Towards constraining soil and vegetation dynamics in land surface models: Modeling ASCAT backscatter incidence-angle dependence with a Deep Neural Network

X. Shan, S. Steele-Dunne, M. Huber, S. Hahn, W. Wagner, B. Bonan, C. Albergel, J.-C. Calvet, O. Ku, S. Georgievska

Remote Sensing of Environment 279 (2022) 113116

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Towards constraining soil and vegetation dynamics in land surface models: Modeling ASCAT backscatter incidence-angle dependence with a Deep Neural Network

Xu Shan [a,b,*], Susan Steele-Dunne [a], Manuel Huber [b,1], Sebastian Hahn [c], Wolfgang Wagner [c], Bertrand Bonan [d], Clement Albergel [d,2], Jean-Christophe Calvet [d], Ou Ku [e], Sonja Georgievska [e]

[a] Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences, TU Delft, Delft, the Netherlands
[b] Department of Water Management, Faculty of Civil Engineering and Geosciences, TU Delft, Delft, the Netherlands
[c] Department of Geodesy and Geoinformation (GEO), Vienna University of Technology, Vienna, Austria
[d] CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[e] Netherlands eScience Center, Amsterdam, the Netherlands

ARTICLE INFO

Edited by Jing M. Chen

Keywords:
ASCAT
Scatterometry
Radar
Vegetation
Land surface model
Machine learning
Deep Neural Network
Plant water dynamics
Soil moisture

ABSTRACT

A Deep Neural Network (DNN) is used to estimate the Advanced Scatterometer (ASCAT) C-band microwave normalized backscatter (σ^o_{40}), slope (σ') and curvature (σ'') over France. The Interactions between Soil, Biosphere and Atmosphere (ISBA) land surface model (LSM) is used to produce land surface variables (LSVs) that are input to the DNN. The DNN is trained to simulate σ^o_{40}, σ' and σ'' from 2007 to 2016. The predictive skill of the DNN is evaluated during an independent validation period from 2017 to 2019. Normalized sensitivity coefficients (NSCs) are computed to study the sensitivity of ASCAT observables to changes in LSVs as a function of time and space. Model performance yields a near-zeros bias in σ^o_{40} and σ'. The domain-averaged values of ρ are 0.84 and 0.85 for σ^o_{40} and σ', compared to 0.58 for σ''. The domain-averaged unbiased RMSE is 8.6% of the dynamic range for σ^o_{40} and 13% for σ', with land cover having some impact on model performance. NSC results show that the DNN-based model could reproduce the physical response of ASCAT observables to changes in LSVs. Results indicated that σ^o_{40} is sensitive to surface soil moisture and LAI and that these sensitivities vary with time, and are highly dependent on land cover type. The σ' was shown to be sensitive to LAI, but also to root zone soil moisture due to the dependence of vegetation water content on soil moisture. The DNN could potentially serve as an observation operator in data assimilation to constrain soil and vegetation water dynamics in LSMs.

1. Introduction

An improved understanding of plant hydraulics, water transport processes and their role in carbon exchanges is essential to understand the role of vegetation in land-atmosphere exchanges, and its response to climate change (Sperry et al., 2016; Lemordant et al., 2018; Li et al., 2021). Most current Land Surface Models (LSMs) represent the canopy water dynamics, i.e. the prognosis of moisture states and fluxes within plants, through a semiempirical parameterization of stomatal aperture (Damour et al., 2010; Bonan et al., 2014; Konings et al., 2017b). Generally, a simple relationship is assumed between leaf net photosynthesis and stomatal conductance (Ball et al., 1987; Jacobs et al., 1996), without linking stomatal responses to plant hydraulic states/fluxes (Sperry and Love, 2015; Sperry et al., 2016; Fisher et al., 2017). This incomplete parameterization may limit the ability of LSMs to represent the response to drought (Powell et al., 2013). Generally, it leads to an overprediction of stomatal conductance under water stress (Anderegg et al., 2017) and the evapotranspiration sensitivity to soil moisture (Liu et al., 2020). It is therefore essential to constrain the canopy water dynamics in LSMs.

Satellite observations from both active and passive microwave remote sensing are sensitive to plant water content variations from sub-daily to interannual scales (Liu et al., 2011; Friesen et al., 2012; Steele-Dunne et al., 2012; van Emmerik et al., 2015, 2017; Konings et al.,

* Corresponding author at: Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences, TU Delft, Delft, the Netherlands.
E-mail address: x.shan-2@tudelft.nl (X. Shan).
[1] Now at European Space Agency, European Space Research and Technology Centre (ESTEC), 2201 AZ, Noordwijk, the Netherlands.
[2] Now at European Space Agency Climate Office, ECSAT, Harwell Campus, Didcot, Oxfordshire, UK.

https://doi.org/10.1016/j.rse.2022.113116
Received 23 July 2021; Received in revised form 10 May 2022; Accepted 6 June 2022
Available online 30 June 2022

2017a). This sensitivity to water is central to the value of microwave remote sensing for vegetation monitoring in a wide range of applications (Teubner et al., 2018; Konings et al., 2019; Moesinger et al., 2020) including the parameterization and validation of LSMs (Matheny et al., 2017; Konings et al., 2017b).

Since 2007, the Advanced Scatterometer (ASCAT), on board the Metop series of satellites operated by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), has been providing a long-term C-band backscatter data record. C-band backscatter has been shown to be influenced by water stress or drought (Friesen, 2008; Schroeder et al., 2016) and seasonal dynamics of vegetation (Frison et al., 1998; Jarlan et al., 2002). Similar correlation has been shown in backscatter at higher frequencies and attributed to vegetation water dynamics (van Emmerik et al., 2017; Konings et al., 2017b). Furthermore, soil moisture and vegetation optical depth products based on ASCAT backscatter have been applied in numerical weather prediction (Wagner et al., 2013), land data assimilation (Albergel et al., 2017; Fairbairn et al., 2017), natural hazard monitoring (Brocca et al., 2017) and agricultural applications (Vreugdenhil et al., 2018).

ASCAT backscatter as a function of incidence angle is described by a second-order Taylor polynomial (Wagner et al., 1999b; Melzer, 2013; Hahn et al., 2017). The incidence angle dependence of backscatter depends on the relative dominance of surface, volume and multiple scattering. Over bare soil σ° decreases sharply as incidence angle increases indicating strong surface scattering. Changes in soil moisture lead to an increase in backscatter at all incidence angles. Increased volumetric scattering in the presence of vegetation reduces sensitivity to incidence angle leading to a change in the slope and curvature. The degree to which volumetric scattering and multiple scattering between the vegetation and ground occur depend on vegetation structure, total water content and the vertical distribution within the vegetation. In a study of ASCAT in the north American grasslands, Steele-Dunne et al. (2019) observed that the seasonal climatology, spatial patterns and interannual variability in the slope varied between land cover types, reflecting variations in the soil moisture availability and growing season length, supporting its interpretation as a measure of vegetation density. Results demonstrated that curvature is clearly influenced by vegetation phenology, with significant variations occurring at the start and end of the growing season. Results also suggested that (in grasslands) curvature is a measure of the relative dominance of direct scattering from vertical vegetation constituents over a ground-bounce contribution, a phenomenon observed in earlier studies on wheat (Ulaby et al., 1986; Stiles et al., 2000; Mattia et al., 2003). Pfeil et al. (2020) observed that dynamics in slope followed structural and phenological changes in broadleaf deciduous forests. A more recent study by Petchiappan et al. (2021) showed that slope and curvature dynamics reflected spatial and temporal patterns in moisture availability and demand.

