Science of Remote Sensing 9 (2024) 100129

Contents lists available at ScienceDirect

Science of Remote Sensing

journal homepage: www.sciencedirect.com/journal/science-of-remote-sensing

Joint assimilation of satellite-based surface soil moisture and vegetation conditions into the Noah-MP land surface model

Zdenko Heyvaert [a,b,*,1], Samuel Scherrer [b,a,1], Wouter Dorigo [b], Michel Bechtold [a], Gabriëlle De Lannoy [a]

[a] Department of Earth and Environmental Sciences, KU Leuven, Heverlee, Belgium
[b] Department of Geodesy and Geoinformation, TU Wien, Vienna, Austria

ARTICLE INFO

Keywords:
Soil moisture
Vegetation
Multi-sensor data assimilation
Multivariate data assimilation

ABSTRACT

This study explores the potential of integrating satellite retrievals of surface soil moisture (SSM) and vegetation conditions into the Noah-MP land surface model. In total, five data assimilation (DA) experiments were carried out. One of the experiments only assimilates SSM retrievals from the Soil Moisture Active Passive mission, two experiments only assimilate retrievals of vegetation conditions: either optical retrievals of leaf area index (LAI) from the Copernicus Global Land Service, or X-band microwave-based retrievals of vegetation optical depth (VOD) from the Advanced Microwave Scanning Radiometer 2. Additionally, two joint DA experiments are performed, each incorporating SSM and one of the vegetation products. The DA experiments are compared with a model-only run, and all experiments are evaluated using independent ground reference data of soil moisture, evapotranspiration, net ecosystem exchange and gross primary production (GPP). Assimilating only SSM improves estimates of the soil moisture profile (median SSM anomaly correlation improves with 0.02 compared to a model-only run), whereas assimilating LAI predominantly improves GPP estimates (reduction in median RMSD of 0.024 gC m^{-2} day^{-1} compared to a model-only run). The joint assimilation of SSM and vegetation conditions captures both of these improvements in a single, physically consistent analysis product. The DA increments show that this combined setup allows one satellite product to compensate for potential degradations introduced into the system by the other product. Furthermore, the joint SSM and VOD DA experiment has the smallest ensemble spread in its estimates (21% reduction in SSM spread compared to a model-only run). Overall, our results underline the potential of multi-sensor and multivariate DA, in which information from different sources is combined to improve the estimates of several land surface states and fluxes simultaneously.

1. Introduction

The importance of soil moisture and vegetation in hydrological processes and land-atmosphere interactions is well-known (Seneviratne et al., 2010; Mahfouf, 2010; Kumar et al., 2022). Soil moisture is the water source for plant transpiration, which contributes to the exchange of energy, water, and carbon between the atmosphere and the land surface and vice versa (Powell et al., 2013; Bonan, 2019). Soil moisture and vegetation on the land surface can be estimated using a range of satellite instruments measuring at various wavelengths of the electromagnetic spectrum (Balsamo et al., 2018). However, satellite observations have the disadvantage of having temporal and spatial discontinuities while not being able to measure fluxes directly (Dorigo et al., 2021a). An alternative way to describe the land surface is to use land surface models (LSMs) to simulate energy, water, and carbon fluxes at the surface and in the root zone. LSMs allow to produce physically consistent estimates of geophysical variables at any time and location (Srinivasan et al., 2000), but individual fields may have large uncertainties (Koster et al., 2002; Dirmeyer et al., 2004; Seneviratne et al., 2016).

Data assimilation (DA) can be used to constrain LSM estimates using remotely sensed observations, combining both sources of information to obtain physically consistent, gap-free estimates of any variable such as soil moisture, evapotranspiration, and sensible heat fluxes (Lahoz and De Lannoy, 2014). In this study, two multi-sensor (using products from different satellites) and multivariate (directly updating multiple

* Corresponding author. Department of Earth and Environmental Sciences, KU Leuven, Heverlee, Belgium.
E-mail address: zdenko.heyvaert@kuleuven.be (Z. Heyvaert).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.srs.2024.100129
Received 17 January 2024; Received in revised form 1 March 2024; Accepted 20 March 2024
Available online 26 March 2024

variables within the LSM) DA systems are introduced. These systems allow constraining of the LSM estimates from both the hydrological and vegetative fronts. Previous studies, e.g., Scherrer et al. (2023), have shown that constraining the LSM using only a single variable (e.g., vegetation) can degrade the results in another variable (e.g., soil moisture) in some areas. Crow et al. (2024) explain such a degradation with the presence of water state–water flux coupling strength biases existing in LSMs. Constraining multiple variables with observations may help to overcome this issue (De Lannoy et al., 2022). The following three paragraphs discuss earlier research on the assimilation of retrievals that are also assimilated in this study.

Surface soil moisture (SSM) refers to moisture in the upper few centimeters of the soil and can be derived from remotely sensed observations in the microwave spectrum (Kerr et al., 2010; Paloscia et al., 2013; Entekhabi et al., 2014; Dorigo et al., 2017). SSM DA has been widely used to improve soil moisture profile estimates, either by using backscatter or brightness temperatures (e.g., Loew et al., 2009; De Lannoy and Reichle, 2016a,b; Lievens et al., 2017; Reichle et al., 2019) or by directly using SSM retrievals from active or passive microwave sensors (e.g., Reichle and Koster, 2005; Draper et al., 2012; Kumar et al., 2014; Lievens et al., 2015; Heyvaert et al., 2023). Because of the different nature of SSM estimated by LSMs and satellite sensors, systematic biases between the model and the observations are typically removed prior to assimilation, e.g., by using climatological or seasonal cumulative density function (CDF) matching (Reichle and Koster, 2004; Barbu et al., 2014; Heyvaert et al., 2023).

Leaf area index (LAI) is a measure of the amount of leaf material in a canopy and is calculated as the ratio of the single-sided leaf area to the ground area. It can be determined using optical sensors (Fang et al., 2019), and has been assimilated into different land surface models for applications focusing on hydrology or vegetation. Like for SSM, biases between observed and simulated LAI are omnipresent. Some studies therefore rescale the LAI prior to assimilation in order to remove the bias (e.g., Jarlan et al., 2008; Khaki et al., 2020). However, several other studies opt to assimilate the LAI as is (Barbu et al., 2014; Albergel et al., 2017; Kumar et al., 2019; Erlingis et al., 2021; Rahman et al., 2022b). Scherrer et al. (2023) examined the impact of bias between observed and modeled LAI on the DA system, and showed that it can cause the model to move away from its equilibrium state and drift between update steps. Therefore, the LAI observations in this work have been rescaled to ensure an unbiased system.

Vegetation optical depth (VOD) is an indicator of the total water content stored in above-ground vegetation (Mo et al., 1982; Jackson and Schmugge, 1991). It can be retrieved from different satellite sensors in the microwave spectrum and for different frequency bands (e.g., L-, C-, X-, and Ku-band). The sensitivity of VOD to different parts of the vegetation depends on the frequency band. The high-frequency bands (X/Ku-band) are mainly sensitive to the upper parts of the canopy, i.e., leafs, and are highly correlated with optical vegetation indices such as LAI (Moesinger et al., 2020). Low-frequency bands (L-band) mostly penetrate the upper canopy and are more sensitive to woody vegetation (Frappart et al., 2020).

Kumar et al. (2020, 2021) have previously assimilated X-band VOD and C-band VOD from the VOD Climate Archive (VODCA; Moesinger et al., 2020) and L-band VOD from Soil Moisture Active Passive (SMAP) into the Noah-MP LSM to obtain better LAI estimates. Since Noah-MP does not directly model VOD, they used monthly CDF matching to a reference LAI as an observation operator, i.e., they replaced the VOD climatology with the climatology of the reference LAI and only used the VOD anomalies from the mean seasonal cycle.

Mucia et al. (2022) used a similar approach, but using the Land Data Assimilation System (LDAS)-Monde (Albergel et al., 2017) platform with a simplified extended Kalman filter (SEKF), and replacing the CDF matching with a monthly linear rescaling. In our study, we develop a custom observation operator that maps root-zone soil moisture (RZSM) and LAI to X-band VOD, based on previous studies that discuss the relation of VOD with other land components (Momen et al., 2017; Rodríguez-Fernández et al., 2018; Konings et al., 2019; Vermunt et al., 2020; Bousquet et al., 2021; Liu et al., 2021).

The novelty of this paper lies in the development of a multivariate, multi-sensor DA system, which has been expressed as a priority within the land DA community (Durand et al., 2021; De Lannoy et al., 2022) because it allows to constrain more variables related to the water cycle (Girotto et al., 2019) and to improve the temporal coverage of observations that go into the DA system (Kumar et al., 2022). To our knowledge, this is the first time that SSM and X-band VOD retrievals from different sensors are simultaneously assimilated over Europe using a VOD observation operator that links both modeled RZSM and LAI to observation predictions of VOD. Among other things, this allows for a detailed evaluation of how soil moisture updates from different satellite products interact with each other. In addition, we explore the potential of the joint SSM and VOD DA system to improve estimates of the soil moisture profile and land surface fluxes. We compare the novel joint SSM and VOD assimilation to a system that jointly assimilates SSM and LAI. Assimilating LAI instead of VOD is arguably a simpler approach, since LAI is directly modeled by an LSM with dynamic vegetation growth. Several studies have already jointly assimilated SSM and LAI (e.g., Albergel et al., 2017; Xu et al., 2021; Rahman et al., 2022a), and this experiment serves in the first place as a benchmark to assess how a joint assimilation of SSM and VOD compares to a joint assimilation of SSM and LAI. The VOD observations have the advantage of a better temporal revisit and the potential of a multivariate state update, i.e., to improve the soil moisture as well, since VOD is related to the water content in the vegetation - which in turn is linked to water in the root zone.

The paper is organized as follows. The different DA experiments are introduced in section 2 in terms of the used model settings, satellite retrievals, and DA setup. This section also introduces the VOD observation operator. The results comparing the different experiments are described in section 3 and discussed in section 4. The main conclusions are summarized in section 5.

2. Material & methods

2.1. Model configuration and study area

For this study, we use the Noah-MP LSM (Niu et al., 2011; Yang et al., 2011) version 4.0.1, implemented within the NASA Land Information System (LIS; Kumar et al., 2006; Peters-Lidard et al., 2007). Noah-MP has been constrained by retrievals of SSM (e.g., Kumar et al., 2014; Ahmad et al., 2022; Heyvaert et al., 2023), LAI (Kumar et al., 2019; Erlingis et al., 2021; Rahman et al., 2022b; Scherrer et al., 2023), and VOD (Kumar et al., 2020, 2021) in multiple earlier studies. The model simulates soil moisture in four layers with depths 0–10 cm, 10–40 cm, 40–100 cm, and 100–200 cm, which we denote by SM_i ($i = 1, \ldots, 4$). The SSM corresponds to SM_1. The RZSM corresponds to the top 100 cm of the soil and is computed as the weighted average of the first three layers, using the thickness of the layers as weights. The dynamic vegetation option, using the maximum vegetation fraction, is enabled in the model. Vegetation carbon is simulated in four carbon mass pools: leafs, green stems, woody biomass, and fine roots. Leaf mass is directly related to LAI through a fixed specific leaf biomass parameter dependent on land cover.

The model simulations are performed from 1 April 2015 through 31 December 2022 over the Euro-Mediterranean region (ranging from 29.875°N, 11.375°W to 71.625°N, 40.125°E) on a regular latitude-longitude grid with a spatial resolution of 0.25°. Fig. S1 in the supplement provides an overview of the domain and shows that it has varying conditions in terms of soil moisture, vegetation, and land cover. The model integration time step is 15 min and daily averaged model output is used for evaluation. The experiments in this study are carried out after a thirty-year spin-up to obtain a reliable climatology of all variables. The final months of the spin-up period (1 January 2013 through 31 March

2015) also include a spin-up of the model ensemble members using the perturbation settings described in section 2.4.

Land cover, soil texture, elevation maps, and parameter tables are taken from the NASA Center for Climate Simulation (NCCS) data portal (Tian et al., 2008), as described in more detail in Scherrer et al. (2023). As meteorological forcing, the lowest model level forecasts of the fifth-generation European Center for Medium-Range Weather Forecasts (ECMWF) Reanalysis (ERA5; Hersbach et al., 2020) are used. This dataset, with a resolution of 0.25°, is matched to our grid using bilinear interpolation.

2.2. Satellite retrievals for assimilation

2.2.1. SMAP surface soil moisture

Version 8 of the SMAP Level-2 (L2) SSM retrievals (Chan et al., 2016; O'Neill et al., 2021), which are derived from L-band radiometer measurements, is used for SSM DA. These retrievals represent the top 5 cm of the soil. Both descending and ascending overpasses are used, which correspond to local collection times of 06:00 and 18:00 h, respectively. The SMAP L2 data are provided on the 36 km Equal-Area Scalable Earth Grid, Version 2.0 (EASE-Grid 2.0; Brodzik et al., 2012) which is resampled via nearest neighbor approach to the model grid before assimilation. Conservative masking is applied, meaning that only retrievals with a recommended quality (i.e., excluding observations accompanied by flags for radiofrequency interference (RFI), urban area, mountainous terrain, and dense vegetation) are used. Additionally, model-based masking is applied so that no assimilation takes place over frozen or snow-covered soils, and meteorological masking is applied so that no assimilation takes place or during rainfall events exceeding 0.0108 mm h^{-1}. Finally, retrievals are only assimilated if a sufficiently robust monthly CDF matching can be applied (Heyvaert et al., 2023). Fig. 1a shows the total amount of assimilated observations and reveals that several areas in the domain (Italy, the Alps and Carpathians, northeastern Europe, and most of Scandinavia) are masked. This means that in these areas no SMAP retrievals are assimilated and land surface estimates are purely model-based, also in the SSM DA experiment.

