comparisons of GEOV1 and MODIS are available from the Copernicus GLS website (http://land.copernicus.eu/global/sites/default/files /products/GIOGL1_VR_LAIV1_I1.10.pdf and http://land.copernicus.eu/global/sites/default/files/products/GIOGL1 _QAR_PROBAV-GEOV1_I3.10.pdf). The direct validation based on in situ LAI observations shows that the GEOV1 products present slightly better scores than MODIS.

   In any case, the aim of this study was not to test new observation datasets but to work with the existing system. The SEKF is already set up to assimilate GEOV1 observations (Barbu *et al.*, 2014).

8. p. 7, l. 25-27: this is confusing, how are the 1.2 m2/m2 and +5% values obtained?

   Response: Please see 1.3 above.

9. 8, l. 5: how are the LAI and WG1 estimates validated against satellite observations? Werent these satellite observations assimilated into the model?

   Response: We agree with the reviewer that this sentence is misleading: "The LAI and WG1 state estimates for the experiments are validated using the satellite observations". We will remove this sentence and replace it with the following: "A system validation was performed by comparing the LAI and WG1 states with the LAI and SSM observations respectively for all the simulations and data assimilation experiments. Note that this is not an independent validation of the performance of the system, for which we would need independent observations. The rationale was to check the effectiveness of the SEKF i.e. to see if it improved the fit between the model simulations and the observations. The fit to the observations was determined by the root mean square difference (RMSD), the correlation coefficient (CC) and the bias. These checks were important because the performance of the SEKF had an important impact on the drainage and runoff fluxes."

10. p. 9, l. 28-29: I don't understand how the good performance of the NITbc is explained by the relationship between the bias in the discharge ratio and the NSE. Doesnt the NITbc just have a bias-corrected radiative forcing? Where is the causality between the simulation configuration and the performance? Wouldn't it make sense that the model with the smaller bias would have better performance in terms of NSE?

    Response: We agree this could be clearer. We explained the causality in the original paper in the following paragraph (starting line 32, page 9) by examining the impact of the different simulations on the soil water fluxes. On line 3 of page 10 we mention that: "The NITbc simulation increases the direct radiative forcing by 5%, which results in increased evapotranspiration and lower WG2 during the year. This significantly reduces the drainage and runoff from October to June." We suggest adding another sentence: "The reduced drainage and runoff feeding into the MODCOU hydrogeological model results in less river discharge, which reduces the positive discharge bias. This in turn improves the Nash efficiency scores."

11. p. 12, l. 14-15: but nothing is said on how the higher LAI parameter was chosen.

    Response: Please see 1.3 above.

Table R2.1: Scores for WG1 (model state compared with observations) averaged over 2007-2014. The closest fit to the observations are shown in bold font.

Experiment	RMSD (m^3/m^3)	CC	Bias (m^3/m^3)
NIT	0.051	0.77	0.00
NIT_m	0.049	0.77	0.00
NIT_{bc}	0.051	0.77	0.00
LDAS1	0.049	0.77	0.00
LDAS2	**0.048**	**0.78**	0.00
$\mathrm{LDAS1}_{bc}$	0.049	0.77	0.00
$\mathrm{LDAS2}_{bc}$	0.049	**0.78**	0.00
$\mathrm{LDAS2}_{QC}$	**0.048**	**0.78**	0.00

Table R2.2: Median Nash efficiency (NE) and discharge ratio (Qs/Qo) scores over the 546 river gauges over France and for the subset of 67 gauges with low anthropogenic influence, calculated over 2007-2014. Also shown are the % of stations with a Nash score above 0.6. The best scores are shown in bold font.