The motivation for the current study is to assimilate a set of ASCAT observables, i.e. the normalized backscatter, slope and curvature, to constrain water transport processes in a LSM. Several studies have assimilated ASCAT soil moisture products alone (Albergel et al., 2017, 2018), or in combination with LAI (Dewaele et al., 2017; Fairbairn et al., 2017) to improve the performance of the the CO2-responsive version of the ISBA (interactions between soil, biosphere, and atmosphere) LSM (Noilhan and Mahfouf, 1996; Calvet et al., 1998; Calvet and Soussana, 2001). The current ISBA model uses with Jacob's A-gs relationship (Jacobs et al., 1996) ("A" stands for net assimilation of CO2, and "gs" for stomatal conductance). This CO2-version of ISBA (i.e. ISBA-A-gs) simulates the dynamic vegetation evolution driven by photosynthesis in response to variations in atmospheric and climate conditions (Calvet et al., 1998; Calvet and Soussana, 2001; Calvet et al., 2004; Gibelin et al., 2006). Therefore, it enables the simulation of carbon and water cycles, their role in interactions between land and atmosphere, as well as their coupling to stomatal conductance. Meanwhile, the physics of ISBA has been implemented in the ECMWF land surface scheme TESSEL (van den Hurk et al., 2000). Assimilation of both soil moisture and LAI into the ISBA LSM resulted in improved estimates of the carbon fluxes (Leroux et al., 2018), and an improved ability to account for the impact of heatwave on vegetation (Albergel et al., 2019). More recent studies have assimilated backscatter directly using a radiative transfer model (RTM) (Lievens et al., 2017; Shamambo et al., 2019).

Our aim is to assimilate ASCAT observables directly rather than assimilating retrieved products to use all information in the incidence-angle dependence, to obviate the need for cumulative distribution function (CDF)-matching or bias correction (Fairbairn et al., 2017; Leroux et al., 2018; Rodríguez-Fernández et al., 2019), and to avoid the potential for cross-correlated errors between retrievals and model simulations (Lannoy and Reichle, 2016). In order to assimilate normalized backscatter, slope and curvature directly into an LSM, an observation operator is needed to predict these observables based on the outputs of the land surface model. Lievens et al. (2017) used the water cloud model (WCM) as an observation operator to assimilate ASCAT normalized backscatter into GLEAM model together with the assimilation of SMOS brightness temperature. Shamambo et al. (2019) also proposed using WCM as an observation operator linking ISBA-based states to ASCAT normalized backscatter. More sophisticated modeling approaches (Ulaby et al., 1990; Ferrazzoli et al., 1997; Kim et al., 2014) implement double bounce terms that are neglected in WCM. However, they require moisture content or dielectric properties of soil and vegetation cover as well as descriptions of size, shape, orientation, and distributions of scatterers in the canopy. Unfortunately, these parameters are generally not simulated by LSM or available globally. Furthermore, despite their complexity, these models still have a highly-idealized and simplified representation of vegetation and the distribution of water within it. In this study, machine learning will be employed to generate a surrogate model to circumvent the challenges and limitations of current radiative transfer modeling approaches.

The number and diversity of applications of machine learning has increased considerably in recent years Reichstein et al. (2019). In remote sensing, neural network (NN) approaches have proved valuable to"emulate" or reproduce the outputs of (e.g.) Monte Carlo radiative transfer code in a computationally efficient and fast way (Chevallier et al., 1998; Castruccio et al., 2014; Verrelst et al., 2016; Fer et al., 2018). Similarly, they have been applied to emulate retrieval models to provide products for assimilation (Rodríguez-Fernández et al., 2019). Aires et al. (2021) recently investigated the merits of applying a NN for retrieval versus forward modeling to assimilate ASCAT σ_{40} into a LSM. In this study, we will follow the approach of Xue and Forman (2015); Forman et al. (2014); Forman and Reichle (2015) where machine learning is used to provide a measurement operator for the direct assimilation of microwave observables to constrain a land surface model. The innovation of our study is that the relationship between the LSM and the observables is indirect. In the studies of Xue and Forman (2015); Forman et al. (2014); Forman and Reichle (2015) focused on snow processes, there is a direct link between the LSM states and the passive microwave brightness temperatures. Here, the ASCAT observables are known to depend on vegetation water content, which is not directly modeled in ISBA. Therefore, our hypothesis is that dynamics in vegetation water content can be captured through the inclusion of ISBA land surface variables related to vegetation water transport processes as inputs to the measurement operator.

Here, a Deep Neural Network (DNN) is used to directly relate the land surface variables from the ISBA Land Surface Model to the ASCAT backscatter, slope and curvature. The objectives are two-fold. First, this DNN-model could be used an observation operator to directly assimilate these ASCAT observables into ISBA. Second, the sensitivity of the observables to the ISBA land surface variables can improve our understanding of, and ability to operationally exploit the ASCAT backscatter, slope and curvature. The study is conducted over France, using the ASCAT data record from 2007 to 2019. The DNN is trained using data from the first 10 years of the ASCAT data record. The performance is

evaluated in terms of the ability of the DNN to reproduce the observed ASCAT backscatter, slope and curvature in the remaining 3 years. The sensitivity of the ASCAT observables to the ISBA soil and vegetation states is then quantified using so-called Normalized Sensitivity Coefficients. Particular attention is paid to the influence of land cover on seasonal variations in performance and sensitivity.

2. Data and methods

2.1. Study area

The study domain extends from 41.625°N to 50.875 N°, and from 4.625°W to 9.375°E Fig. 1. The ASCAT data and land surface variables (LSVs) are available almost daily from 2007 to 2019 with a spatial resolution of 0.25° x 0.25°. The study domain comprises 1020 grid points (GPIs).

The albedo and LAI values required by the ISBA model are obtained from the ECOCLIMAP II database, as well as vegetation fractions of 12 plant functional types and other land cover types at each GPI (Faroux et al., 2013). Data from this global surface parameter database are provided at 1 km resolution. The land cover type assigned to individual grid points is the dominant land cover types within a 25 km × 25 km window centred on the ASCAT GPI (Fig. 1 (a)). The corresponding maximum value of land cover fraction in each GPI is shown in Fig. 1 (b). Clearly all grid points comprise a mixture of different land cover types. Four representative regions of interest (ROIs) are identified with a comparatively high fraction of the dominant land cover (Fig. 1 (c)).

Fig. S1 shows the fraction of vegetation cover types of the GPIs in four regions of interest. This mixed composition is accounted for in ISBA as each grid cell is composed of patches of the different cover types.

2.2. ASCAT backscatter-incidence angle dependence

ASCAT backscatter dependence on incidence angle can be described as a second order polynomial:

$$\sigma^o(\theta) = \sigma^o(\theta_r) + \sigma'(\theta_r)(\theta-\theta_r) + \frac{1}{2}\sigma''(\theta_r)(\theta-\theta_r)^2, \qquad [\mathrm{dB}] \tag{1}$$

where the zeroth order term $\sigma^o(\theta_r)$ is the normalized backscatter at reference angle (40°), and the 1st and 2nd order coefficient $\sigma'(\theta_r)$ and $\sigma''(\theta_r)$ are referred to as slope and curvature (Hahn et al., 2017). The measured backscatter at any incidence angle can be extrapolated to the θ_r as following given the values of $\sigma'(\theta_r)$ and $\sigma''(\theta_r)$:

$$\sigma^o(\theta_r) = \sigma^o(\theta) - \sigma'(\theta_r)(\theta-\theta_r) - \frac{1}{2}\sigma''(\theta_r)(\theta-\theta_r)^2 \tag{2}$$

In this study, backscatter at 40 degrees is obtained using this equation and referred to as "normalized backscatter" or σ^o_{40}.