2.2.2. CGLS LAI

For the LAI DA, the Copernicus Global Land Service (CGLS) LAI product version 1 (Verger et al., 2014) is used, which is derived from the Ocean and Land Color Instrument (OLCI) onboard Sentinel-3 and from the Project for On-Board Autonomy - Vegetation (PROBA-V) satellite. The product is provided approximately every 10 days with a 300 m spatial resolution, and was upscaled to the model grid by averaging over model grid cells, and seasonally rescaled to the model climatology as in Scherrer et al. (2023). Fig. 1b shows that CGLS LAI retrievals are not assimilated over bare ground (e.g., the Sahara desert), where the modeled LAI is zero by definition.

In addition to this, the number of assimilated observations N_{obs} has a latitudinal dependence due to seasonal snow cover, with fewer observations over Scandinavia compared to the rest of Europe. Also, note that N_{obs} is an order of magnitude larger for SMAP SSM than for CGLS LAI due to the 10-day resolution of the latter.

2.2.3. AMSR2 VOD

For the VOD DA, the Advanced Microwave Scanning Radiometer 2 (AMSR2) X-band VOD (10.65 GHz) retrievals of the Land Parameter Data Record algorithm (LPDR; Du et al., 2017a,b) version 3 are used. Data are provided as daily images on a global EASEv2.0-Grid format with 25 km spatial resolution. For consistency with the model output, the data were resampled to a 0.25° regular grid. Only night-time retrievals are used, as these are typically more robust (e.g., Moesinger et al., 2020, 2022) and are often assumed to represent an equilibrium state of plant hydraulics (Liu et al., 2021). This corresponds to the descending overpass at 01:30 h local time. Since VOD is not a prognostic or diagnostic variable modeled by Noah-MP, an observation operator was developed to map the modeled RZSM and LAI to VOD. This is discussed in the next section.

Fig. 1c shows that no observations are assimilated in England and Italy, mainly due to missing retrievals caused by RFI. Parts of Russia and Ukraine are also masked due to the poor performance of the observation operator in these regions (section 2.3). Similarly to CGLS LAI, no retrievals are assimilated over bare ground.

2.3. VOD observation operator

Kumar et al. (2020) - and based on this study, Mucia et al. (2022) - assimilated VOD by monthly CDF-matching VOD observations to LAI values of a reference product. That is, they created 12 different observation operators, one for each month. This approach implicitly removes the mean seasonal cycle of the VOD observations, and replaces it with the reference LAI's mean seasonal cycle. Therefore, the approach (i) only uses the anomaly information of the VOD observations and (ii) is only unbiased if the reference LAI is unbiased with respect to the model. The latter is not always the case, as shown in Scherrer et al. (2023), for example.

Our approach differs from Kumar et al. (2020) in three main aspects: (i) instead of creating separate observation operators for each month, we create a single model valid for all months, (ii) we do not use an external 'reference LAI' product but only modeled LAI and therefore obtain an unbiased DA system, and (iii) we also account for the dependence of VOD anomalies on RZSM anomalies. Both dependencies are modeled via a linear dependence as follows:

$$\mathrm{VOD}' = \alpha \cdot \mathrm{LAI}' + \beta(\mathrm{LAI}, \mathrm{RZSM}) \cdot \mathrm{RZSM}', \tag{1}$$

where the prime symbol denotes the anomalies, α is a long-term time-invariant fitted parameter and β(LAI, RZSM) is a time- and state-dependent coefficient that links the night-time VOD anomalies and RZSM anomalies in Equation (1) based on current RZSM and LAI:

$$\beta(\mathrm{LAI}, \mathrm{RZSM}) = \beta_0 + \beta_1 \cdot \mathrm{LAI} + \beta_2 \cdot \mathrm{RZSM} + \beta_3 \cdot \mathrm{RZSM} \cdot \mathrm{LAI}, \tag{2}$$

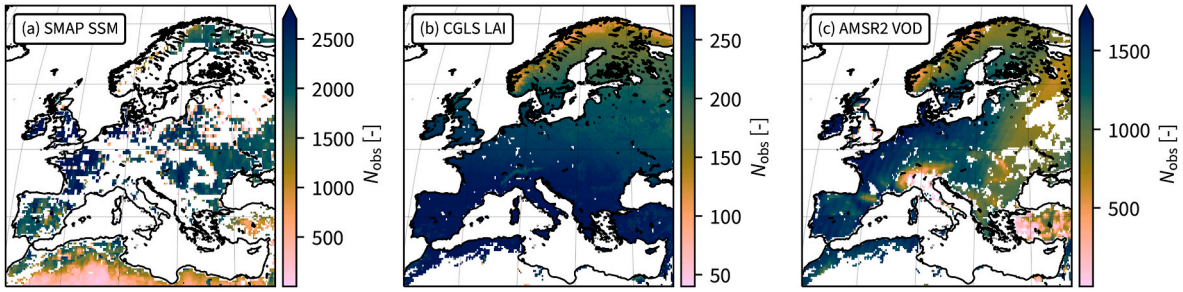

Fig. 1. Number of assimilated observations N_{obs} for the three satellite products between 1 April 2015 and 31 December 2022: (a) SMAP SSM, (b) CGLS LAI, and (c) AMSR2 VOD. No assimilation is performed over the white masked grid cells.

with long-term time-invariant fitted parameters β_i.

In situations with little vegetation, VOD is more sensitive to the effects of rain, such as interception (Vermunt et al., 2020) and surface ponding or flooding (Bousquet et al., 2021). The latter can lead to strong decreases in VOD; therefore, we expect a negative relationship between VOD anomalies and RZSM anomalies for low LAI and high RZSM. In contrast, for high LAI, when the vegetation is fully developed and active, the VOD is more sensitive to vegetation water content (VWC), and night-time VWC is in equilibrium with RZSM (or more precisely, the leaf water potential is in equilibrium with the root water potential) (Konings and Gentine, 2017; Konings et al., 2019; Liu et al., 2021). Therefore, we expect a positive relationship between VOD anomalies and RZSM anomalies for high LAI. This effect will be stronger at lower RZSM because of saturation effects in the relationship of moisture content and water potential at high moisture contents, hence the inclusion of an interaction term. Fig. S2 in the supplement presents the changes of β as a function of LAI and RZSM for a specific grid cell in the domain. As expected, the interaction effect is negative at low LAI and high RZSM, and positive at high LAI and low RZSM.

Inserting Equation (2) into Equation (1) leads to a linear model with interaction effects. The parameters α and β_i have been calibrated for each grid cell separately using VOD retrievals and model output for LAI and RZSM over the period July 2012 (start of availability of AMSR2 observations) through December 2022. The anomalies of the modeled RZSM and LAI used in Equation (1) are calculated by subtracting the long-term mean seasonal cycle of the model over the same period. Similarly, VOD (bulk signal) estimates are obtained by adding the long-term seasonal cycle to the VOD anomaly estimates. Fig. 2 shows evaluation metrics of the fit of the observation operator using VOD retrievals that were not used in the observation operator calibration. Over large parts of the domain, the coefficient of determination R^2 scores higher than 0.7, which means that the operator can explain more than 70% of the total variance. Lower scores are found in mountainous areas and in northern Turkey and Finland. The median root mean square error (RMSE) of the observation operator is 0.045 [-], but there is a strong east-west gradient, with good performance in western Europe and poor performance in eastern Europe. Furthermore, northern Turkey has a high RMSE, which could be caused by RFI. The correlation of anomalies, R_{anom}, is higher than 0.5 in most of the domain, with lower scores in Scandinavia, the Alps, and Turkey.

To avoid assimilating low-quality retrievals or in areas where the observation operator does not perform well, we mask pixels where $R^2 \leq 0.5$ [--] and RMSE ≥ 0.08 [-]. This occurs mainly in some larger areas of western Ukraine and Russia (Fig. 1c).

Note that the RMSE is used here to quantify the explicit model *error*, whereas the evaluation of simulation results below will focus on *differences* with reference datasets. This is why we will use the term root mean square deviation (RMSD) in the remainder of the text.

2.4. Data assimilation

Six experiments are carried out in total and are listed in Table 1: one open loop (OL; model-only) and five DA experiments. In the DA_{SSM}, DA_{LAI}, and DA_{VOD} experiments, retrievals of SMAP L2 SSM, CGLS LAI, and AMSR2 X-band VOD are assimilated, respectively. We refer to them as the 'single-sensor' DA experiments since only one satellite product is assimilated in each of them. In the remaining two experiments a 'joint' DA of SMAP L2 SSM retrievals and vegetation retrievals from CGLS LAI ($DA_{SSM+LAI}$) or AMSR2 VOD ($DA_{SSM+VOD}$) is performed. All six experiments use an ensemble of 24 members and the same perturbation scheme as listed in Table 2. The perturbations represent the uncertainty in the forcings and the state variables of the model. Perturbations to precipitation and downward short-wave radiation are applied multiplicatively and are drawn from lognormal distributions, while perturbations to downward long-wave radiation and the model prognostic state variables are drawn from a normal distribution and applied additively. Perturbations are truncated if they exceed 2.5 standard deviations. Temporal autocorrelations are only applied to forcings, using a first-order autoregressive model with a time scale of 24 h. Spatial correlations are not applied to the perturbations. Settings for the cross-correlations of errors in the forcing and soil moisture state variables are taken from Kumar et al. (2014), while the cross-correlations between the soil moisture and LAI state variables are based on the anomaly correlations between these variables in a deterministic model run with the same domain and temporal coverage as our experiments. Perturbations in prognostic variables are not cross-correlated with those of forcings. Any unintentional perturbation bias in the forecasts is removed using a correction described by Ryu et al. (2009).

The five DA experiments use a one-dimensional ensemble Kalman Filter (EnKF; Reichle et al., 2002; Evensen, 2003) to assimilate satellite retrievals into the LSM. Whenever an observation is available at a specific location, the EnKF updates the state variables to reduce the deviation between model forecasts and observations. For experiments DA_{SSM} and DA_{LAI}, the four soil moisture state variables or the LAI state

Table 1
Overview of the performed experiments, indicating for each of them which of the three satellite product(s) is (are) assimilated, and whether the soil moisture SM_i ($i = 1, \ldots, 4$) and/or the LAI state variables are updated.

experiment	assimilated products			updated variables	
	SMAP SSM	CGLS LAI	AMSR2 VOD	SM_i	LAI
OL					
DA_{SSM}	•			•	
DA_{LAI}		•			•
DA_{VOD}			•	•	•
$DA_{SSM+LAI}$	•	•		•	•
$DA_{SSM+VOD}$	•		•	•	•

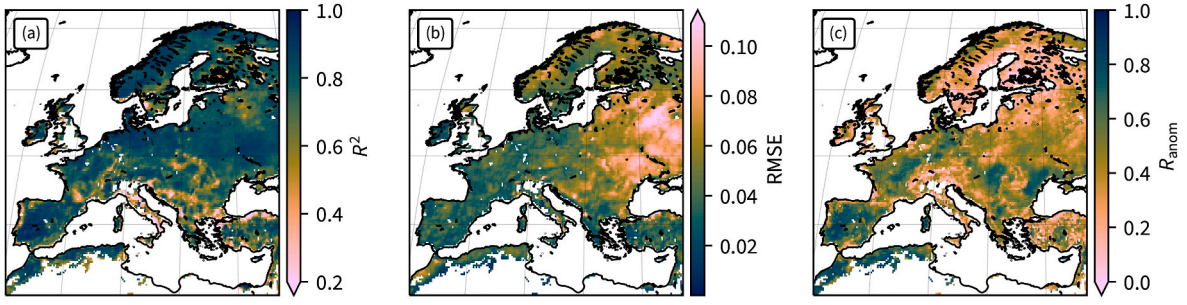

Fig. 2. Evaluation of the VOD observation operator model fit in terms of (a) the coefficient of determination R^2 [-], (b) the root mean square error RMSE [-], and (c) the anomaly correlation R_{anom} [-] between model predictions and observation data. Darker colors indicate better performance.

Table 2
Overview of the ensemble model perturbations, valid for all experiments. Perturbed forcing fields are the downward short-wave radiation (SW), downward long-wave radiation (LW), and precipitation (*P*). For the prognostic variables, soil moisture in the four soil layers (SM_i, $i = 1, \ldots, 4$) and LAI are perturbed. Additive (+) perturbations have a mean of 0 and are drawn from a normal distribution, while multiplicative (×) perturbations have a mean of 1 and are drawn from a lognormal distribution. Temporal correlations of the time series are applied through a first-order autoregressive model. Perturbations to prognostic variables are not correlated with forcing perturbations.

	type	mean	standard deviation	temporal correlation	cross-correlations with other perturbations							
					SW	LW	*P*	SM_1	SM_2	SM_3	SM_4	LAI
SW	×	1	0.3	24 h		−0.50	−0.80					
LW	+	0	50 W m^{-2}	24 h	−0.50		0.50					
P	×	1	0.5	24 h	−0.80	0.50						
SM_1	+	0	0.00400 m^3m^{-3}	0					0.60	0.40	0.20	0.08
SM_2	+	0	0.00007 m^3m^{-3}	0				0.60		0.60	0.40	0.13
SM_3	+	0	0.00004 m^3m^{-3}	0				0.40	0.60		0.60	0.19
SM_4	+	0	0.00002 m^3m^{-3}	0				0.20	0.40	0.60		0.10
LAI	+	0	0.01	0				0.08	0.13	0.19	0.10	

variables are updated respectively, as shown in Table 1. For the DA_{VOD} and joint DA experiments, both soil moisture and LAI state variables are updated. In the case of DA_{VOD}, the multivariate update is driven by the observation operator used to map the LAI and RZSM simulated by Noah-MP to VOD (section 2.3). For the joint DA experiments, the multivariate update is a natural consequence of our intention to use multi-sensor information to improve soil moisture and vegetation state variables simultaneously.