Experiment	NE for 546/67 stations	Discharge ratio for 546/67 stations	% stations with NE > 0.6 for 546/67 stations
NIT	0.44/0.48	1.19/1.16	26%/44%
NIT$_m$	0.48/0.54	1.15/1.12	30%/48%
NIT$_{bc}$	**0.56/0.60**	**1.02/0.99**	**42**%/**59**%
LDAS1	0.44/0.48	1.18/1.15	27%/44%
LDAS2	0.41/0.45	1.21/1.18	23%/40%
LDAS1$_{bc}$	**0.56/0.60**	**1.02/1.00**	**42**%/**57**%
LDAS2$_{bc}$	0.53/0.54	1.08/1.06	38%/53%
LDAS2$_{QC}$	0.40/0.45	1.21/1.18	21%/39%

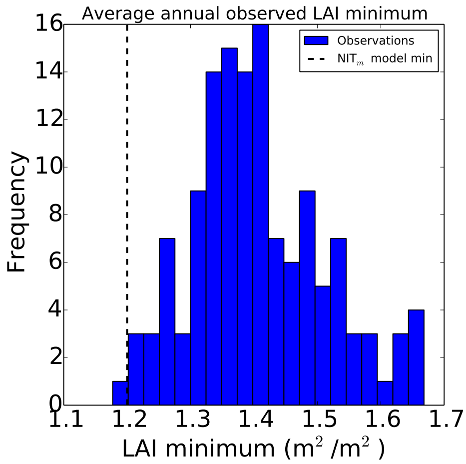

Figure R2.1: Histogram of the average annual LAI minimum (2007-2014) values for the observations over predominantly grassland points ($>$ 70% grasslands) (m^2/m^2) over France. Also shown is the NIT$_m$ LAI minimum parameter for grasslands.

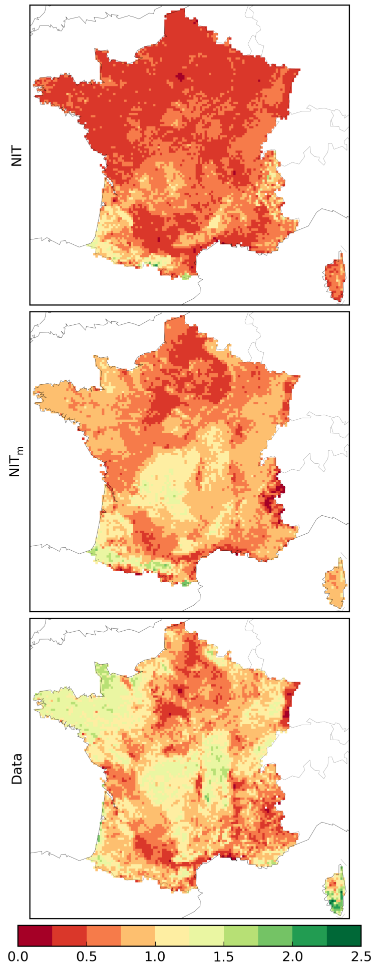

Figure R2.2: Map showing the average annual LAI minimum (2007-2014) for NIT, NIT$_m$ and the observations (m^2/m^2) over France.

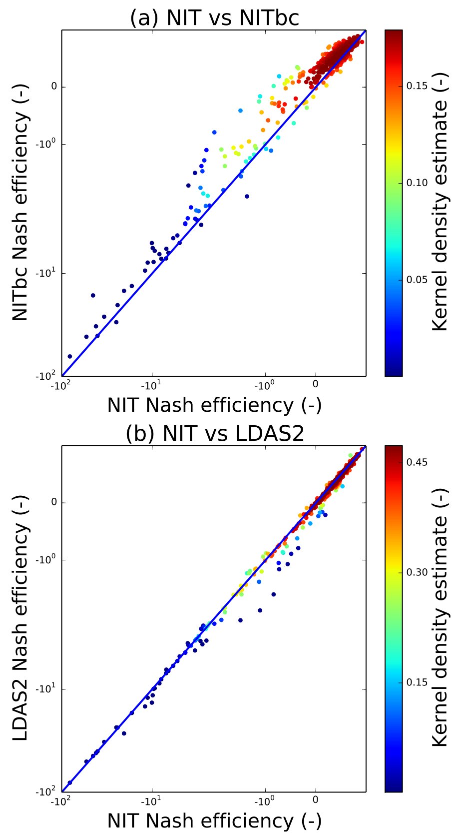

Figure R2.3: Scatter plots of the SIM discharge Nash efficiency scores for all 546 stations for (a) NIT vs NITbc and (b) for NIT vs LDAS2. The scores are calculated over 2007-2014.