The slope and curvature values are estimated from backscatter triplets measured by Metop ASCAT. ASCAT observes backscatter using two sets of three side-looking antennas each illuminating a 550 km wide swath on either side of the satellite track. On each side, 3 antennas are oriented at 45° (fore), 90° (mid) and 135° (aft) to the satellite track. The incidence angles are 34–65° and 25–55° for fore−/aft- and mid-antennas, respectively. Backscatter triplets [σ_f, σ_{mid}, σ_a] (fore-, mid-, and aft-beam) at the incidence angles [θ_f, θ_{mid}, θ_a] are used to produce two local slope estimates:

$$\sigma'\left(\frac{\theta_{mid}-\theta_{a/f}}{2}\right) = \frac{\sigma^o_{mid}(\theta_{mid}) - \sigma^o_{a/f}(\theta_{a/f})}{\theta_{mid}-\theta_{a/f}}, \qquad [\mathrm{dB}/deg] \tag{3}$$

A large number of local slope values, distributed over the entire incidence angle range (Hahn et al., 2017), are combined to estimate the slope and curvature Hahn et al. (2017). Following Melzer (2013), the slope and curvature parameters for day d are computed using all local slope values within a window with half-width λ centered at d. The value of λ is chosen to balance the bias and variance of estimates. Hahn et al. (2017) tested the performance and robustness of the kernel smoother (KS) globally, by comparing the results obtained from ASCAT on-board Metop-A and Metop-B independently. Overall the results from Metop-A and Metop-B are in a good agreement, confirming a robust interannual estimation of the incidence angle dependence of backscatter using KS.

Fig. 2 summarizes the dynamics in ASCAT normalized backscatter, slope and curvature during the study period. The interannual variability and seasonal variability are estimated by assuming that the time series is the additive combination of a seasonal cycle (period = 365 days), trend and residual which is decomposed using moving averages (Seabold and Perktold, 2010). The interannual variability (IAV) and seasonal variability (SV) indicate the standard deviation of the trend and seasonal cycle series respectively.

Spatial variability in normalized backscatter, slope and curvature

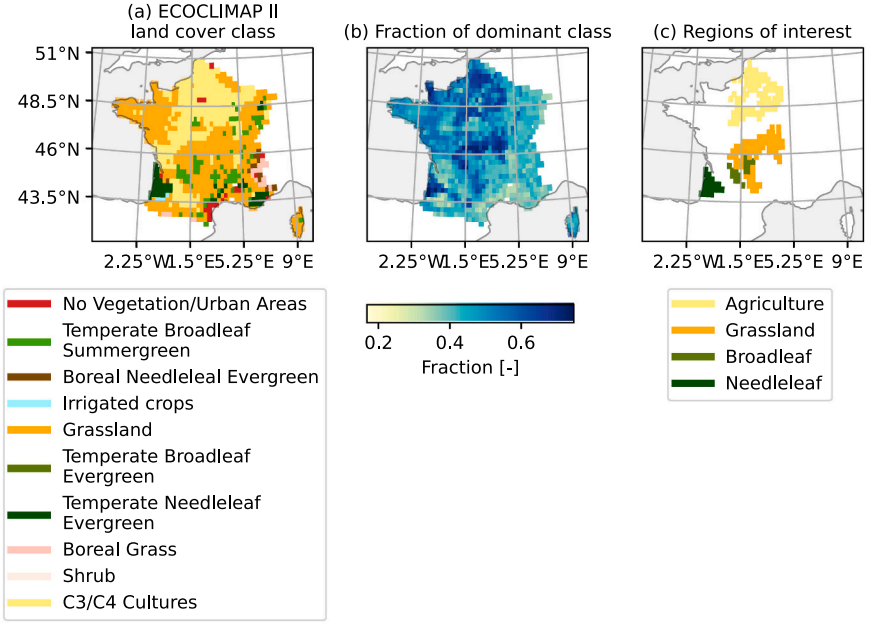

Fig. 1. The main land cover types and fraction of France based on ECOCLIMAP II. (a) the dominant vegetation types of France; (b) the vegetation fraction of the dominant vegetation types; (c) 4 regions of interest (ROIs) representing 4 vegetation types: Broadleaf (20 GPIs in lon 3–0, lat 44–46.5 with the fraction of temperate broadleaf cold-deciduous summergreen (TEBD) ≥ 0.35), Agriculture (98 GPIs in lon 0.2–5, lat 47.5–51 with the fraction of C3 cultures (C3) ≥ 0.5), Grassland (76 GPIs in lon 0.4–5, lat 43–48 with the fraction of grass (GRAS) ≥ 0.5) and Needleleaf forests (21 GPIs with the fraction of temperate needleleaf evergreen (TENE) ≥ 0.4).