The uncertainty in the observations is tuned by optimizing innovation diagnostics (Reichle et al., 2002) through Desroziers' metrics (Desroziers et al., 2005) and making a systematic comparison of different perturbation sizes as was done in Heyvaert et al. (2023). The perturbations of all observations are applied additively with standard deviations of 0.0125 m^3 m^{-3} for the SMAP SSM retrievals, 0.05 [-] for the CGLS LAI retrievals, and 0.03 [-] for the AMSR2 VOD retrievals. Because these sensors are all based on different wavelengths (respectively, L-band microwave, X-band microwave, and optical), no cross-correlations between the observation errors are considered.

The timing of state updates in the DA experiments coincides with the overpass times of the satellites, as pictured in Fig. 3 for the $DA_{SSM+VOD}$ experiment: 01:30 h local time for the AMSR2 VOD retrievals, and 06:00 h and 18:00 h for the SMAP SSM retrievals. Note that observations from either sensor may occasionally be masked, meaning that two SMAP retrievals do not necessarily follow each AMSR2 retrieval in reality.

The EnKF is a method to correct random errors in the model and observations, but should not be used to correct for systematic deviations (bias) (Baek et al., 2006). To obtain an unbiased DA system, a monthly rescaling of the observations to the model climatology is performed through CDF matching (Reichle and Koster, 2004; Barbu et al., 2014; Heyvaert et al., 2023) in the case of SSM DA. For the LAI DA, a seasonal rescaling is performed as described in Scherrer et al. (2023). For the VOD DA, the bias is removed during the construction of the observation operator, guaranteeing an unbiased DA system. Therefore, effectively only anomalies from the mean seasonal cycle are assimilated.

2.5. Impact analysis

In order to visualize the impact of the respective DA experiments on geophysical estimates, maps are created that show the RMSD between the OL experiment on the one hand and the various DA experiments (listed in Table 1) on the other. The RMSD between both time series of length *N* days is defined as

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i^{\mathrm{DA}} - x_i^{\mathrm{OL}}\right)^2}, \tag{3}$$

with x_i^{DA} (x_i^{OL}) the ensemble mean realizations of a geophysical variable, e.g., SSM or LAI, estimated by the DA (OL) experiment at day *i*. A larger RMSD between OL and DA signifies larger deviations from the model by assimilating the satellite retrievals, i.e., a stronger impact of the DA on the land surface estimates.

Fig. 3. Depiction of which retrieval is assimilated at what time in the $DA_{SSM+VOD}$ experiment.

2.6. Evaluation with reference data

The performance of the different DA setups is evaluated by comparing the results with independent reference data. Since we use an unbiased DA system, our objective is to improve the representation of anomalies in the DA analysis compared to the OL experiment. These improvements are measured with the anomaly correlation R_{anom}, i.e., the Pearson correlation between anomalies from the mean seasonal cycles of observations and analysis. Only evaluation data in grid cells without masking for any of the three satellite products (Fig. 1) are used to ensure a fair comparison of the different DA experiments. A two-sided Wilcoxon signed-rank test is performed for each evaluation to find whether the distributions of R_{anom} from the DA experiments are significantly different from those of the OL experiment. The test was not performed for evaluations with X-BASE data (see section 2.6.2) because one of the assumptions of the statistical test (independent samples) is not fulfilled when using this product, due to spatial autocorrelation between the gridded values. In addition to the evaluation with R_{anom}, we also evaluate in terms of RMSD for the flux variables to assess whether DA manages to improve their absolute values.

2.6.1. International Soil Moisture Network

SSM and RZSM are evaluated with in situ data from the International Soil Moisture Network (ISMN; Dorigo et al., 2011; 2013; 2021b). This database provides ground station measurements of soil moisture at various depths for locations spread throughout the study domain. When multiple sites are available within one grid cell, the skill metrics are computed for each of the sites, after which the metrics are averaged to obtain a single value for the grid cell to avoid giving too much weight to such areas in the analysis.

Only measurements that have been flagged as 'good' are used in the evaluation. We require each site to have at least 200 observations (after daily aggregation) over at least three different years to be used in the evaluation. Sites should have at least one measurement deeper than 10 cm for them to be used in the evaluation of RZSM. In total, 38 sites are

used in the SSM evaluation and 17 in the RZSM evaluation. An overview of all the networks used and their references is provided in the supplement, Table S1.

2.6.2. Warm Winter 2020 and X-BASE

Vegetation-related fluxes are evaluated using in situ measurements from the Warm Winter 2020 database (Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre, 2022). These provide, among other variables, estimates of evapotranspiration (ET), gross primary production (GPP), and net ecosystem exchange (NEE), based on flux tower stations through eddy covariance measurements. For evaluation, the daily means provided in the data product are compared with the daily mean model output.

The same preprocessing steps that are applied to the ISMN in situ data are also applied to station data from the Warm Winter 2020 database. Of the 73 stations in the database, only 10 can be used for the evaluation of ET and GPP in model grid cells where data from all three satellite products are assimilated. Only one NEE station is available where all three satellite products are assimilated, so the Warm Winter 2020 database is not used for the evaluation of NEE. Additionally, the data are available only until the end of 2020, whereas our experiments run until the end of 2022. Therefore, we also use the X-BASE dataset (Nelson et al., 2023, 2024) for the evaluation. These data are based on eddy covariance flux towers and remotely sensed observations, which are linked by machine learning techniques to provide gridded estimates of ET, GPP, and NEE. The version we use has a temporal resolution of one day and uses a 0.25° regular grid, which is transformed into the model grid through bilinear interpolation.

2.7. Data assimilation diagnostics

In this study, the ensemble spread is defined as the range of the ensemble, i.e., the difference between the maximum and minimum values of the ensemble members. Its temporal average is computed for the SSM, RZSM, and LAI state variables. The spread for each DA experiment (section 2.4) can then be compared with that of the OL. A reduction in the mean ensemble spread for the DA compared to the OL reflects a reduction in the uncertainty of the estimate as a result of the analysis update (Reichle et al., 2017).

The observation-minus-forecast residuals, or innovations, denote the difference between rescaled satellite retrieval and model estimate in 'observation space'. Therefore, the rescaled SMAP SSM and CGLS LAI retrievals are compared to the modeled SM_1 and LAI, respectively, whereas a VOD model estimate must be derived from modeled LAI and RZSM using the observation operator (section 2.3) before a comparison with the AMSR2 VOD retrievals can be made.

The analysis-minus-forecast residuals, or increments, are the equivalent of the innovations in the 'model space'. SMAP SSM assimilation results in increments of SM_i ($i = 1, \ldots, 4$), whereas VOD assimilation results in increments of SM_i ($i = 1, \ldots, 4$), and LAI. We define an 'SSM increment' as an increment of SM_1, and an 'RZSM increment' as the weighted average of soil moisture increments in the top three layers of the model, with the thickness of the layers as weights.

3. Results

3.1. Impact analysis

Fig. 4a shows the impact of the five different DA experiments on SSM using the RMSD between the OL and DA experiments. The three single-sensor DA experiments in the upper row show that DA_{LAI} does not have a substantial impact on SSM compared to DA_{SSM} or DA_{VOD} (mean RMSD of 0.001 m^3m^{-3} compared to 0.013 m^3m^{-3} and 0.010 m^3m^{-3}, respectively). Note that DA_{LAI} is the only DA experiment in which soil moisture is not explicitly updated (Table 1), and updates in LAI do not propagate much to the SSM via the model. The impact of DA_{LAI} on RZSM is slightly

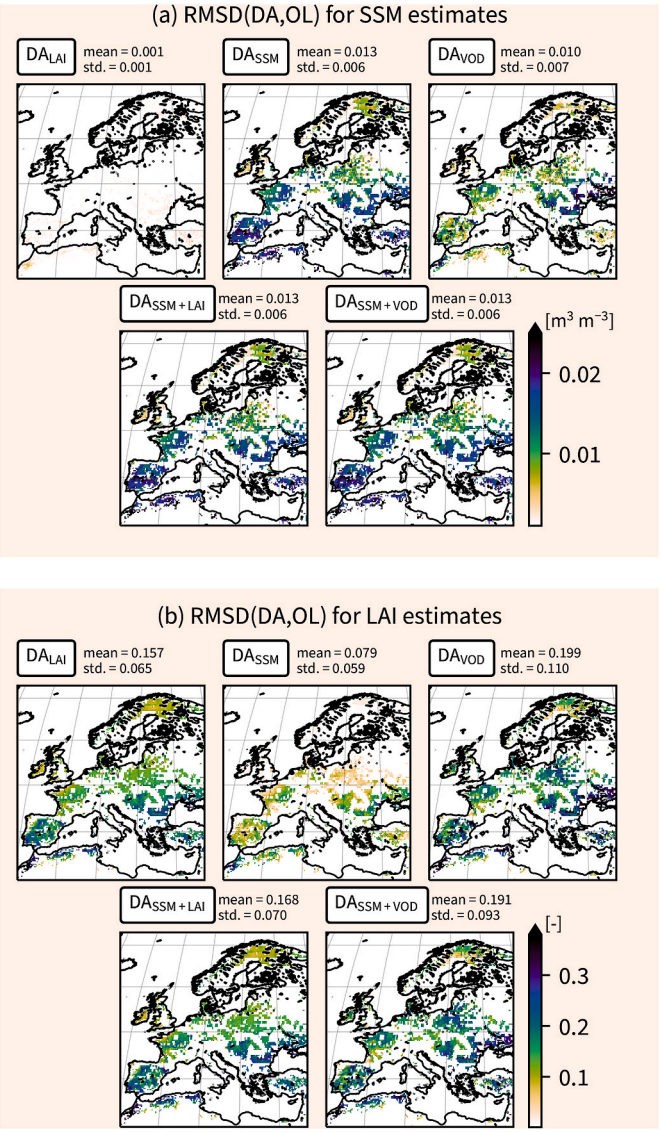

Fig. 4. (a) RMSD between the estimated SSM in the OL and the DA experiments indicated above the panels. The mean and standard deviation of the maps are computed over the non-masked pixels. (b) As in (a), but for estimates of LAI.

higher (Fig. S3 in the supplement). DA_{SSM} affects SSM estimates more in lower latitudes, that is, for drier areas such as Spain and Turkey, and over southern Ukraine. In this last region, the DA_{VOD} also impacts SSM, but the impact of DA_{VOD} is generally lower in other areas compared to DA_{SSM}. The joint DA experiments, $DA_{SSM+LAI}$ and $DA_{SSM+VOD}$, show similar results as DA_{SSM} in terms of impact on SSM (same mean, standard deviation, and spatial distribution).

The same analysis is performed for the LAI estimates in Fig. 4b. While LAI is most impacted by DA_{LAI} and DA_{VOD} (mean RMSD of 0.157 and 0.199 respectively), there is also a relatively high impact of DA_{SSM} (mean RMSD of 0.079), even though LAI is not explicitly updated in DA_{SSM} (Table 1). This impact is thus coming purely from model propagation. The three single-sensor experiments in the upper row show a similar spatial pattern, with the largest DA impact on LAI in eastern Europe (Ukraine) and, to a lesser extent, the Iberian peninsula. The impact map of $DA_{SSM+LAI}$ resembles that of DA_{LAI} and the map of $DA_{SSM+VOD}$ resembles that of DA_{VOD}. This means that for LAI, the dominant impact in the joint DA experiments comes from the assimilation of vegetation retrievals.

In summary, both joint DA experiments resemble the SSM-only DA

experiment in terms of impact on SSM estimates, and the vegetation-only DA experiment in terms of LAI estimates. The impact analyses are also performed for RZSM, GPP, ET, and runoff (Figs. S3–S6 in the supplement). The results for RZSM, ET, and runoff are comparable to those of SSM (Fig. 4a) with most impact from DA_{SSM}, and lower impact from DA_{VOD} and DA_{LAI}. The impact of the joint experiments closely resembles that of DA_{SSM}. For GPP, the impact of the joint DA depends on which vegetation product is assimilated: the map of $DA_{SSM+LAI}$ corresponds to that of DA_{SSM}, while the map of $DA_{SSM+VOD}$ corresponds to that of DA_{VOD}.

3.2. Evaluation with reference data

The performance of the different DA experiments is shown in Table 3 for a number of prognostic and diagnostic variables. The spatial median of the improvement (positive values) or degradation (negative values) in skill with respect to the OL, i.e. ΔR_{anom}, is computed using independent data at sparse in situ locations or from satellite products. For SSM and RZSM, a comparison with ISMN stations shows the highest median improvements for the DA_{SSM} experiment. Assimilating vegetation products does not show any impact (DA_{LAI}) or degrades soil moisture estimates at in situ sites (DA_{VOD}). Both $DA_{SSM+LAI}$ and $DA_{SSM+VOD}$ show improvements in soil moisture estimates as well, although slightly smaller than for the single-sensor DA_{SSM} experiment. Note that the number of available stations is substantially smaller for RZSM than for SSM, hence the statistics of the latter are more reliable. However, the findings of both are consistent with each other.

Fig. 5a–e shows the spatial distribution of skill changes at the in situ stations and Fig. 5f compares the boxplots of the distributions for the different DA experiments. Only when SSM DA is included, small overall improvements are found with respect to the OL, and particularly in eastern Europe. This is consistent with the findings of Heyvaert et al. (2023) who assimilate the ESA CCI Soil Moisture product, which contains SMAP SSM retrievals, among others (Dorigo et al., 2017; Gruber et al., 2019). The spatial distribution of the ISMN stations for RZSM is presented in the supplementary material, Fig. S7.