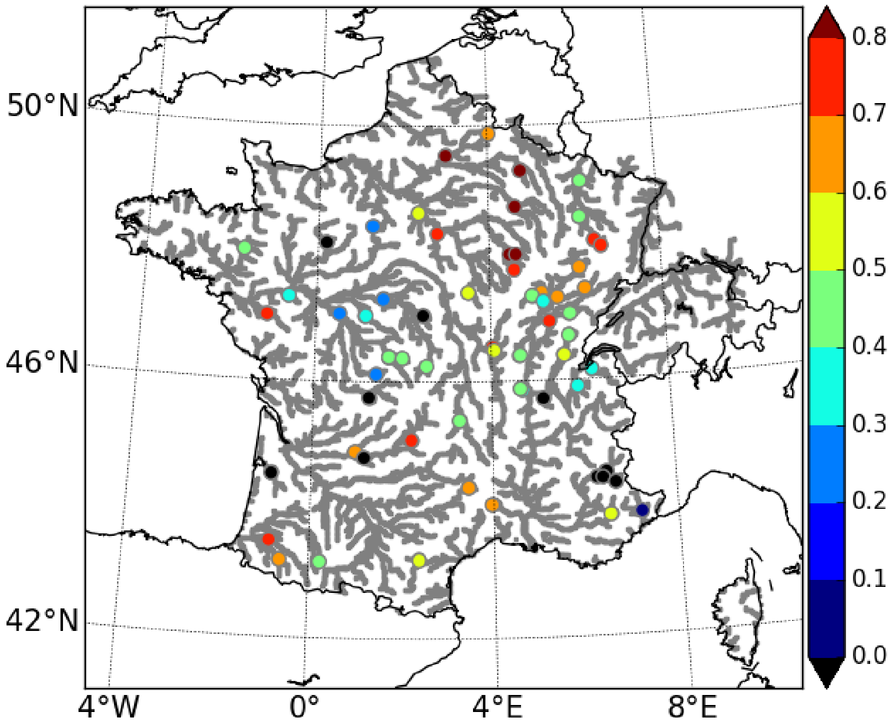

Figure R2.4: Map of the SIM discharge Nash efficiency scores for the 67 stations with low-anthropogenic influence over France for the NIT simulation, calculated over the period 2007-2014. The river network is also shown.

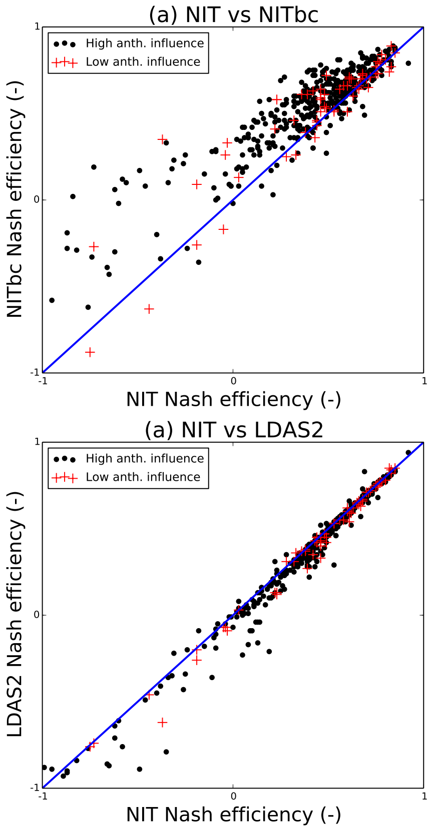

Figure R2.5: Same as Fig. R2.3, but the stations are classified with either low (67 stations) or high anthropogenic influence (479 stations). For the sake of clarity, the Nash scores are shown between -1.0 and 1.0.