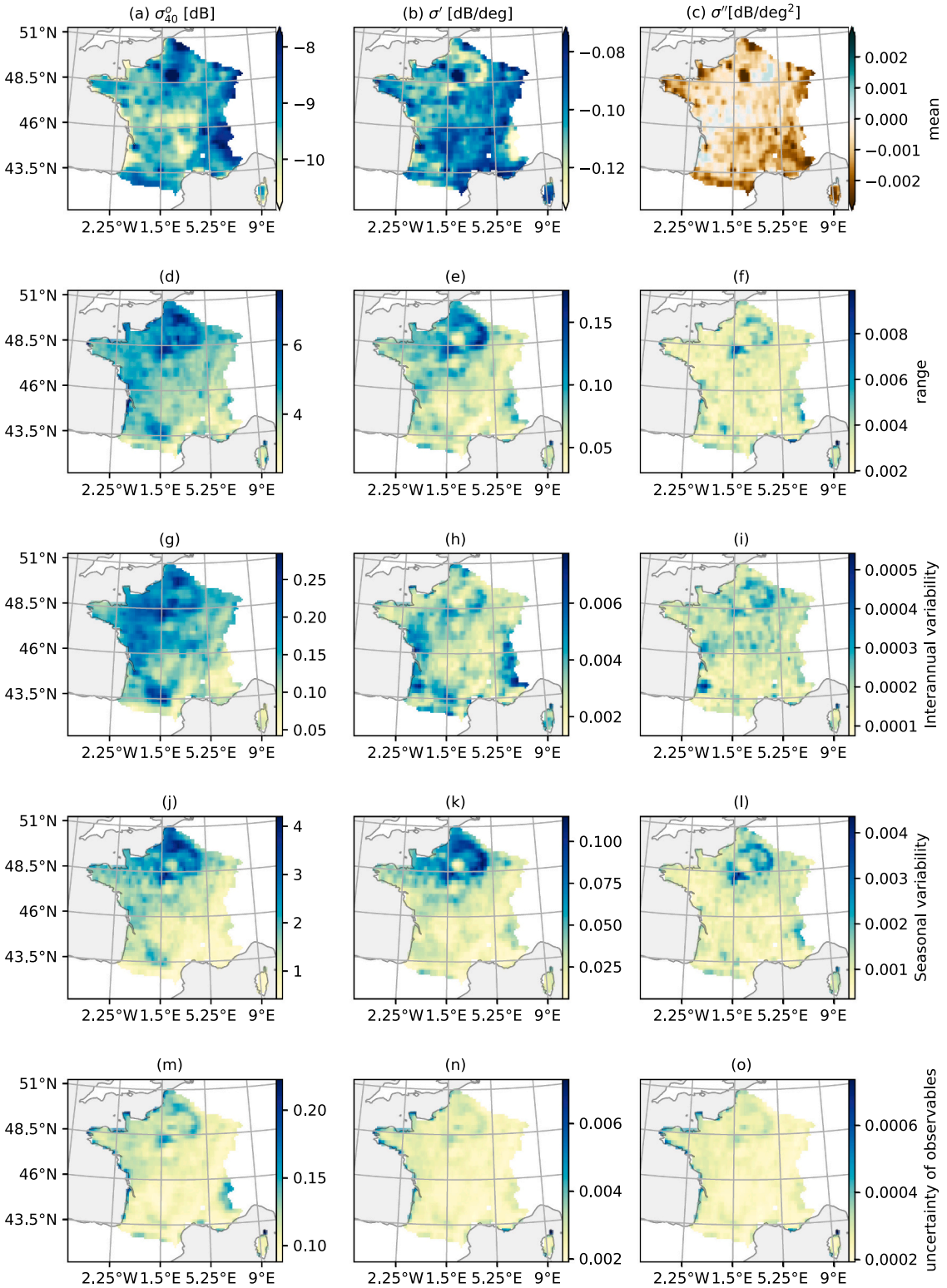

Fig. 2. The mean values, range values and uncertainty of ASCAT backscatter and DVP from 2007 to 2019. Different columns show backscatter, slope and curvature, respectively. Mean values (a - c), the range values (d - f), interannual variability (IAV, g - i), seasonal variability (j - l) and uncertainty (m-o) are shown for the study domain. The mean values >98% and smaller than 2% have been masked as the 98% and 2% values respectively.

reflect patterns in dominant land cover. Mean values range from −10.75 to −7.805 dB, from −0.13 to −0.075 dB/deg. and from −0.0027 to 0.00050 dB/deg^2 for σ^o_{40}, σ' and σ'', respectively.

Aside from urban areas, the highest normalized backscatter values are observed (−7.80 dB) in northern France where the main land cover type is Agricultural C3 crops, and lowest (−10.75 dB) in central France which is dominated by grassland land cover. The highest mean value of σ' (−0.075 dB/deg) is found in grassland, while the lowest value (−0.13 dB/deg) is found in northern France as shown in Fig. 2 (b). Agricultural crop in northern France has the smallest mean value of σ'.

The IAV of normalized backscatter, slope and curvature range from 0.042 to 0.29 dB, from 0.0013 to 0.0078 dB/deg. and from 7.5e-5 to 0.00054 dB/deg^2, respectively. The corresponding domain-averaged values of IAV are 0.15 dB, 0.0033 dB/deg. and 0.00019 dB/deg^2. The IAV of normalized backscatter is largest in northern agricultural land and is generally higher in the north and west part of France, as shown in Fig. 2 (g). Meanwhile, in Fig. 2 (h - i), it is observed that IAV of σ' and σ'' are largest in the Les Landes forest area. This is likely due to the impact of Cyclone Klaus in January 2009, after which the forest degraded from 2009 to 2012 and started regeneration during 2013–2017 (Shamambo et al., 2019; Teuling et al., 2017). As for the seasonal cycle, SV of normalized backscatter, slope and curvature are generally largest in agricultural areas, with values up to 4.2 dB, 0.11 dB/deg. and 0.0043 dB/deg^2 for σ^0_{40}, σ' and σ'', respectively. SV are specifically largest in the areas of intensive agricultural land surrounding Paris. The spatial pattern of SV is generally similar to the spatial pattern of range values.

The uncertainty of ASCAT σ^0_{40}, σ' and σ'' varies with land cover but is overall small compared to the dynamic range. Uncertainties in σ^0_{40} are highest in agricultural areas in northern France, mountainous areas such as the Alps and coastal areas.

2.3. Interactions between soil, biosphere and atmosphere (ISBA) model

Here, the ISBA land surface model (Noilhan and Planton, 1989; Noilhan and Mahfouf, 1996) is used to simulate water, energy and carbon exchanges at the land surface. The CO2-responsive version of ISBA (ISBA-A-gs, here referred as ISBA) is built within the SURFEX platform (version 8.1) (Masson et al., 2013; Albergel et al., 2017, 2018). With the "NIT" plant biomass monitoring option, ISBA simulates plant physiological states and plant growth (Calvet et al., 1998, 2004, 2007).

Exchanges of water and heat fluxes in the soil are simulated based on a multi-layer diffusion scheme (Boone et al., 2000; Decharme et al., 2013). Here, ISBA simulates soil moisture in 14 layers, with depths of 0.01 m, 0.04 m, 0.1 m, 0.2 m, 0.4 m, 0.6 m, 0.8 m, 1.0 m, 1.5 m, 2.0 m, 3.0 m, 5.0 m, 8.0 m, 12.0 m (Albergel et al., 2017; Leroux et al., 2018). The ISBA parameters are defined for 19 generic land surface patches as described in Table S1.

Parameters of the photosynthesis model and plant growth model are patch-dependent (Lafont et al., 2012). LSV values for the GPI are weighted combinations of the values obtained for each patch. The nominal values of LSVs from an open loop run of ISBA are used as input for DNN. The model was forced by the latest ERA-5 atmospheric reanalysis from ECWMF (Hersbach et al., 2020) from 1996 to 2019. The meteorological forcing data are available on a 0.25° x 0.25° grid, and include rainfall rate, 2 m air temperature, 2 m specific humidity, wind speed, wind direction, surface pressure, downward direct shortwave radiation, downward diffuse radiation, downward long wave radiation, snowfall rate and CO2 concentration. All ERA-5 atmospheric variables were interpolated using bi-linear interpolation to match the grid points of ISBA. The model was initiated by spinning up with 20 repetitions using the 1996 forcing data. The open-loop simulation was obtained using the ERA5 forcing data from 1997 to 2019.

2.4. Deep Neural Network

Here, a DNN is used to relate the ISBA LSVs to the ASCAT normalized backscatter, slope and curvature. DNNs are based on artificial neural networks (ANN), and increasingly used for machine learning applications involving large datasets with high dimensionality (Schmidhuber, 2015; Reichstein et al., 2019) including in the geosciences (e.g. Hong et al. (2004); Tao et al. (2016); Tang et al. (2018)). A DNN contains an input layer, multiple hidden layers, and an output layer. Each layer contains multiple neurons connected to neurons in adjacent layers using the corresponding weight and bias matrix (Goodfellow et al., 2016). During back-propagation, the weights and biases are updated by a stochastic optimization algorithm, Adam (Kingma and Ba, 2014), that minimizes (maximizes) the loss (accuracy) between the DNN outputs and the training targets. In order to prevent over-fitting, early-stopping was used during the training process. DNNs contain more parameters (e.g. number of layers and nodes) than other machine learning algorithms, so they have the flexibility to capture complex features from high dimensional systems (Bengio, 2009; Bianchini and Scarselli, 2014).

One DNN is trained per GPI to obtain normalized backscatter, slope and curvature simultaneously. The structure of DNN is fine-tuned by Bayesian optimization (Snoek et al., 2012). The parameters to be optimized include the number of layers, the number of neurons of each layer, batch size, the choice of activation function and learning rate. The optimal parameter combination is that which minimizes RMSE between observed and estimated σ^0_{40}, σ' and σ'' simultaneously during the testing period.