The evaluation results for GPP are depicted in the third and fourth rows of Table 3 for the Warm Winter 2020 and X-BASE evaluation data, respectively. Over the ten Warm Winter 2020 stations, we find degradations for each DA experiment except DA_{LAI}, which is also the experiment that has the largest improvements in GPP when evaluating with X-BASE. Fig. 6a shows that these improvements occur throughout the domain and are largest in southern Europe. DA_{SSM} overall slightly degrades the estimates of GPP, especially in Ukraine and northern Spain (Fig. 6b). Assimilating both SSM and LAI ($DA_{SSM+LAI}$) results in slightly smaller improvements than assimilating only LAI (Table 3), mainly due to small degradations in the areas mentioned above (Fig. 6d). Overall, the spatial ΔR_{anom} distribution of $DA_{SSM+LAI}$ resembles that of DA_{LAI}, with locations at the lower end of the distribution performing slightly better for the single-sensor DA_{LAI} experiment (Fig. 6f). Although DA_{VOD} performs better overall than DA_{SSM} in terms of GPP, Fig. 6c highlights some areas where it degrades the results compared to OL, particularly the Pannonian basin, Romania, and Ukraine. The results improve slightly in these first two areas when SSM is assimilated alongside VOD ($DA_{SSM+VOD}$, Fig. 6e). Similar conclusions can be drawn from an evaluation in terms of RMSD, with the largest improvements for the DA_{LAI} ($\Delta RMSD = -0.024$ gC m^{-2} day^{-1}) and $DA_{SSM+LAI}$ ($\Delta RMSD = -0.022$ gC m^{-2} day^{-1}) experiments (Fig. S8).

The evaluation results for NEE are similar to those for GPP, with the largest improvements in the experiments in which LAI is assimilated. Assimilating VOD slightly degrades the estimates. Figs. S9–S10 show that the spatial patterns of the NEE improvements are similar to those of the GPP, both in terms of R_{anom} and RMSD.

Finally, for ET, we find that the impact of any DA experiment is relatively small in terms of R_{anom}, both for an evaluation with the ten stations of the Warm Winter 2020 dataset and with the X-BASE product (last two rows of Table 3). As was the case for GPP and NEE, the best overall performance is achieved when a vegetation product is assimilated: both DA_{LAI} and DA_{VOD} show small improvements when evaluating with X-BASE data. DA_{LAI} has a relatively small impact throughout the domain, whereas the effect of DA_{VOD} is quite different throughout the domain, with some areas showing stronger improvements and others showing degradations, especially in southern Ukraine (Fig. S11). Assimilating SSM slightly degrades ET results, both in the single-sensor DA_{SSM} and the joint $DA_{SSM+LAI}$ and $DA_{SSM+VOD}$ experiments. The evaluation in terms of RMSD is similar, with stronger degradations for DA_{VOD} (Fig. S12).

A remarkable result from Table 3 may be that DA_{SSM} improves the SSM and RZSM estimates substantially, yet slightly degrades turbulent fluxes from the land surface. This may be attributed to imperfect coupling mechanisms (e.g., errors in the partitioning between transpiration and evaporation) in the LSM, something that has been raised by multiple previous studies (Dong et al., 2022; Zhou et al., 2023; Crow et al., 2024).

In summary, soil moisture estimates are most improved by DA_{SSM}, whereas estimates of GPP, NEE, and ET are most improved by assimilating a vegetation product (DA_{LAI} in particular). The joint $DA_{SSM+LAI}$ does not have the highest performance for any specific variable, but is capable of capturing the improvements of both single-sensor DA experiments (DA_{SSM} and DA_{LAI}) into a single physically consistent product.

3.3. Ensemble spread

The impact of DA on the ensemble spread is visualized in Table 4. It shows how the spread for SSM, RZSM, and LAI differs for the various DA experiments with respect to the OL.

For SSM, the average OL ensemble spread over the domain is 0.065 m^3 m^{-3}, with the largest spreads in drier areas such as the Iberian peninsula, Anatolia, and northern Africa (see Fig. S13a for the spatial pattern). For RZSM, the pattern is quite different, with smaller spreads over the sandy soils of the Sahara desert and mountainous areas (Alps, Carpathians), and larger spreads over Ukraine (Fig. S13g). Its spatially averaged value is 0.028 m^3 m^{-3}; this smaller spread for the root zone compared to the surface can be explained by the smaller perturbations

Table 3
Median ΔR_{anom} for the different DA experiments with respect to OL. The experiment with the largest improvement for each geophysical variable is indicated in bold. Only model grid cells for which all three products are assimilated are included in the computation of the metrics.

	Product	period	N_{sites}	skill improvement (median ΔR_{anom})				
				DA_{SSM}	DA_{LAI}	DA_{VOD}	$DA_{SSM+LAI}$	$DA_{SSM+VOD}$
SSM	ISMN	2015–2022	38	**0.016***	0.000	−0.011*	0.013*	**0.016**
RZSM	ISMN	2015–2022	17	**0.015**	0.001	−0.017*	0.011	0.005
GPP	Warm Winter 2020	2015–2020	10	−0.002	**0.000**	−0.023	−0.002	−0.016
	X-BASE	2015–2021		−0.004	**0.022**	−0.001	0.014	−0.002
NEE	X-BASE	2015–2021		−0.003	**0.009**	−0.006	0.002	−0.004
ET	Warm Winter 2020	2015–2020	10	**−0.001**	**−0.001**	−0.007	−0.003	−0.011
	X-BASE	2015–2021		−0.003	**0.003**	0.001	−0.003	−0.002

*$p < 0.05$ (two-sided Wilcoxon signed-rank test, not performed for X-BASE).

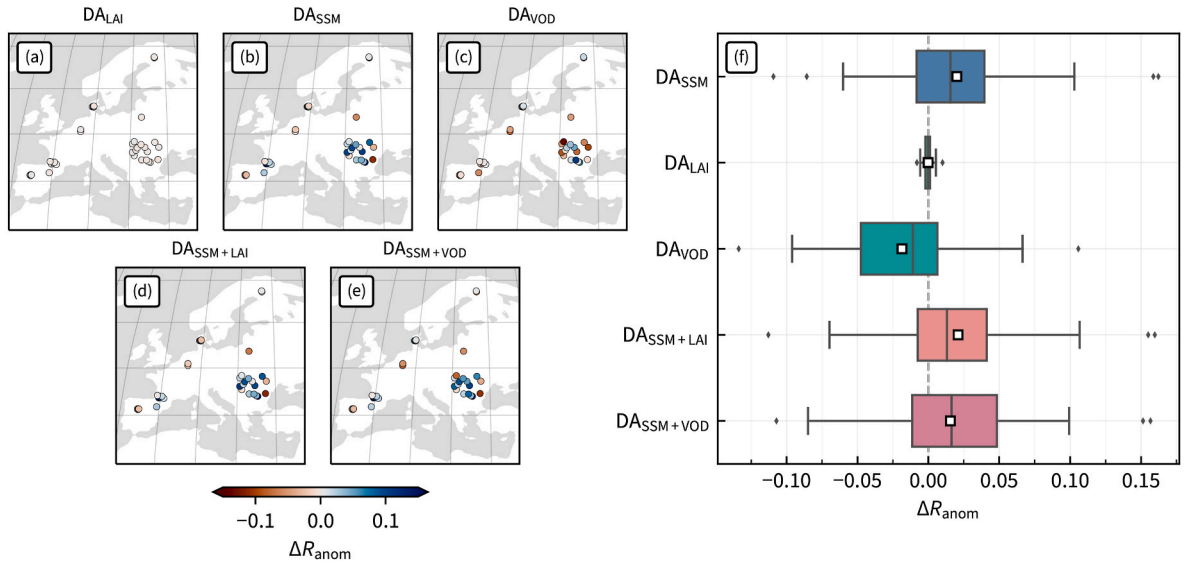

Fig. 5. DA skill difference ΔR_{anom} relative to the OL for SSM at the ISMN stations within the region where all three satellite products are assimilated, for (a) DA_{SSM}, (b) DA_{LAI}, (c) DA_{VOD}, (d) $DA_{SSM+LAI}$, and (e) $DA_{SSM+VOD}$ experiments. Blue indicates improvements by the DA, whereas red indicates degradations. (f) Corresponding boxplots summarizing the spatial distributions for the different experiments. The central line denotes the median and the white square the mean of each distribution.

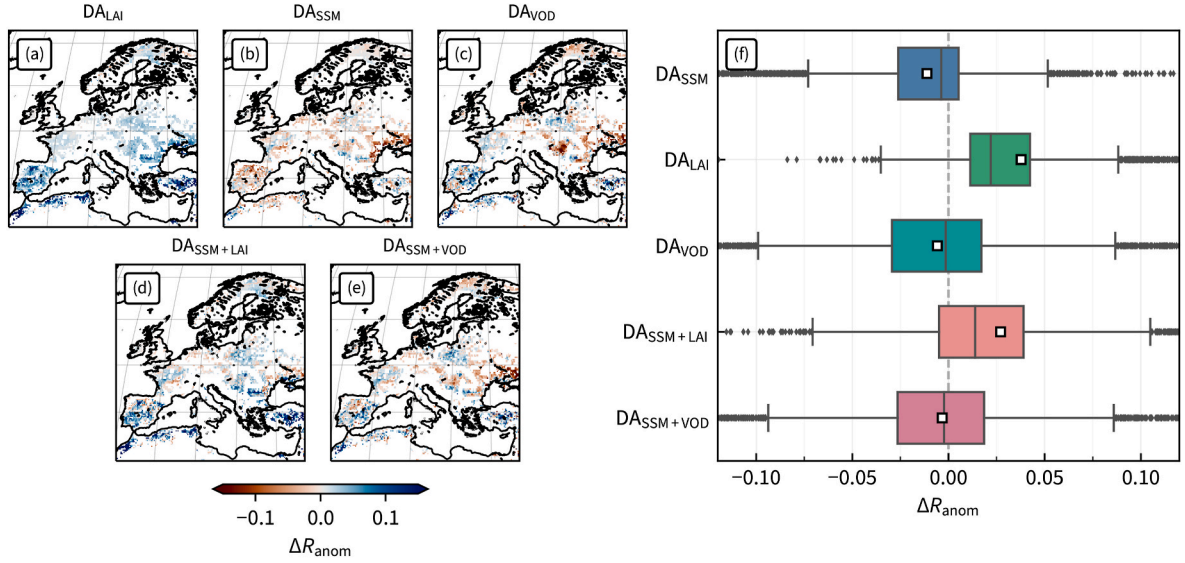

Fig. 6. Same as Fig. 5, but now using gridded X-BASE GPP as reference data.

Table 4
Relative mean ensemble spread of the DA experiments with respect to the mean ensemble spread of the OL. Values smaller than one indicate that the spread is smaller for the DA. The experiment with the smallest spread is indicated in bold for each variable. Only pixels where observations from all three sensors are assimilated are considered.

	DA_{SSM}	DA_{LAI}	DA_{VOD}	$DA_{SSM+LAI}$	$DA_{SSM+VOD}$
SSM	**0.79**	1.00	0.96	0.80	**0.79**
RZSM	0.62	1.01	0.86	0.62	**0.60**
LAI	0.90	0.76	0.74	0.74	**0.71**

applied to soil moisture in the deeper layers (Table 2) and the lower sensitivity to atmospheric variables. Table 4 shows that the ensemble spread corresponding to DA_{SSM} is only 79% (62%) that of the OL for SSM (RZSM), whereas DA_{LAI} does not impact SSM and RZSM spread, and DA_{VOD} mainly decreases the spread in the RZSM. The joint DA experiments lead to similar spreads in SSM and RZSM as single-sensor DA_{SSM}, with a slightly stronger spread reduction for $DA_{SSM+VOD}$, making it the experiment with the least uncertainty in soil moisture estimates.

In the case of LAI, we find the largest spreads in the same dry areas as for SSM, except for the desert where Noah-MP LAI is zero by definition (Fig. S13m). For single-sensor DA, the relative spread is the smallest for DA_{VOD} (0.74) and DA_{LAI} (0.76), but also DA_{SSM} has a considerably smaller spread compared to OL (0.90). The slightly larger reduction in spread with DA_{VOD} could be an effect of more frequent updates, since VOD retrievals are available almost daily, while LAI retrievals are available only every ten days. As is the case for SSM and RZSM, the largest decrease in the average spread can be found in the joint DA experiments, with $DA_{SSM+VOD}$ showing the strongest reduction in spread.

Summarized, the DA_{SSM} mostly decreases the uncertainty in the SSM and RZSM estimates, and DA_{LAI} and DA_{VOD} mostly decrease the uncertainty in LAI estimates. The joint $DA_{SSM+LAI}$ and $DA_{SSM+VOD}$ experiments

are able to combine both in such a way that the uncertainty in all variables is maximally decreased.

3.4. Innovations and increments

3.4.1. $DA_{SSM+LAI}$

This section compares the innovations between the joint $DA_{SSM+LAI}$ and single-sensor DA_{SSM} and DA_{LAI}. Fig. 7a shows no clear difference between the SSM innovations from the joint $DA_{SSM+LAI}$ and single-sensor DA_{SSM} experiment, with the time series of the difference of their absolute values closely resembling a zero-mean white-noise sequence. Fig. 7b, visualizing the temporally averaged absolute differences between both experiments, also shows a zero-mean noisy pattern. Assimilating CGLS LAI alongside SMAP SSM hence does not alter SSM innovations in any systematic way.