References

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

C. Albergel, C. Rüdiger, T. Pellarin, J.-C. Calvet, N. Fritz, F. Froissard, D. Suquia, A. Petitpa, and B. Piguet. From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations. *Hydrol. Earth Syst. Sci*, 12:1323–1337, 2008.

D. Aubert, C. Loumagne, and L. Oudin. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model. *J. Hydrol.*, 280:145161, 2003.

A.L. Barbu, J.C. Calvet, J.F. Mahfouf, C. Albergel, and S. Lafont. Assimilation of soil wetness index and leaf area index into the isba-a-gs land surface model: grassland case study. *Biogeosciences*, 8:1971–1986, 2011.

A.L. Barbu, J.C. Calvet, J.F. Mahfouf, and S. Lafont. Integrating ascat surface soil moisture and geov1 leaf area index into the surfex modelling platform: a land data assimilation application over france. *Hydrol. Earth Syst. Sci.*, 18:173192, 2014.

B. Candy, K. Bovis, I. Dharssi, and B. Macpherson. Development of an Extended Kalman Filter for the Land Surface. Technical report, Met Office, Exeter, UK, 2012.

M.L. Carrera, S. Bélair, and B. Bilodeau. The Canadian Land Data Assimilation System (CaLDAS): Description and Synthetic Evaluation Study. *J. Hydrometeorol.*, 16:1293–1294, 2015.

M.P. Clark, D.E. Rupp, R.A. Woods, X. Zheng, R.P. Ibbitt, A.G. Slater, J. Schmidt, and M.J. Uddstrom. Hydrological data assimilation with the ensemble kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Adv. Water Resour.*, 31:13091324, 2008.

R. A. M. de Jeu, W. Wagner, T. R. H. Holmes, A.J. Dolman, N. C. de Giesen, and J. Friesen. Global soil moisture patterns observed by space borne microwave radiometers and scatterometers. *Surv. Geophys.*, 29:399420, 2008.

P. de Rosnay, Matthias Drusch, Drasko Vasiljevic, Gianpaolo Balsamo, Clment Albergel, and Lars Isaksen. A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Q. J. R. Meteorol. Soc.*, 139:1199–1213, 2013.

B. Decharme, E. Martin, and S. Faroux. Reconciling soil thermal and hydrological lower boundary conditions in land surface models. *J. Geophys. Res. Atmos.*, 118: 78197834, 2013.

C.S. Draper, J.-F. Mahfouf, J.-C. Calvet, E. Martin, and W. Wagner. Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France. *Hydrol. Earth Syst. Sci*, 15:3829 – 3841, 2011.

C.S. Draper, J.-F. Mahfouf, and J.P. Walker. An EKF assimilation of AMSR-E soil moisture into the ISBA surface scheme. *Journal of Geophysical Research*, 114:D20104, 2009.

C.S. Draper, R.H. Reichle, R. de Jeu, V. Naeimi, R. Parinussa, and W. Wagner. Estimating root mean square errors in remotely sensed soil moisture over continental scale domains. *Remote Sensing of Environment*, 137:288–298, 2013.

C.S. Draper, R.H. Reichle, G.J.M. De Lannoy, and Q. Liu. Assimilation of passive and active microwave soil moisture retrievals. *Geophys. Res. Lett.*, 39:L04401, 2012.

ECMWF. Annual report 2015. ECMWF website, 2016. URL `http://www.ecmwf.int/sites/default/files/elibrary/2016/16478-annual-report-2015.pdf`. Last accessed July 2016.

D. Fairbairn, A.L. Barbu, J.-C. Calvet J.-F. Mahfouf, and E. Gelati. Comparing the ensemble and extended Kalman filters for in situ soil moisture assimilation with contrasting conditions. *Hydrol. Earth. Syst. Sci.*, 19:4811–4830, 2015.