Table S1 lists the ISBA LSVs that are used as inputs to the DNN. These are selected to account for the influence of soil and vegetation on the ASCAT observables. It is widely understood that surface soil moisture has a significant influence on C-band backscatter (Wagner et al., 1999a, 1999b), therefore WG2 is included. Note that WG2 (5-10 cm) rather than WG1 (0-5 cm) is used as surface soil moisture following previous studies (Albergel et al., 2017; Draper et al., 2011; Barbu et al., 2014). The sensing depth of C-band backscatter is generally assumed to be on the order of 2 cm in non-arid environments, therefore there is unlikely to be a *direct* relationship between backscatter, slope or curvature and root zone soil moisture. However, root zone soil moisture terms (WG3 - WG10) are included as inputs, due to the relationship between root zone soil moisture, plant hydraulics and plant water content (Konings et al., 2021; Li et al., 2021; Liu et al., 2021; Bittner et al., 2012; Janott et al., 2010; Bohrer et al., 2005; Scott et al., 2003; Matheny et al., 2015, 2016). The vegetation interception reservoir water storage (WR) is included to allow for the fact that intercepted water (from dew/irrigation) can have an effect on backscatter, and hence slope and curvature. This has recently been shown for L-band backscatter (Vermunt et al., 2020; Khabbazan et al., 2022), but it is reasonable to suspect that it also has an effect on ASCAT DVP in C-band (Steele-Dunne et al., 2021). Plant transpiration (LETR) and stomatal conductance (XRS) are included as they are indicators of the transport of moisture from the root zone to the atmosphere, and therefore reflect dynamics in plant water content. Leaf Area Index (LAI) and Gross primary production (GPP) are included to account for canopy growth and biomass accumulation as both backscatter, and its incidence angle dependence are influenced by seasonal changes in both biomass itself and the associated changes in canopy architecture. Net radiation (RN) is included to account for the influence of radiation availability on phenology, and particularly structural changes in the vegetation (e.g. leaf/bud formation, leaf loss etc.) (Petchiappan et al., 2021), and changes in constituent water content associated with seasonality (Pfeil et al., 2020). The LSV are obtained from the open loop simulations of the ISBA model. The reader is referred to the studies of (Draper et al., 2011; Albergel et al., 2017; Barbu et al., 2011) for information regarding ISBA model error. Note, however, that the nominal open loop values are used without any consideration of model error.

The training targets of the DNN are the ASCAT normalized backscatter, slope and curvature. Rather than training one generic DNN, or one DNN per land cover type, a DNN is generated for each individual grid point. This is motivated by the coarse resolution of the ASCAT observations which means that the ASCAT footprint always contains a combination of several land cover types (Fig. 1(b) and Fig. S1). This heterogeneity is highly relevant because (1) the relationship between ISBA LSVs and the quantities affecting microwave backscatter varies per vegetation type (2) microwave interactions with vegetation are different per vegetation type and (3) one needs to consider how contributions from different surface types are "mixed" to yield the observables at footprint scale. Note that the proposed DNN is not simply emulating a RTM. To use most RTMs as a measurement operator, one would need to translate the outputs of ISBA (the LSVs) to the required inputs of the

RTM (e.g. roughness, soil moisture, vegetation water contents, vegetation parameters), run an RTM for each of the vegetation types, and then apply a mixing model to obtain the observables at the ASCAT scale. This is particularly challenging as the required inputs to the RTM are generally not simulated by most land surface models (e.g. VWC). The proposed DNN circumvents these challenges by mapping the ISBA LSVs directly to the ASCAT observables at the scale of interest.

The data record was split into a training-testing dataset (2007–2016) and an independent validation dataset (2017–2019). Before the training process, the inputs and training targets of the DNN are normalized as follows to account for the difference in magnitudes among the variables:

$$x_{i,norm} = \frac{x_i - mean(x_i)}{std(x_i)} \tag{4}$$

where x_i is the i-th input variable or labels and $i = 1, \ldots,$m for input LSVs or i = 1,2,3 for ASCAT data. During the cross-validation process, a"jackknife" approach was employed where training cycled through the 10-year period excluding the current year from the training-testing set (McCuen et al., 2005; Forman and Xue, 2016). The best model was chosen as that which minimized RMSE among those submodels according to their performances on the testing dataset. Finally, this best model was validated on the independent validation dataset to evaluate the performance of DNN based on unbiased Root Mean Square Error (ubRMSE), Pearson Correlation Coefficient (ρ) and bias (Entekhabi et al., 2014).

2.5. Normalized sensitivity coefficient

A sensitivity analysis was conducted to investigate the response of each DNN output to small perturbations in the DNN inputs, i.e. the sensitivity of the ASCAT observables to the ISBA LSVs. As a DNN was generated for each grid point, physical consistency in space and time between adjacent and similar grid cells provides an indication of robustness. Following Xue and Forman (2015), the normalized sensitivity coefficients (NSC) defined by (Willis and Yeh, 1987) were obtained using:

$$NSC(j,i) = \frac{\partial y_j}{\partial x_i}\frac{x_i}{|y_j|} = \frac{y_j^i - y_j^0}{\delta x_i}\frac{x_i^0}{|y_j^0|} \tag{5}$$

where x_i^0 and y_j^0 are the nominal values of input i and output j, y_j^i is the perturbed output value, δx_i is the amount of perturbation; i = 1,…,n; j = 1, …, m; n is the number of input LSVs and m is the number of outputs. The absolute value of y_j^0 is used to normalize the Jacobian term $\frac{\partial y_j}{\partial x_i}$ because the values of σ_{40}^o and σ' are negative. In that case, NSC keeps the same sign as the Jacobian term $\frac{\partial y_j}{\partial x_i}$.

The perturbation is applied to one LSV at a time in order to calculate the NSC. If the perturbation is too small, it will merely amplify model noise. If the perturbation is too large, the marginal function will not be a local estimate of the rate of change in the model output with respect to the change in the input (Xue and Forman, 2015). Hence, a perturbation of size ±5% of the range of x_i^0 was used.

The NSC is an indicator of the sensitivity of the measurement operator (the DNN-model). These are calculated to confirm that the sensitivities of the DNN-based measurement are physically plausible. NSC are normalized Jacobian values, analysis of which is commonly performed in data assimilation studies to understand the sensitivity of the observations to the states (Rüdiger et al., 2010; Chevallier and Mahfouf, 2001; Fillion and Mahfouf, 2003; Garand et al., 2001). This is key to identifying which states will be updated using new observation types (Xue et al., 2018), and to quantify any seasonal or other variation in this sensitivity. Note that the NSC uses the local linear relationship between the observables and states and therefore allows for a non-linear relationship between the two.