The two bottom panels, Fig. 7c and d, show the results of the LAI innovations. Innovations in the joint DA are on average smaller than in the single-sensor DA_{LAI}. This means that the LAI estimates are closer to the observed CGLS LAI retrievals when SMAP SSM is assimilated alongside CGLS LAI. Some periods, such as the growing seasons of 2017, 2018 and 2020, show larger innovations for the joint DA. Spatially, we find that the smaller innovations for the joint DA are predominantly located in southern latitudes.

A similar figure comparing increments, rather than innovations, between $DA_{SSM+LAI}$ and the single-sensor DA experiments is presented in the supplement (Fig. S14). Unsurprisingly, differences in SSM and RZSM increments (coming from the SSM innovations) are not systematically different between the joint $DA_{SSM+LAI}$ and single-sensor DA_{SSM} experiment. Likewise, the behavior for the LAI increments closely resembles that of the LAI innovations.

3.4.2. $DA_{SSM+VOD}$

In contrast to the findings of $DA_{SSM+LAI}$, Fig. 8a and b shows that the SSM innovations are on average slightly larger for $DA_{SSM+VOD}$ than for DA_{SSM}, especially in southern Ukraine. This result implies that assimilating AMSR2 VOD retrievals moves the SSM estimates further away from the SMAP SSM retrievals. As can be seen in Fig. S15, this also translates in larger increments for SSM and RZSM from the SMAP SSM retrievals in $DA_{SSM+LAI}$, i.e., the DA is making larger corrections to the soil moisture state variables in the joint DA. This is confirmed by the strong positive 'spikes' after the two outages of the SMAP instrument (19 June 2019–23 July 2019; 20 September 2022–6 October 2022), indicating how the analysis deviates from the SMAP SSM retrievals when only AMSR2 VOD is being assimilated. However, at most other times, we find that the differences in the SSM innovations are relatively small.

Innovations of VOD show a very pronounced difference between $DA_{SSM+VOD}$ and DA_{VOD} in Fig. 8c and d: they are substantially larger in the joint than in the single-sensor DA, particularly in the second half of the year and in the east of Europe (up to 40% larger innovations here, with an average of 20% over the domain as a whole). Assimilation of SMAP SSM thus introduces larger differences between the observed and estimated VOD.

Innovations in VOD are transformed to increments of soil moisture SM_i ($i = 1, \ldots 4$, summarized by means of the RZSM) and LAI in the state space. Fig. 9 shows how the absolute values of the increments compare between the joint $DA_{SSM+VOD}$ and single-sensor DA_{VOD}. Increments of LAI tend to be larger in the joint DA, as is the case for the VOD innovations, whereas increments of RZSM are actually smaller in the joint DA: on average 48% smaller throughout the domain.

Unique to the $DA_{SSM+VOD}$ experiment is that increments of the RZSM come from two sources: the SSM innovations from SMAP (at 6:00 and 18:00 h local time) and the VOD innovations from AMSR2 (at 1:30 h local time). Fig. 10 computes the temporal correlation between the increments introduced by both sensors. Only days for which increments from both sensors are available are considered in the computation of the correlation. It shows that increments introduced by both sources are anticorrelated, meaning they tend to 'counteract' each other in the joint DA system. This is particularly the case in the east of the domain (southern Ukraine), an area that also appears in maps comparing innovation magnitudes of $DA_{SSM+VOD}$ and single-sensor experiments (Fig. 8).

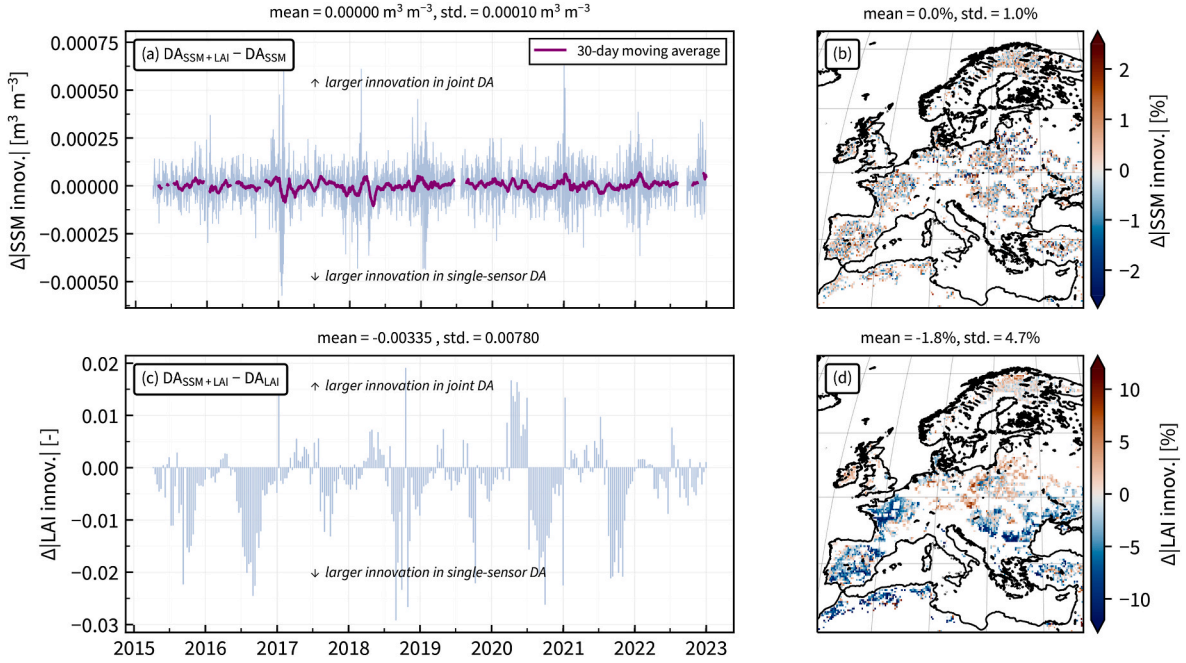

Fig. 7. Overview of differences in absolute values of innovations between $DA_{SSM+LAI}$ and (a)–(b) DA_{SSM} or (c)–(d) DA_{LAI}. Left column: time series of the difference of spatial averages, right column: maps of the difference of temporal averages.

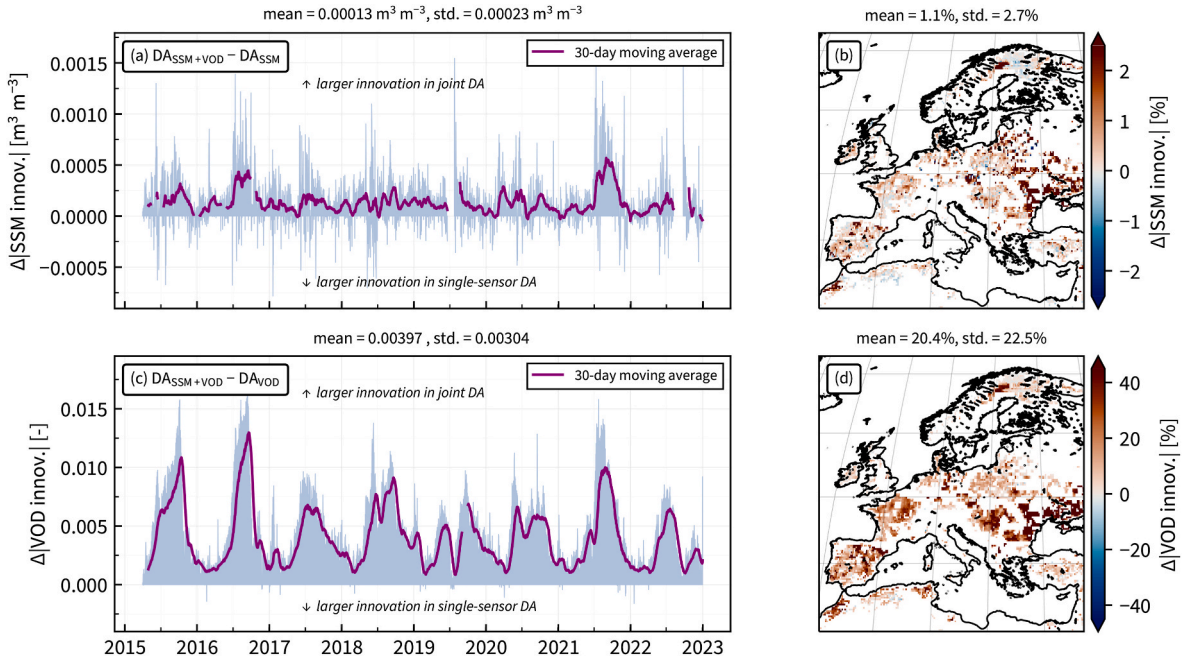

Fig. 8. Overview of differences in absolute values of innovations between $DA_{SSM+VOD}$ and (a)–(b) DA_{SSM} or (c)–(d) DA_{VOD}. Left column: time series of the difference of spatial averages, right column: maps of the difference of temporal averages.

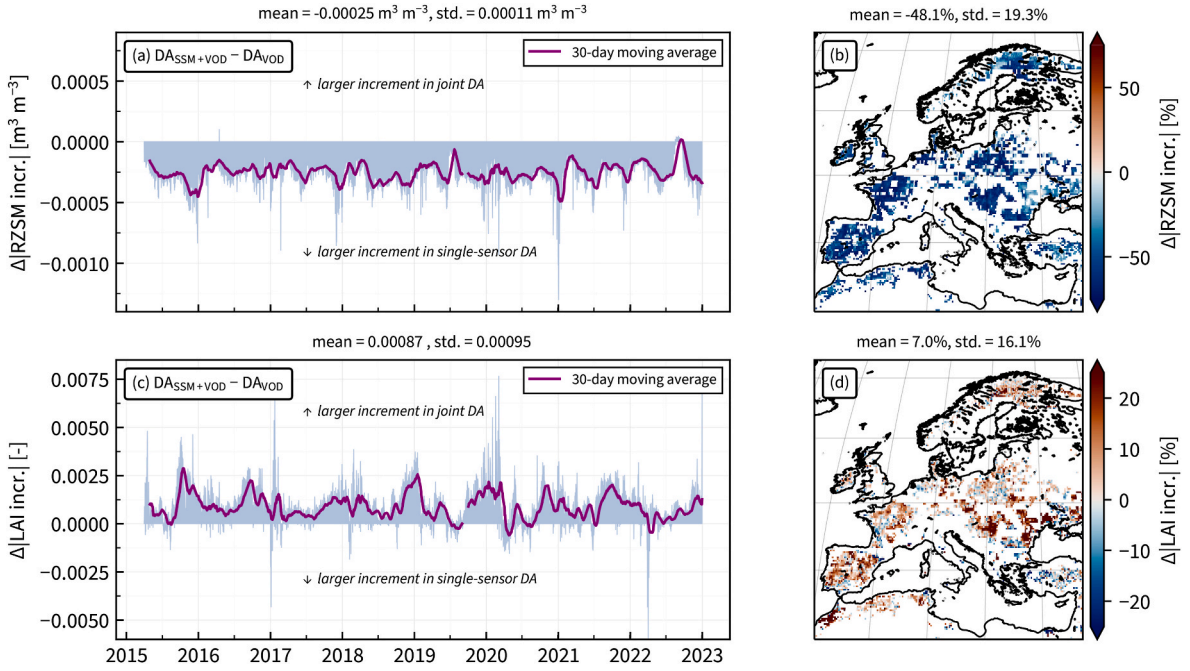

Fig. 9. Overview of differences in absolute values of increments between $DA_{SSM+VOD}$ and DA_{VOD} for (a)–(b) RZSM and (c)–(d) LAI. Left column: time series of the difference of spatial averages, right column: maps of the difference of temporal averages.

4. Discussion

4.1. Synthesis of the results

The evaluation of DA_{SSM}, DA_{LAI}, and $DA_{SSM+LAI}$ shows that the two single-sensor DA systems complement each other as expected. The reason for the complementarity of the single-sensor DA systems lies in the way they operate: DA_{SSM} only updates soil moisture while DA_{LAI} only updates LAI. Ignoring the effects of model propagation, they thus operate more or less independently of each other in the joint $DA_{SSM+LAI}$.

This is not the case when comparing DA_{SSM}, DA_{VOD} and $DA_{SSM+VOD}$, as VOD DA updates both the LAI and RZSM and more complicated interactions therefore occur. The innovations of both SSM and VOD are greater in $DA_{SSM+VOD}$ than in either DA_{SSM} or DA_{VOD}, and soil moisture increments introduced by SMAP SSM and AMSR2 VOD are negatively correlated throughout the domain. This indicates that the two DA systems counteract each other. However, larger innovations in $DA_{SSM+VOD}$ do not directly translate into larger increments either: larger RZSM and LAI increments are introduced by SMAP SSM and AMSR2 VOD, respectively, but AMSR2 VOD introduces smaller RZSM increments for

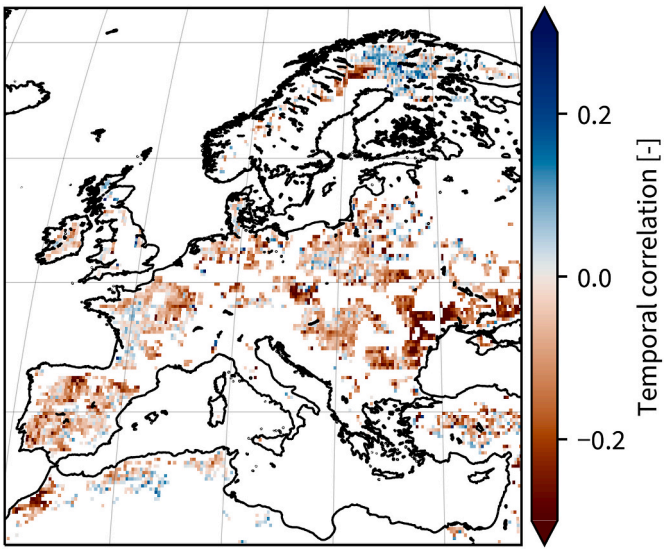

Fig. 10. Temporal correlation between RZSM increments from alternating innovations of SSM and VOD in the $DA_{SSM+VOD}$ experiment.