I. Giuntoli, B. Renard, and M. Lang. *Changes in Flood risk in Europe*, pages 199–211. Kundzewicz, Z. K., Ed., 2012. ISBN 9780415621892. IAHS Special Publication 10.

I. Giuntoli, B. Renard, J.-P. Vidal, and A. Bard. Low flows in france and their relationship to large-scale climate indices. *Journal of Hydrology*, 482:105–118, 2013.

A. Gruber, W. Crow, W. Dorigo, and W. Wagner. The potential of 2D Kalman filtering for soil moisture data assimilation. *Remote Sensing of Environment*, 171:137–148, 2015.

F. Habets, A. Boone, J. L. Champeaux, P. Etchevers, L. Franchistéguy, E. Leblois, E. Ledoux, P. Le Moigne, E. Martin, S. Morel, J. Noilhan, P. Quintana Segu, F. Rousset-Regimbeau, and P. Viennot. The safran-isba-modcou hydrometeorological model applied over france. *J. Geophys. Res.*, 113:D06113, 2008.

H. Hess. Assimilation of screen-level observations by variational soil moisture analysis. *Meteorol. Atmos. Phys.*, 77:145–154, 2001.

P. Le Moigne. Description de l'analyse des champs de surface sur la France par le système SAFRAN. Technical report, Meteo-France/CNRM, 2002.

V. Maggioni, R.H. Reichle, and E.N. Anagnostou. The effect of satellite rainfall error modeling on soil moisture prediction uncertainty. *J. Hydrometeor*, 12:413428, 2011.

J.-F. Mahfouf. Assimilation of satellite-derived soil moisture from ascat in a limited-area nwp model. *Q.J.R. Meteorol. Soc.*, 136:784798, 2010.

J.-F. Mahfouf, K. Bergaoui, C. Draper, C. Bouyssel, F. Taillefer, and L. Taseva. A comparison of two off-line soil analysis schemes for assimilation of screen-level observations. *J. Geophys. Res*, 114:D08105, 2009.

H. Moradkhani, C. M. DeChant, and S. Sorooshian. Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-markov chain monte carlo method. *Water Resour. Res.*, 48:W2520, 2012.

H. Moradkhani, K.-L. Hsu, H. Gupta, and S. Sorooshian. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resour. Res.*, 41:W05012, 2005.

D.N. Moriasi, J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, and T.L. Veith. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers*, 50:885–900, 2007.

J. Muñoz Sabater, L. Jarlan, J.-C. Calvet, and F. Boyssel. From near-surface to root-zone soil moisture using different assimilation techniques. *J. Hydrometeor*, 8:194–206, 2007.

R.H. Reichle, W.T. Crow, and C.L. Keppenne. An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resour. Res*, 44:WO3243, 2008.

R.H. Reichle, J.B. Walker, R.D. Koster, and P.R. Houser. Extended vs Ensemble Kalman Filtering for Land Data Assimilation. *J. Hydrometeor*, 3:728–740, 2002.

D.W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization.* John Wiley & Sons, New York, Chicester, 1992.

C. Szczypta, J.-C. Calvet, C. Albergel, G. Balsamo, S. Boussetta, D. Carrer, S. Lafont, and C. Meurey. Verification of the new ECMWF ERA-Interim reanalysis over France. *Hydrol. Earth Syst. Sci.*, 15:647–666, 2011.

G. Thirel, E. Martin, J.-F. Mahfouf, S. Massart, and F. Habets. A past discharges assimilation system for ensemble streamflow forecasts over france part 1: Description and validation of the assimilation system. *Hydrol. Earth Syst. Sci.*, 14:1623–1637, 2010.

Y. Zhou, D. McLaughlin, and D. Entekhabi. Assessing the Performance of the Ensemble Kalman Filter for Land Surface Data Assimilation. *Mon. Wea. Rev.*, 134:21282142, 2006.