3. Results and discussion

3.1. Model performance

In Fig. 3, the three performance metrics are mapped for the study domain to evaluate the performance of the DNN model. Note that the unbiased RMSE and bias are normalized to account for the variation in dynamic range across grid cells due to the difference in land cover. In Fig. S5 and S6, the original values of ubRMSE and bias are shown to compare with the uncertainty of ASCAT normalized backscatter, slope and curvature. ubRMSE is always higher than the uncertainties of the observed variables. Pearson Correlation Coefficient, ρ, is generally higher for slope and normalized backscatter than it is for curvature. The domain-averaged values of ρ are 0.84 and 0.85 for σ_{40}^o and σ', compared to 0.58 for σ''. The spatial patterns in ρ appear to be related to land cover and climate. ρ values for σ_{40}^o and σ' are higher in northern and northwestern France where seasonal variability in both quantities is largest. Comparatively low ρ values for σ_{40}^o are found in the Pyrenees and Alps. In addition to the influence of elevation and complex terrain, this poor performance is likely due to the fact that snow cover was not included in the list of inputs to the DNN in this study. Generally ρ for σ'' reaches the largest value in central grassland and northern agricultural areas. In a small number of grid cells ρ is around zero for one or more of the observables, but these are often grid cells with urban areas. For normalized backscatter, ubRMSE is <11% of the range for 98% of grid cells, with most exceptions occurring in the Alps or Pyrenees. Bias is also generally low, with 98% of values <3.3% of the dynamic range. Generally, the bias is positive in agricultural areas and negative elsewhere with particularly large negative biases occur in urban areas. For slope, the lowest ubRMSE values occur in C3/C4 crop areas where ρ is highest. The ubRMSE is <19% of the range for 98% of grid cells, with larger values limited to urban areas. Bias in slope is particularly low, with 95% below 3.7% of the dynamic range. Bias is slightly positive in grasslands and negative everywhere else.

The poorest performance is observed in the curvature estimates. In addition to the lower ρ values, the ubRMSE and bias are both higher fractions of dynamic range than in the case of backscatter or slope. The median values of ubRMSE and bias are 18% and − 10% of the range, with no clear spatial pattern. The curvature is the second derivative of the relationship between backscatter and incidence angle. It is estimated from the local slope values and is inherently more variable than the first derivative (the slope). From a microwave perspective, it has been argued that the curvature contains information on the relative dominance of various scattering mechanisms, which depends on both the structure and the vertical distribution of moisture within the vegetation. Neither of these quantities is modeled by ISBA or closely related to states simulated by ISBA. For example, Table S2 shows that the maximum correlation between the LSVs and σ'' is just 0.38. Therefore, it is expected that the curvature is more difficult to predict than the backscatter and curvature.

Additionally, bias shown in Fig. 3(g - i) is generally lower than ubRMSE for σ_{40}^o and σ', suggesting that the error is primarily due to variance rather than bias. This suggests that using the DNN as a measurement operator in assimilation could reduce or eliminate the need for bias correction. Figs. S2 and S3 show that the DNN compresses the range of the ASCAT observables. The median range compression of σ_{40}^o, σ' and σ'' is 13%, 13% and 17% (median), respectively.

Fig. 4 shows boxplots of performance matrices binned for the four ROIs indicated in Fig. 1. Performance is best in the agricultural ROI where ρ is around 0.9 and the ubRMSE is around 10% of the range for both σ_{40} and σ'. The bias is also lower in the agricultural ROI than in the other cover types. This may be due to the strong seasonal cycles in both σ_{40} and σ' in agricultural areas, indicated by a larger range in σ_{40} (Fig. 4 (m) and σ' (Fig. 4(n). In contrast, ρ for both σ_{40} and σ' are lowest in the broadleaf forest ROI, in which the range of σ_{40} and σ' is much lower.

Note that the performance is generally poor and varies considerably

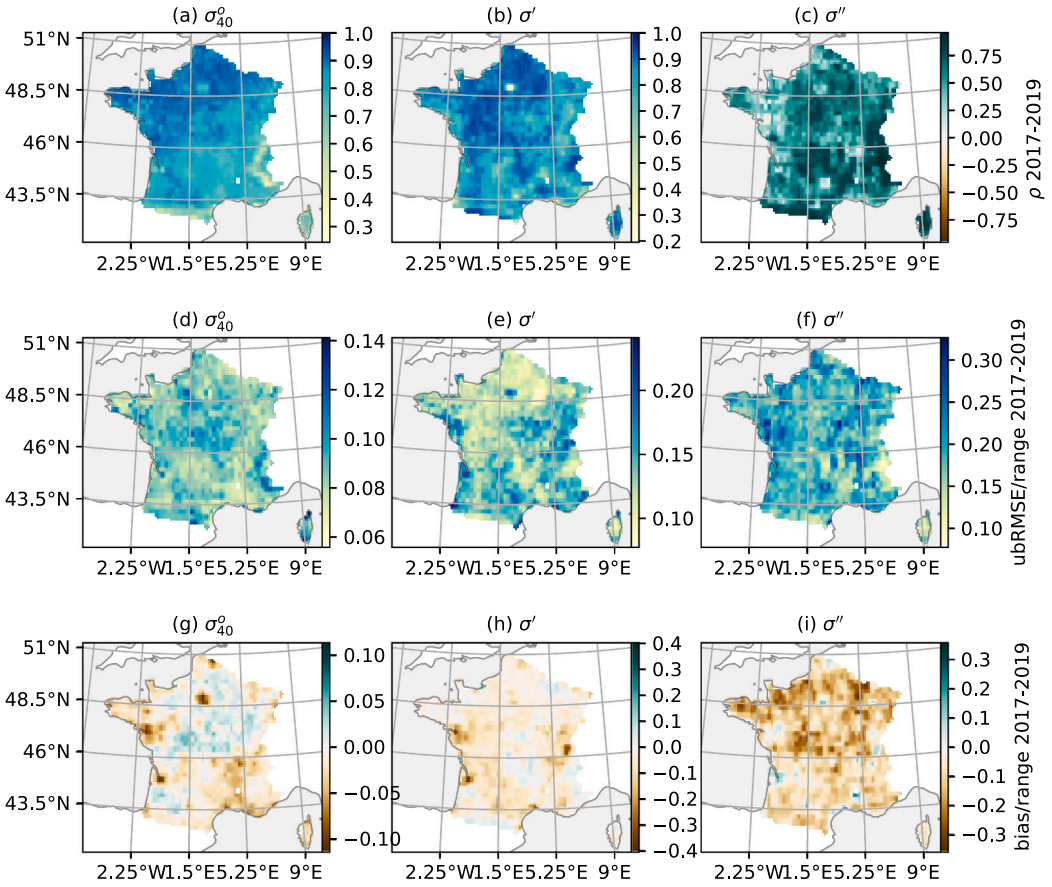

Fig. 3. The performance of the DNN during the validation period (2017–2019). The columns show normalized backscatter, slope and curvature respectively. The rows show the Pearson correlation coefficient (a -c), unbiased RMSE normalized by dynamic range (d - f) and bias normalized by dynamic range (g - i). Note that the colorbars are different on each plot.

within the needleleaf forest ROI despite its limited spatial extent. This spatial heterogeneity may be explained by the damage caused by Cyclone Klaus in January 2009. Damage due to the storm itself and the subsequent recovery of the forest means that interannual variability is high, rendering the observations difficult to predict.

Fig. 5 shows the time series of predicted and observed normalized backscatter, slope and curvature for two grid points in the Agricultural ROI. Results in the left column are from the grid point within this region that has the best performance in terms of RMSE. The RSMEs in backscatter, slope and curvature are just 5%, 8% and 13% of the range observed in these quantities. The corresponding values for the grid point presented on the right (worst performance) are almost twice as high.