$DA_{SSM+VOD}$ than for DA_{VOD}. This is because SSM DA substantially decreases the spread of SSM and RZSM (Table 4). As a consequence, the EnKF gives more weight to the SSM and RZSM model estimates than to the AMSR2 observations in the joint DA. The larger innovations are preferentially attributed to larger LAI increments, because the LAI spread is not reduced by the SSM DA. Similarly, the SMAP RZSM increments are slightly reduced in $DA_{SSM+VOD}$ compared to DA_{SSM}, since VOD assimilation has reduced the spread in the modeled RZSM.

4.1.1. Optimality of the DA systems

Despite the strong reduction in ensemble spread for the joint DA experiments, we find through internal diagnostics of the DA system (Desroziers et al., 2005) that the perturbation settings of Table 2 are well suited for both single-sensor and joint DA experiments. Fig. 11 shows the distributions of the pixel-wise standard deviations of the normalized innovations, which should be centered on unity, if the model and observation errors are assumed correctly (Reichle et al., 2002; De Lannoy and Reichle, 2016a, 2016b; Heyvaert et al., 2023). There is no substantial difference in this diagnostic between single-sensor and joint DA experiments. Therefore, the random ensemble model and observation error characterization in the DA system is well chosen, especially for the SMAP SSM and AMSR2 VOD retrievals.

By design, any bias between assimilated observations and model predictions was removed for the single-sensor DA experiments, but this does not guarantee a bias-free situation in a joint DA system. A bias in the joint DA system could explain systematic differences in some time series of innovations and increments (e.g., the fact that VOD innovations tend to be larger for $DA_{SSM+VOD}$ than for DA_{VOD}, Fig. 8c). This concern is refuted in Fig. S16, which shows that the innovations for joint DA experiments are not biased and have distributions similar to those of single-sensor DA experiments.

4.1.2. Complementarity of the DA systems

In the joint $DA_{SSM+VOD}$, the sizes of the SSM increments from AMSR2 VOD are approximately an order of magnitude smaller than the SSM increments from SMAP SSM. This may be explained by the chosen observation errors and by the fact that the VOD is related to RZSM rather than SSM via the observation operator (section 2.3). Despite this, the DA impact on SSM estimates shown in Fig. 4 is only slightly less for DA_{VOD} than for DA_{SSM}. Although this appears to be a contradiction at first sight, it may be explained by considering the lagged autocorrelations of the total error (model and observations) in the DA system. We find that this quantity is substantially larger for VOD than for SSM retrievals (Fig. S17), likely due to the smoothing performed by the LPDR algorithm. As such, a potential error can return in multiple small increments and over time result in (i) the high RMSD between OL and DA observed in the impact analysis and (ii) significant degradations in SSM by DA_{VOD} (Table 3 and Fig. 5).

$DA_{SSM+VOD}$ and DA_{SSM} show an interesting difference in the magnitude of SSM innovations after the two outages of the SMAP instrument in the summers of 2019 and 2022 (Fig. 8a). These downtime periods are followed by large positive peaks in the innovations, which means that the joint SSM estimates drift away from the SSM observed by the satellite during the SMAP outage. The large innovation spikes also translate to larger increments for SSM in the joint (Fig. S15a), indicating that SSM DA has to make more corrections if VOD is assimilated than when this is not the case. However, the opposite is true for RZSM increments, which show negative spikes (Fig. S15c), indicating that the assimilation of SSM needs to make fewer corrections when jointly assimilated with VOD. This could signify a poor vertical coupling between the surface and the root zone of the Noah-MP LSM, which does not allow a transfer of the improvements made by VOD DA from the root-zone to the surface. Another explanation may be that assimilating VOD decreases the model spread of RZSM three times more than that of SSM (Table 4), causing smaller increments in the joint DA.

It should be noted that the time series of innovations and increments, as shown in Figs. 7–9, show spatially averaged values over very different domains depending on the season, as observations over snow-covered or frozen soil are not assimilated. Figs. S18a and S19a in the supplement show time series of the number of pixels over which assimilation of SMAP SSM and AMSR2 VOD occurred, respectively.

4.2. Limitations of the study

4.2.1. Rescaling of the satellite retrievals

Since the EnKF expects no bias between the model estimates and the observations, CDF matching was applied to match the retrieval

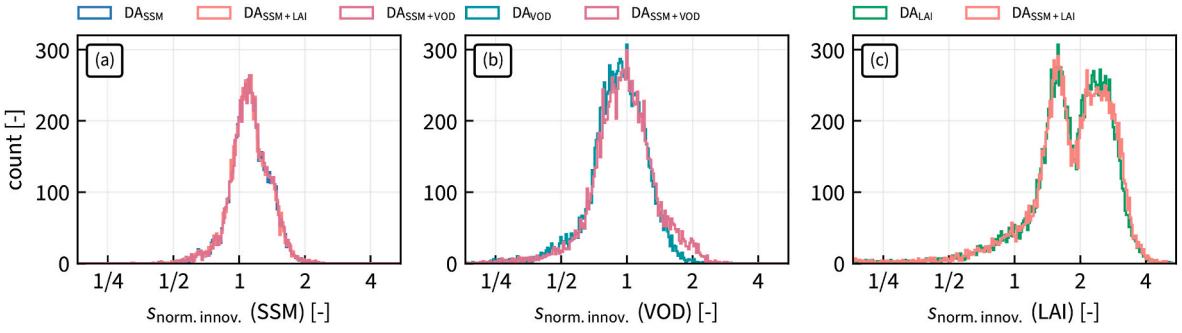

Fig. 11. Distributions of the temporal standard deviations of the normalized innovations $s_{norm.\ innov.}$ of (a) SSM, (b) VOD, and (c) LAI. For each panel, the relevant DA experiments are shown. The horizontal axis is logarithmic.

climatology to that of the LSM (Reichle and Koster, 2004; Barbu et al., 2014; Heyvaert et al., 2023). While this is a common approach in literature, it is known to potentially result in an important loss of information when the bias in the innovations is the result of anthropogenic processes such as irrigation, which are not well represented in LSMs (Kumar et al., 2015) and therefore not modeled in our study. For such a situation, some studies present alternative approaches with a reduced bias correction in which for example only the mean climatological cycle is bias corrected (e.g., Kwon et al., 2022). This might further improve the results of DA_{SSM} since our domain contains many agricultural areas.

Regarding LAI, it is common practice not to rescale the retrievals before assimilation (e.g., as is done by Albergel et al., 2017; Erlingis et al., 2021; Rahman et al., 2022b). In contrast to this, we only update differences in anomalies by rescaling the LAI as in Scherrer et al. (2023) to respect the assumptions of the EnKF. As a consequence, the impact of LAI DA on the soil moisture estimates (Fig. 4a and Fig. 5a) is much smaller in our study compared to studies assimilating biased LAI in Noah-MP (e.g., Kumar et al., 2019; Mocko et al., 2021). Overall, LAI in Noah-MP appears to be more sensitive to small changes in soil moisture than the other way around (impact of DA_{SSM} on LAI is large compared to impact of DA_{LAI} on SSM and RZSM). This apparent insensitivity of the Noah-MP hydrology to vegetation dynamics has also been reported by other studies (e.g., Hosseini et al., 2022).

4.2.2. Reference data

We aim to facilitate a fair comparison between all DA experiments by only evaluating over grid cells where the three satellite products are available for assimilation. Unfortunately, this limits the availability of independent reference data, especially in the case of the Warm Winter 2020 database. We compensated for this by performing an additional evaluation using X-BASE, but this probably overestimates the improvements introduced by DA_{LAI}, because X-BASE has been trained using data from optical remote sensing. Despite this, we found in Table 3 that DA_{LAI} performs best in terms of GPP and ET not only with respect to X-BASE, but also to the Warm Winter 2020 database. To further examine the correspondence between both datasets, we compare R_{anom} for GPP over grid cells for which a ground station is available (Fig. S20). While X-BASE appears to systematically overestimate the performance, especially at Warm Winter 2020 ground stations with poor R_{anom} values, there is a strong correspondence between both. This is consistent with the findings of Wild et al. (2022) and an additional motivation to use the X-BASE product to supplement the limited ground station data in the evaluation of DA experiments.

4.2.3. Regional effects

An area that often stands out is eastern Europe, particularly the region of Ukraine north of the Black Sea. This area is known to have a relatively poor performance of the VOD observation operator (Fig. 2b). Despite this, there is a large impact of DA_{VOD} on model estimates in this region (Fig. 4). Therefore, it could be useful to use a spatially varying observation error in the DA system in future work, so that larger uncertainties can be applied to VOD observations in such areas. This would be particularly relevant since the VOD assimilation degrades estimates of, for example, GPP in this area (Fig. 6c). With this in mind, it is not surprising that the SMAP SSM innovations are larger in this area for $DA_{SSM+VOD}$ compared to DA_{SSM} (Fig. 8b): assimilating VOD drives SSM estimates further away from SMAP observations. A good characteristic of the joint $DA_{SSM+VOD}$ setup is the fact that these larger innovations translate into larger increments for SSM and RZSM in this area (Fig. S15) and that these increments aim to counteract the effects of the VOD assimilation (Fig. 10). The result is that the largest degradations by DA_{VOD} in geophysical estimates are less pronounced in the joint $DA_{SSM+VOD}$ (compare panels c and e of Fig. 6 for GPP).

4.2.4. Limitations of VOD DA

Contrary to Kumar et al. (2020) and Mucia et al. (2022), we did not find significant improvements of RZSM or GPP estimates compared to reference data with VOD DA. DA_{VOD} even degrades the estimates of RZSM compared to in situ sites (Table 3), even though including RZSM in the VOD observation operator improves the performance of the observation operator. Therefore, the smaller SMAP RZSM increments in the joint DA (Figs. S15c–d) are counterproductive, and $DA_{SSM+VOD}$ shows less improvement in RZSM than DA_{SSM}.

One major difference of our setup to Kumar et al. (2020) and Mucia et al. (2022) is that both of these studies use a biased VOD DA setup where a large part of the DA effects are due to bias correction. Since both rescale the VOD data to a reference LAI, it is difficult to separate the relative contributions of VOD anomaly DA and reference LAI climatology DA in their setup, especially considering that they found very similar performance for both VOD DA and LAI DA. Other reasons for the poor performance in our setup could be related to data quality, the observation operator, or the setup of the joint DA system, which will be discussed in more detail below.

The LPDR data product used here applies a strong smoothing to the observations. While this can be useful for studies using the data directly, it can be counterproductive for data assimilation. Smoothed observations result in substantially higher total error autocorrelation from VOD DA compared to SSM DA (Fig. S19). This may explain why we do not see a positive effect of the more frequent updates in VOD DA compared to LAI DA (daily versus 10-daily) that has been reported by Mucia et al. (2022). To overcome this problem, data thinning or an alternative VOD product that employs less smoothing could be used in future studies.

A potential shortcoming of the observation operator is the calibration using LAI and RZSM modeled by Noah-MP. This makes it sensitive to how well the coupling of LAI and RZSM is represented in this model. If Noah-MP overestimates the coupling, so will the observation operator, and it will therefore overcorrect RZSM based on VOD innovations. Such an overcoupling can degrade DA performance (Crow et al., 2020). Additionally, a coupling that is too strong can present problems during the calibration of the observation operator, since the multicolinearity of LAI and RZSM can make it difficult to assign the correct effect size for LAI and RZSM.

Finally, it might be possible to improve $DA_{SSM+VOD}$ performance by further improving the perturbations to the model state variables. Contrary to observation perturbations, which have been tuned based on innovation diagnostics, we used literature values from Kumar et al. (2019) and our own estimates for the error cross-correlation of LAI and SM. In the case of $DA_{SSM+LAI}$ (a similar setup as in Kumar et al. (2019; 2020)) this did not cause problems, because RZSM has only a weak effect on LAI in the model. For LAI DA, the relative sizes of RZSM and LAI perturbations also only have a weak effect, since updates of RZSM due to LAI DA is limited by the low ensemble cross-correlation between model predicted observable (LAI) and RZSM. Instead, in the case of DA_{VOD} and $DA_{SSM+VOD}$, the model predicted observable (VOD) has a higher correlation with RZSM, since RZSM factors into the observation operator. As a consequence, the relative sizes of the perturbations of RZSM and LAI matter, because they govern how strongly the VOD innovations are translated into RZSM and LAI increments, respectively. Underestimating the LAI perturbation size (i.e., being overconfident in the model LAI estimates), will consequently lead to higher RZSM updates, with similar effects as an overestimation of LAI-RZSM coupling in the model.

As discussed above, for $DA_{SSM+VOD}$ the VOD DA leads to higher increments of LAI and lower increments of RZSM compared to DA_{VOD}, due to the lower spread in RZSM caused by the additional SSM DA. Therefore, increasing the LAI perturbation size of the model could help decouple the vegetation and soil moisture updates in $DA_{SSM+VOD}$, and reduce the negative impact of VOD DA on RZSM.

5. Conclusions

This work evaluates single-sensor and joint data assimilation of three satellite products into the Noah-MP land surface model: (1) a microwave

SMAP L2 surface soil moisture product, (2) the optical CGLS leaf area index product, and (3) an AMSR2 X-band vegetation optical depth product. For each DA experiment, the impact of assimilation on the model estimates was examined in terms of RMSD with respect to the OL or model-only experiment, and the improvements in the DA estimates were quantified using independent evaluation data. The single-sensor experiments showed that soil moisture estimates are most improved by assimilating SMAP SSM. The uncertainty in the soil moisture estimates (ensemble spread) was also decreased the most by assimilating this product. Similarly, vegetation estimates such as LAI or GPP are mainly affected by assimilating a vegetation product, i.e., either CGLS LAI or AMSR2 VOD. The largest improvements in estimates of GPP, NEE, and ET were observed when CGLS LAI was assimilated.

We find that the joint assimilation of SMAP SSM and CGLS LAI 'inherits' the properties of the two single-sensor DA experiments in the sense that it shows relatively large improvements in both the soil moisture and vegetation state variables. Overall, the innovations are slightly smaller in the joint experiment, meaning that the observations and model estimates are closer together at times of updates.

For the joint assimilation of SMAP SSM and AMSR2 VOD, the story is more complex. While this particular DA experiment shows the largest decrease in ensemble spread for all considered state variables, we also find that soil moisture increments introduced by both sensors actively attempt to counteract each other. However, this counteracting is present predominantly over areas where the VOD assimilation does not perform well, and demonstrates that DA of one product can counteract the flaws of another product in a multi-sensor DA setup.

Our results emphasize the value of integrating data from different satellite sources to improve self-consistent estimates of land surface state variables and fluxes. This will ultimately improve applications in, for example, land and water management or numerical weather prediction. Future work can focus on improving the observation operator such that the performance of the VOD DA increases, and further optimizing the perturbation settings of the model, particularly for the LAI prognostic variable and its cross-correlation with soil moisture.

CRediT authorship contribution statement

Zdenko Heyvaert: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Samuel Scherrer:** Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Wouter Dorigo:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Michel Bechtold:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Gabriëlle De Lannoy:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was funded via the CONSOLIDATION project by the Research Foundation Flanders (FWO, grant no. G0A7320N) and the Austrian Science Fund (FWF, grant no. I 4489-N). The computer resources and services used in this work were provided by the High-Performance Computing system of the Vlaams Supercomputer Center (VSC), funded by FWO and the Flemish Government (Storage4Climate collaborative grant). The authors thank the editor and the four reviewers for their constructive feedback.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.srs.2024.100129.

References

Ahmad, J.A., Forman, B.A., Kumar, S.V., 2022. Soil moisture estimation in South Asia via assimilation of SMAP retrievals. Hydrol. Earth Syst. Sci. 26, 2221–2243.

Albergel, C., Munier, S., Leroux, D.J., Dewaele, H., Fairbairn, D., Barbu, A.L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., et al., 2017. 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. (GMD) 10, 3889–3912.

Baek, S.J., Hunt, B.R., Kalnay, E., Ott, E., Szunyogh, I., 2006. Local ensemble Kalman filtering in the presence of model bias. Tellus Dyn. Meteorol. Oceanogr. 58, 293–306.

Balsamo, G., Agusti-Parareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Blyth, E., Bousserez, N., Boussetta, S., Brown, A., et al., 2018. Satellite and in situ observations for advancing global earth surface modelling: a review. Rem. Sens. 10, 2038.

Barbu, A., Calvet, J.C., Mahfouf, J.F., Lafont, S., 2014. 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, 173–192.

Bonan, G., 2019. Terrestrial Biosphere Models. Climate Change and Terrestrial Ecosystem Modeling. Cambridge University Press, Cambridge, UK, pp. 1–24.

Bousquet, E., Mialon, A., Rodriguez-Fernandez, N., Prigent, C., Wagner, F.H., Kerr, Y.H., 2021. Influence of surface water variations on VOD and biomass estimates from passive microwave sensors. Remote Sens. Environ. 257, 112345 https://doi.org/10.1016/j.rse.2021.112345. URL: https://www.sciencedirect.com/science/article/pii/S0034425721000638.

Brodzik, M.J., Billingsley, B., Haran, T., Raup, B., Savoie, M.H., 2012. Ease-grid 2.0: incremental but significant improvements for earth-gridded data sets. ISPRS Int. J. Geo-Inf. 1, 32–45.

Chan, S.K., Bindlish, R., O'Neill, P.E., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., et al., 2016. Assessment of the SMAP passive soil moisture product. IEEE Trans. Geosci. Rem. Sens. 54, 4994–5007.

Crow, W.T., Gomez, C.A., Sabater, J.M., Holmes, T., Hain, C.R., Lei, F., Dong, J., Alfieri, J.G., Anderson, M.C., 2020. Soil moisture–evapotranspiration overcoupling and L-band brightness temperature assimilation: sources and forecast implications. J. Hydrometeorol. 21, 2359–2374.

Crow, W.T., Kim, H., Kumar, S., 2024. Systematic modeling errors undermine the application of land data assimilation systems for hydrological and weather forecasting. J. Hydrometeorol. 25, 3–26.

De Lannoy, G.J., Bechtold, M., Albergel, C., Brocca, L., Calvet, J.C., Carrassi, A., Crow, W.T., De Rosnay, P., Durand, M., Forman, B., et al., 2022. Perspective on satellite-based land data assimilation to estimate water cycle components in an era of advanced data availability and model sophistication. Frontiers in Water 4, 981745.

De Lannoy, G.J., Reichle, R.H., 2016a. Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. Hydrol. Earth Syst. Sci. 20, 4895–4911.

De Lannoy, G.J., Reichle, R.H., 2016b. Global assimilation of multiangle and multipolarization SMOS brightness temperature observations into the GEOS-5 catchment land surface model for soil moisture estimation. J. Hydrometeorol. 17, 669–691.

Desroziers, G., Berre, L., Chapnik, B., Poli, P., 2005. Diagnosis of observation, background and analysis-error statistics in observation space. Q. J. R. Meteorol. Soc. 131, 3385–3396. https://doi.org/10.1256/qj.05.108. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.108.

Dirmeyer, P.A., Guo, Z., Gao, X., 2004. Comparison, validation, and transferability of eight multiyear global soil wetness products. J. Hydrometeorol. 5, 1011–1033.

Dong, J., Lei, F., Crow, W.T., 2022. Land transpiration-evaporation partitioning errors responsible for modeled summertime warm bias in the central United States. Nat. Commun. 13, 336.

Dorigo, W., Dietrich, S., Aires, F., Brocca, L., Carter, S., Cretaux, J.F., Dunkerley, D., Enomoto, H., Forsberg, R., Güntner, A., et al., 2021a. Closing the water cycle from observations across scales: where do we stand? Bull. Am. Meteorol. Soc. 102, E1897–E1935.

Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., Preimesberger, W., Xaver, A., Annor, F., Ardö, J., et al., 2021b. The international soil moisture network: serving earth system science for over a decade. Hydrol. Earth Syst. Sci. 25, 5749–5804.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., et al., 2017. ESA CCI soil moisture for improved earth system understanding: state-of-the art and future directions. Remote Sens. Environ. 203, 185–215.

Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., et al., 2011. The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. Hydrol. Earth Syst. Sci. 15, 1675–1698.

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiova, A., Sanchis-Dufau, A., Zamojski, D., Cordes, C., Wagner, W., Drusch, M., 2013. Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. Vadose Zone J. 12.

Draper, C., Reichle, R., De Lannoy, G., Liu, Q., 2012. Assimilation of passive and active microwave soil moisture retrievals. Geophys. Res. Lett. 39.

Du, J., Jones, L., Kimball, J., 2017a. Daily Global Land Surface Parameters Derived from AMSR-E and AMSR2". https://doi.org/10.5067/JIKQZ6WO5C5M version 2.

Du, J., Kimball, J.S., Jones, L.A., Kim, Y., Glassy, J., Watts, J.D., 2017b. A global satellite environmental data record derived from AMSR-E and AMSR2 microwave earth observations. Earth Syst. Sci. Data 9, 791–808. https://doi.org/10.5194/essd-9-791-2017. URL: https://essd.copernicus.org/articles/9/791/2017/.

Durand, M., Barros, A., Dozier, J., Adler, R., Cooley, S., Entekhabi, D., Forman, B.A., Konings, A.G., Kustas, W.P., Lundquist, J.D., et al., 2021. Achieving breakthroughs in global hydrologic science by unlocking the power of multisensor, multidisciplinary earth observations. AGU Advances 2, e2021AV000455.

Entekhabi, D., Yueh, S., De Lannoy, G., 2014. SMAP Handbook.

Erlingis, J.M., Rodell, M., Peters-Lidard, C.D., Li, B., Kumar, S.V., Famiglietti, J.S., Granger, S.L., Hurley, J.V., Liu, P.W., Mocko, D.M., 2021. A high-resolution land data assimilation system optimized for the western United States. JAWRA Journal of the American Water Resources Association 57, 692–710. https://doi.org/10.1111/1752-1688.12910. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1752-1688.12910.

Evensen, G., 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dynam. 53, 343–367.

Fang, H., Baret, F., Plummer, S., Schaepman-Strub, G., 2019. An overview of global leaf area index (LAI): methods, products, validation, and applications. Rev. Geophys. 57, 739–799.

Frappart, F., Wigneron, J.P., Li, X., Liu, X., Al-Yaari, A., Fan, L., Wang, M., Moisy, C., Le Masson, E., Aoulad Lafkih, Z., Vallé, C., Ygorra, B., Baghdadi, N., 2020. Global monitoring of the vegetation dynamics from the vegetation optical depth (VOD): a review. Rem. Sens. 12 https://doi.org/10.3390/rs12182915. URL: https://www.mdpi.com/2072-4292/12/18/2915.

Girotto, M., Reichle, R.H., Rodell, M., Liu, Q., Mahanama, S., De Lannoy, G.J., 2019. Multi-sensor assimilation of smos brightness temperature and grace terrestrial water storage observations for soil moisture and shallow groundwater estimation. Remote Sens. Environ. 227, 12–27.

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., Dorigo, W., 2019. Evolution of the ESA CCI soil moisture climate data records and their underlying merging methodology. Earth Syst. Sci. Data 11, 717–739.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049.

Heyvaert, Z., Scherrer, S., Bechtold, M., Gruber, A., Dorigo, W., Kumar, S., De Lannoy, G., 2023. Impact of design factors for ESA CCI satellite soil moisture data assimilation over Europe. Journal of Hydrometeorology 24 (7), 1193–1208.

Hosseini, A., Mocko, D.M., Brunsell, N.A., Kumar, S.V., Mahanama, S., Arsenault, K., Roundy, J.K., 2022. Understanding the impact of vegetation dynamics on the water cycle in the Noah-MP model. Frontiers in Water 4, 925852.

Jackson, T., Schmugge, T., 1991. Vegetation effects on the microwave emission of soils. Remote Sens. Environ. 36, 203–212. https://doi.org/10.1016/0034-4257(91)90057-D. URL: http://www.sciencedirect.com/science/article/pii/003442579190057D.

Jarlan, L., Balsamo, G., Lafont, S., Beljaars, A., Calvet, J.C., Mougin, É., 2008. Analysis of leaf area index in the ECMWF land surface model and impact on latent heat and carbon fluxes: application to West Africa. J. Geophys. Res. Atmos. 113.

Kerr, Y.H., Waldteufel, P., Wigneron, J.P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.J., Font, J., Reul, N., Gruhier, C., et al., 2010. The SMOS mission: new tool for monitoring key elements ofthe global water cycle. Proc. IEEE 98, 666–687.

Khaki, M., Hendricks Franssen, H.J., Han, S., 2020. Multi-mission satellite remote sensing data for improving land hydrological models via data assimilation. Sci. Rep. 10, 18791.

Konings, A.G., Gentine, P., 2017. Global variations in ecosystem-scale isohydricity. Global Change Biol. 23, 891–905. https://doi.org/10.1111/gcb.13389. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.13389.

Konings, A.G., Rao, K., Steele-Dunne, S.C., 2019. Macro to micro: microwave remote sensing of plant water content for physiology and ecology. New Phytol. 223, 1166–1172. https://doi.org/10.1111/nph.15808. URL: https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.15808.

Koster, R.D., Dirmeyer, P.A., Hahmann, A.N., Ijpelaar, R., Tyahla, L., Cox, P., Suarez, M.J., 2002. Comparing the degree of land–atmosphere interaction in four atmospheric general circulation models. J. Hydrometeorol. 3, 363–375.

Kumar, S., Kolassa, J., Reichle, R., Crow, W., de Lannoy, G., de Rosnay, P., MacBean, N., Girotto, M., Fox, A., Quaife, T., et al., 2022. An agenda for land data assimilation priorities: realizing the promise of terrestrial water, energy, and vegetation observations from space. J. Adv. Model. Earth Syst. 14, e2022MS003259.

Kumar, S., Peters-Lidard, C., Santanello, J., Reichle, R., Draper, C., Koster, R., Nearing, G., Jasinski, M., 2015. Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. Hydrol. Earth Syst. Sci. 19, 4463–4478.

Kumar, S., Peters-Lidard, C., Tian, Y., Houser, P., Geiger, J., Olden, S., Lighty, L., Eastman, J., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E., Sheffield, J., 2006. Land Information System: an interoperable framework for high resolution land surface modeling. Environ. Model. Software 21, 1402–1415. https://doi.org/10.1016/j.envsoft.2005.07.004. URL: https://www.sciencedirect.com/science/article/pii/S1364815205001283.

Kumar, S.V., Holmes, T., Andela, N., Dharssi, I., Vinodkumar, Hain, C., Peters-Lidard, C., Mahanama, S.P., Arsenault, K.R., Nie, W., Getirana, A., 2021. The 2019–2020 Australian drought and bushfires altered the partitioning of hydrological fluxes. Geophys. Res. Lett. 48, e2020GL091411 https://doi.org/10.1029/2020GL091411. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL091411. agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL091411. e2020GL091411 2020GL091411.