From a visual comparison, their performance does not seem so different. At both grid points, the DNN captures the seasonal cycle very well. Interannual variability in backscatter and slope is captured reasonably well. This is easiest to see in the slope time series, where the predicted values match those observed in terms of both magnitude and timing. The slope is underestimated in the winters of 2015 and 2016 at the"best" grid point (Fig. 5). Note that the observed slope is higher in these two years than in the rest of the training period. Their unusually high values may be harder to capture, particularly if they are due to an extreme value of one of the inputs or an anomalous variations in a quantity not included in the list of inputs. In general, performance is poorest for curvature. However, given our limited physical understanding of what the curvature captures in terms of geophysical variables, it is encouraging that the DNN captures the seasonal cycle and interannual variability so well. Clearly there is scope for improvement through, for example, inclusion of additional input variables. However, part of the poor performance may be due to the fact that curvature is generally noisier than slope. The predicted slope and curvature are noisier than those observed. This may be attributed to the fact that their values on a given day are based on all local slope values within a 42-day window, effectively smoothing the observations. Applying an Epanechnikov kernel smoother to the predicted slope and curvature a posteriori could reduce the random errors.

3.2. *Model sensitivity: a qualitative analysis*

Fig. 6 shows the remarkable agreement between the predicted and observed ASCAT observables in all ROIs, despite their diversity in terms of SM and LAI. The strongest seasonal cycles in WG2 and LAI are observed in the Agricultural ROI. Fig. 6(b) shows that σ^o_{40} closely follows WG2 from day 200 onwards due to the low water content during crop senescence, and the sensitivity to soil moisture in the winter months in the absence of vegetation cover. This sensitivity is lost in spring as LAI rises. Fig. 6(b) also shows that the DNN models the strong drop in normalized backscatter from around February to April. This drop occurs long before LAI starts to increase (Fig. 6(f)) and starts before surface soil moisture starts to decrease (Fig. 6(b)). Such a rapid decrease in the real observed backscatter cannot be explained by the change in WG2 alone. It may be amplified by agricultural management practices e.g. tillage and planting in the spring, though the absence of ground data makes this difficult to prove conclusively. The decrease in σ^o_{40} with increasing LAI has been observed in narrow-leaved crops (e.g. wheat and barley) in previous studies (Macelloni et al., 2001; Fontanelli et al., 2013) and can be attributed to the transition from soil dominating backscatter in the winter, to vegetation dominating during the growing season. Fig. 6(j) shows that the slope generally follows the overall seasonal cycles of LAI

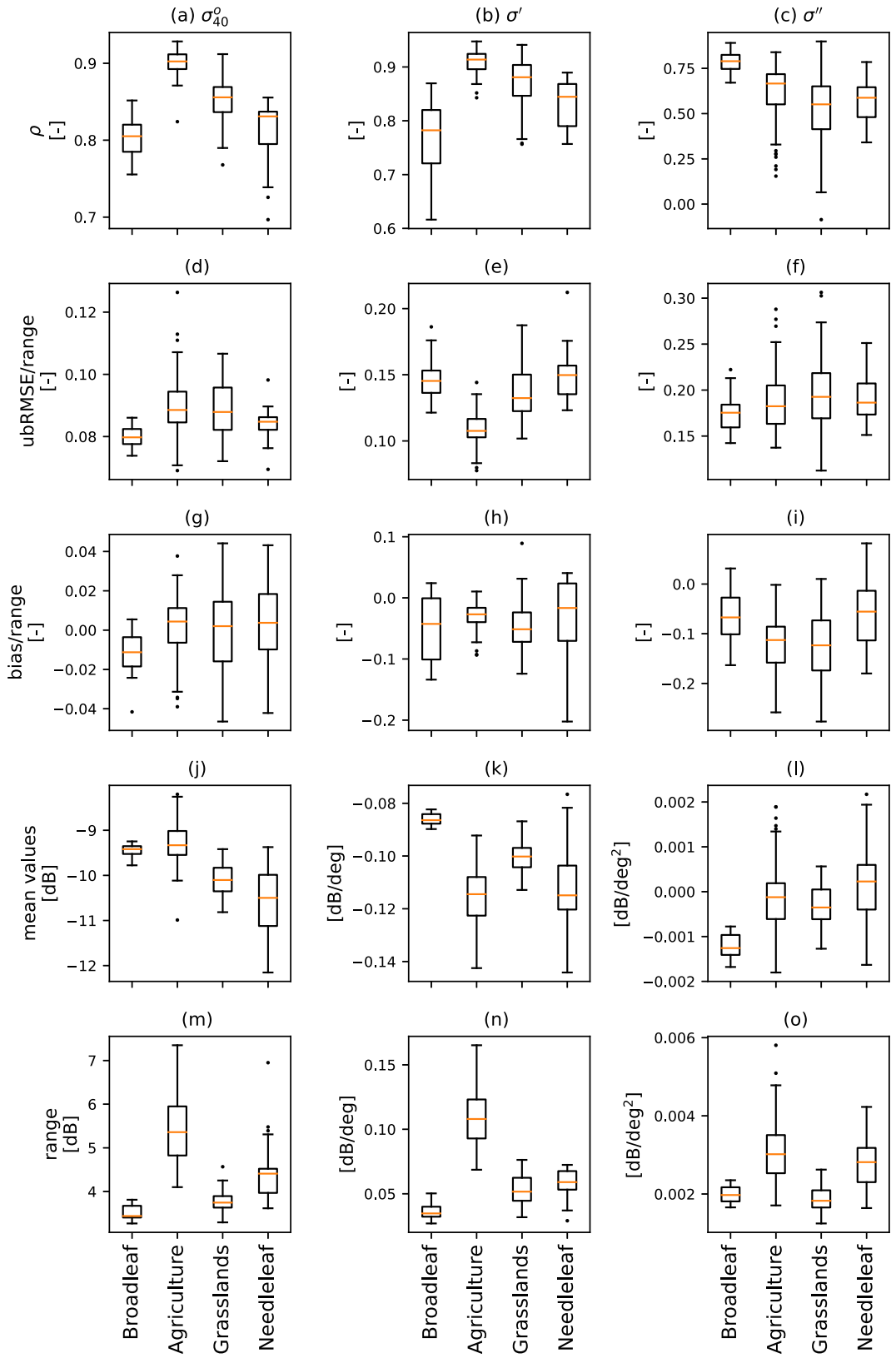

Fig. 4. DNN performance during the validation period (2017–2019) for each of the four ROIs. The columns (left to right) correspond to σ^o_{40}, σ' and σ'' respectively. The top three rows correspond to the Pearson correlation coefficient (a-c), ubRMSE normalized by dynamic range (d-f), bias normalized by dynamic range (g-i). The last two rows show the mean and range of the quantities of interest per ROI. The box extends from the lower to upper quartile values of the data, with a line at the median. The lower whisker is at the lowest datum above Q1–1.5*(Q3-Q1), and the upper whisker at the highest datum below Q3 + 1.5*(Q3-Q1), where Q1 and Q3 are the first and third quartiles. Outliers are indicated as grey dots.