Kumar, S.V., Holmes, T.R., Bindlish, R., de Jeu, R., Peters-Lidard, C., 2020. Assimilation of vegetation optical depth retrievals from passive microwave radiometry. Hydrol. Earth Syst. Sci. 24, 3431–3450. https://doi.org/10.5194/hess-24-3431-2020. URL: https://hess.copernicus.org/articles/24/3431/2020/.

Kumar, S.V., Mocko, D.M., Wang, S., Peters-Lidard, C.D., Borak, J., 2019. Assimilation of remotely sensed leaf area index into the Noah-MP land surface model: impacts on water and carbon fluxes and states over the continental United States. J. Hydrometeorol. 20, 1359–1377.

Kumar, S.V., Peters-Lidard, C.D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K.R., Xia, Y., Ek, M., Riggs, G., Livneh, B., et al., 2014. Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation. J. Hydrometeorol. 15, 2446–2469.

Kwon, Y., Kumar, S.V., Navari, M., Mocko, D.M., Kemp, E.M., Wegiel, J.W., Geiger, J.V., Bindlish, R., 2022. Irrigation characterization improved by the direct use of SMAP soil moisture anomalies within a data assimilation system. Environ. Res. Lett. 17, 084006.

Lahoz, W.A., De Lannoy, G.J., 2014. Closing the gaps in our knowledge of the hydrological cycle over land: conceptual problems. Surv. Geophys. 35, 623–660.

Lievens, H., Reichle, R.H., Liu, Q., De Lannoy, G.J., Dunbar, R.S., Kim, S., Das, N.N., Cosh, M., Walker, J.P., Wagner, W., 2017. Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. Geophys. Res. Lett. 44, 6145–6153.

Lievens, H., Tomer, S.K., Al Bitar, A., De Lannoy, G.J., Drusch, M., Dumedah, G., Franssen, H.J.H., Kerr, Y.H., Martens, B., Pan, M., et al., 2015. SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. Remote Sens. Environ. 168, 146–162.

Liu, Y., Holtzman, N.M., Konings, A.G., 2021. Global ecosystem-scale plant hydraulic traits retrieved using model–data fusion. Hydrol. Earth Syst. Sci. 25, 2399–2417. https://doi.org/10.5194/hess-25-2399-2021. URL: https://hess.copernicus.org/articles/25/2399/2021/.

Loew, A., Schwank, M., Schlenz, F., 2009. Assimilation of an L-band microwave soil moisture proxy to compensate for uncertainties in precipitation data. IEEE Trans. Geosci. Rem. Sens. 47, 2606–2616.

Mahfouf, J.F., 2010. Assimilation of satellite-derived soil moisture from ASCAT in a limited-area NWP model. Q. J. R. Meteorol. Soc.: A journal of the atmospheric sciences, applied meteorology and physical oceanography 136, 784–798.

Mo, T., Choudhury, B., Schmugge, T., Wang, J.R., Jackson, T., 1982. A model for microwave emission from vegetation-covered fields. J. Geophys. Res.: Oceans 87, 11229–11237.

Mocko, D.M., Kumar, S.V., Peters-Lidard, C.D., Wang, S., 2021. Assimilation of vegetation conditions improves the representation of drought over agricultural areas. J. Hydrometeorol. 22, 1085–1098.

Moesinger, L., Dorigo, W., de Jeu, R., van der Schalie, R., Scanlon, T., Teubner, I., Forkel, M., 2020. The global long-term microwave vegetation optical depth climate archive (VODCA). Earth Syst. Sci. Data 12, 177–196. https://doi.org/10.5194/essd-12-177-2020. URL: https://essd.copernicus.org/articles/12/177/2020/.

Moesinger, L., Zotta, R.M., van der Schalie, R., Scanlon, T., de Jeu, R., Dorigo, W., 2022. Monitoring vegetation condition using microwave remote sensing: the standardized vegetation optical depth index SVODI. Biogeosciences 19, 5107–5123. https://doi.org/10.5194/bg-19-5107-2022, 10. 5194/bg-19-5107-2022.

Momen, M., Wood, J.D., Novick, K.A., Pangle, R., Pockman, W.T., McDowell, N.G., Konings, A.G., 2017. Interacting effects of leaf water potential and biomass on vegetation optical depth. J. Geophys. Res.: Biogeosciences 122, 3031–3046.

Mucia, A., Bonan, B., Albergel, C., Zheng, Y., Calvet, J.C., 2022. Assimilation of passive microwave vegetation optical depth in LDAS-Monde: a case study over the continental USA. Biogeosciences 19, 2557–2581. https://doi.org/10.5194/bg-19-2557-2022. URL: https://bg.copernicus.org/articles/19/2557/2022/.

Nelson, J.A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., Ibrom, A., Papale, D., Göckede, M., Duveiller, G., Knohl, A., Hörtnagl, L., Scott, R.L., Zhang, W., Hamdi, Z.M., Reichstein, M., Aranda-Barranco, S., Ardö, J., Op de Beeck, M., Billdesbach, D., Bowling, D., Bracho, R., Brümmer, C., Camps-Valls, G., Chen, S., Cleverly, J.R., Desai, A., Dong, G., El-Madany, T.S., Euskirchen, E.S., Feigenwinter, I., Galvagno, M., Gerosa, G., Gielen, B., Goded, I., Goslee, S., Gough, C.M., Heinesch, B., Ichii, K., Jackowicz-Korczynski, M.A., Klosterhalfen, A., Knox, S., Kobayashi, H., Kohonen, K.-M., Korkiakoski, M., Mammarella, I., Mana, G., Marzuoli, R., Matamala, R., Metzger, S., Montagnani, L., Nicolini, G., O'Halloran, T., Ourcival, J.-M., Peichl, M., Pendall, E., Ruiz Reverter, B., Roland, M., Sabbatini, S., Sachs, T., Schmidt, M., Schwalm, C.R., Shekhar, A., Silberstein, R., Silveira, M.L., Spano, D., Tagesson, T., Tramontana, G., Trotta, C., Turco, F., Vesala, T., Vincke, C., Vitale, D., Vivoni, E.R., Wang, Y., Woodgate, W., Yepez, E.A., Zhang, J., Zona, D., Jung, M.. X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, FLUXCOM-X, EGUsphere [preprint]. https://doi.org/10.5194/egusphere-2024-165.

Nelson, J.A., Walther, S., Jung, M., Gans, F., Kraft, B., Weber, U., Hamdi, Z., Duveiller, G., Zhang, W., 2023. FLUXCOM-X-BASE. https://doi.org/10.18160/5NZG-JMJE.

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. model description and evaluation with local-scale measurements. J. Geophys. Res. Atmos.

116 https://doi.org/10.1029/2010JD015139. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD015139.
O'Neill, P.E., Chan, S., Njoku, E.G., Jackson, T., Bindlish, R., Chaubell, J., 2021. SMAP L2 radiometer half-orbit 36 km EASE-grid soil moisture. Version 8. https://doi.org/10.5067/LPJ8F0TAK6E0.
Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., Reppucci, A., 2013. Soil moisture mapping using Sentinel-1 images: algorithm and preliminary validation. Remote Sens. Environ. 134, 234–248.
Peters-Lidard, C.D., Houser, P.R., Tian, Y., Kumar, S.V., Geiger, J., Olden, S., Lighty, L., Doty, B., Dirmeyer, P., Adams, J., et al., 2007. High-performance earth system modeling with NASA/GSFC's Land Information System. Innovat. Syst. Software Eng. 3, 157–165.
Powell, T.L., Galbraith, D.R., Christoffersen, B.O., Harper, A., Imbuzeiro, H.M., Rowland, L., Almeida, S., Brando, P.M., da Costa, A.C.L., Costa, M.H., et al., 2013. Confronting model predictions of carbon fluxes with measurements of amazon forests subjected to experimental drought. New Phytol. 200, 350–365.
Rahman, A., Maggioni, V., Zhang, X., Houser, P., Sauer, T., Mocko, D.M., 2022a. The joint assimilation of remotely sensed leaf area index and surface soil moisture into a land surface model. Rem. Sens. 14, 437.
Rahman, A., Zhang, X., Houser, P., Sauer, T., Maggioni, V., 2022b. Global assimilation of remotely sensed leaf area index: the impact of updating more state variables within a land surface model. Frontiers in Water 3. https://doi.org/10.3389/frwa.2021.789352. URL: https://www.frontiersin.org/article/10.3389/frwa.2021.789352.
Reichle, R.H., De Lannoy, G.J., Liu, Q., Koster, R.D., Kimball, J.S., Crow, W.T., Ardizzone, J.V., Chakraborty, P., Collins, D.W., Conaty, A.L., et al., 2017. Global assessment of the SMAP Level-4 surface and root-zone soil moisture product using assimilation diagnostics. J. Hydrometeorol. 18, 3217–3237.
Reichle, R.H., Koster, R.D., 2004. Bias reduction in short records of satellite soil moisture. Geophys. Res. Lett. 31.
Reichle, R.H., Koster, R.D., 2005. Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. Geophys. Res. Lett. 32.
Reichle, R.H., Liu, Q., Koster, R.D., Crow, W.T., De Lannoy, G.J., Kimball, J.S., Ardizzone, J.V., Bosch, D., Colliander, A., Cosh, M., et al., 2019. Version 4 of the SMAP level-4 soil moisture algorithm and data product. J. Adv. Model. Earth Syst. 11, 3106–3130.
Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002. Hydrologic data assimilation with the ensemble Kalman filter. Mon. Weather Rev. 130, 103–114.
Rodríguez-Fernández, N.J., Mialon, A., Mermoz, S., Bouvet, A., Richaume, P., Al Bitar, A., Al-Yaari, A., Brandt, M., Kaminski, T., Le Toan, T., Kerr, Y.H., Wigneron, J. P., 2018. An evaluation of SMOS L-band vegetation optical depth (L-VOD) data sets: high sensitivity of L-VOD to above-ground biomass in Africa. Biogeosciences 15, 4627–4645. https://doi.org/10.5194/bg-15-4627-2018. URL: https://bg.copernicus.org/articles/15/4627/2018/.
Ryu, D., Crow, W.T., Zhan, X., Jackson, T.J., 2009. Correcting unintended perturbation biases in hydrologic data assimilation. J. Hydrometeorol. 10, 734–750.
Scherrer, S., De Lannoy, G., Heyvaert, Z., Bechtold, M., Albergel, C., El-Madany, T.S., Dorigo, W., 2023. Bias-blind and bias-aware assimilation of leaf area index into the Noah-MP land surface model over Europe. Hydrol. Earth Syst. Sci. 27, 4087–4114. https://doi.org/10.5194/hess-27-4087-2023. URL: https://hess.copernicus.org/articles/27/4087/2023/.
Seneviratne, C.D., Oki, T.J., Douville, H.M., Colin, J.M., Ducharne, A., Cheruy, F.M., et al., 2016. LS3MIP (v1. 0) contribution to CMIP6: the land surface, snow and soil moisture model intercomparison project-aims, setup and expected outcome. Geosci. Model Dev. (GMD) 9, 2809.
Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J., 2010. Investigating soil moisture–climate interactions in a changing climate: a review. Earth Sci. Rev. 99, 125–161.
Srinivasan, G., Robock, A., Entin, J.K., Luo, L., Vinnikov, K.Y., Viterbo, P., 2000. Soil moisture simulations in revised AMIP models. J. Geophys. Res. Atmos. 105, 26635–26644.
Tian, Y., Peters-Lidard, C.D., Kumar, S.V., Geiger, J., Houser, P.R., Eastman, J.L., Dirmeyer, P., Doty, B., Adams, J., 2008. High-performance land surface modeling with a Linux cluster. Comput. Geosci. 34, 1492–1504.
Verger, A., Baret, F., Weiss, M., 2014. Near real-time vegetation monitoring at global scale. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 7, 3473–3481. https://doi.org/10.1109/JSTARS.2014.2328632.
Vermunt, P.C., Khabbazan, S., Steele-Dunne, S.C., Judge, J., Monsivais-Huertero, A., Guerriero, L., Liu, P.W., 2020. Response of subdaily L-band backscatter to internal and surface canopy water dynamics. IEEE Trans. Geosci. Rem. Sens. 1–16doi. https://doi.org/10.1109/TGRS.2020.3035881.
Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre, 2022. Warm winter 2020 ecosystem eddy covariance flux product for 73 stations in fluxnet-archive format—release 2022-1. https://doi.org/10.18160/2G60-ZHAK.
Wild, B., Teubner, I., Moesinger, L., Zotta, R.M., Forkel, M., van der Schalie, R., Sitch, S., Dorigo, W., 2022. VODCA2GPP–a new, global, long-term (1988–2020) gross primary production dataset from microwave remote sensing. Earth Syst. Sci. Data 14 (3).
Xu, T., Chen, F., He, X., Barlage, M., Zhang, Z., Liu, S., He, X., 2021. Improve the performance of the Noah-MP-crop model by jointly assimilating soil moisture and vegetation phenology data. J. Adv. Model. Earth Syst. 13, e2020MS002394.
Yang, Z.L., Niu, G.Y., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 2. evaluation over global river basins. J. Geophys. Res. Atmos. 116 https://doi.org/10.1029/2010JD015140. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD015140.
Zhou, J., Yang, K., Crow, W.T., Dong, J., Zhao, L., Feng, H., Zou, M., Lu, H., Tang, R., Jiang, Y., 2023. Potential of remote sensing surface temperature-and evapotranspiration-based land-atmosphere coupling metrics for land surface model calibration. Remote Sens. Environ. 291, 113557.