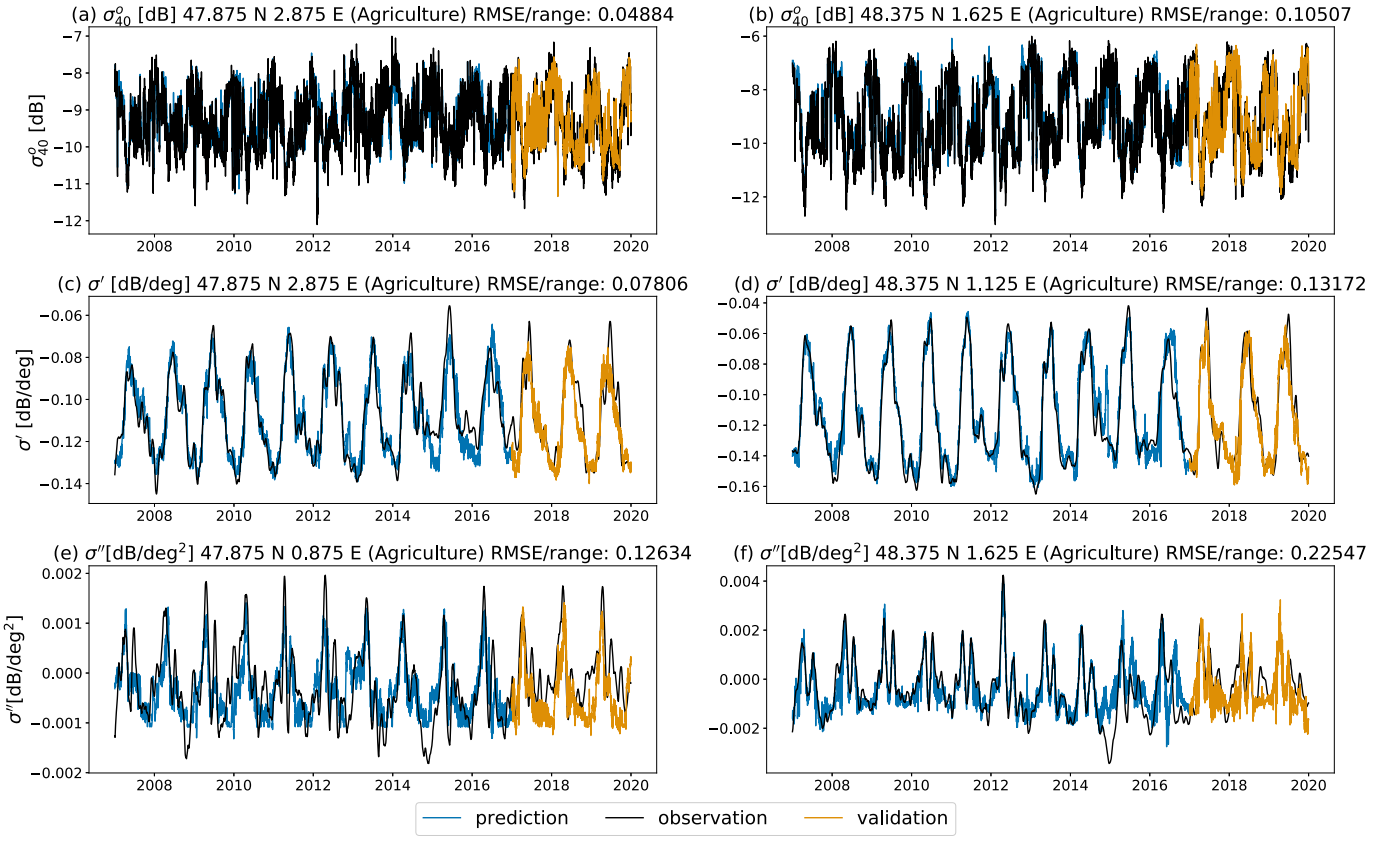

Fig. 5. Time series of predicted (blue) vs observed (black) values of ASCAT normalized backscatter, slope and curvature in the agriculture ROI. The validation period (yellow) is from 2017 to 2019, shown as the brown one. The"best" (left column) and"worst" (right column) grid points were selected based on the maximum and minimum values of RMSE/range among GPIs in the agricultural ROI. The rows show ASCAT normalized backscatter, slope and curvature respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

closely in the Agricultural ROI. However, they diverge after the LAI peak because ISBA does not account for agricultural management practices (including harvest) and their effect on simulated LAI (Shamambo, 2020). In contrast, the limited seasonal variation in WG2 and LAI in the Needleleaf ROI translate to limited variations in both $\mathring{\sigma}_{40}$ and σ'. In Fig. 6 (d), $\mathring{\sigma}_{40}$ primarily follows WG2 with some seasonal influence of LAI. The agreement between σ' and LAI in Fig. 6 (l) is consistent with slope being an indicator of vegetation density.

The seasonality of WG2 and LAI is similar in Broadleaf and Grassland ROIs due to their geographical proximity and similar climate. More importantly, Fig. S1 shows that the Broadleaf ROI comprised 38% (median value) temperate broadleaf cold-deciduous summergreen forest (TEBD) and 30% C4 grass. Similarly, around 20% of the grassland ROI consists of TEBD. Therefore, $\mathring{\sigma}_{40}$ and σ' in both ROIs are effectively a weighted mixture of the scattering from both TEBD and C4 grass. Comparing Fig. 6 (a) and (c), the seasonal cycle of $\mathring{\sigma}_{40}$ is larger in grassland because the vegetation is less dense, allowing a stronger response to soil moisture variations. The seasonal cycle in LAI is, nonetheless, sufficient to introduce some change in the sensitivity to soil moisture. In Fig. 6 (k) the slope follows LAI reasonably well in Grassland, but there is clearly some influence from the TEBD fraction. Slope is generally higher and more constant in the Broadleaf ROI (Fig. 6(i) due to the presence of higher biomass throughout the year.

Fig. 6(i) shows a clear"spring peak" in σ' before the LAI peak is clearly visible in the broadleaf ROI. A similar feature is visible in all cover types but it is on the rising limb of a much stronger seasonal cycle and therefore less significant in the overall cycle. It is also apparent in Fig. 6 (k) due to the presence of some TEBD in the grassland grid cells. Fig. 6 shows that the spring peak is not merely an artefact due to the uncertainty of ASCAT processing to obtain σ'·Pfeil et al. (2020) has identified this spring peak in broadleaf deciduous forests across Europe and attributed it to the elevated water content of twigs and branches prior to leaf out.

3.3. *Model sensitivity: normalized sensitivity coefficient*

The values of NSC(σ^o_{40}, WG2) in Fig. 7 (a - f) are consistent with our expectations in terms of the sensitivity of σ^o_{40} to surface soil moisture. NSC(σ^o_{40}, WG2) is generally positive because backcatter increases with soil moisture, and it is highest during the winter months when the reduced vegetation cover leads to an increased sensitivity of σ^o_{40} to WG2. Comparing Fig. 7(a - f) to Fig. 7 (g - l), it is clear that σ^o_{40} is considerably more sensitive to WG2 than to LAI. Values of NSC(σ^o_{40}, LAI) are generally closer to zero. NSC(σ^o_{40}, LAI) is highest, and comparable to the values of NSC(σ^o_{40}, WG2) in May and July in the west indicating that σ^o_{40} during this period is therefore sensitive to both.

Fig. 7(m - r) show that σ' is highly sensitive to LAI. NSC(σ', LAI) is generally positive in winter with the exception of areas with needleleaf forest. Striking spatial patterns emerge in May (Fig. 7(o)) and July (Fig. 7(p)) that correspond to the land cover patterns in Fig. 1. NSC(σ', LAI) reaches exceptionally high values in May, before becoming equally large in magnitude but negative in July (Fig. 7(p).

Fig. 8 shows seasonal variations in the sensitivity of normalized backscatter to WG2 and LAI in the four ROIs. In all cover types, NSC(σ^o_{40}, WG2) is highest during winter and lowest during the summer months when higher biomass reduces the sensitivity of C-band backscatter to soil moisture. NSC(σ^o_{40}, LAI) is generally lower than NSC(σ^o_{40}, WG2), indicating that σ^o_{40} is more sensitive to WG2 than LAI. The only

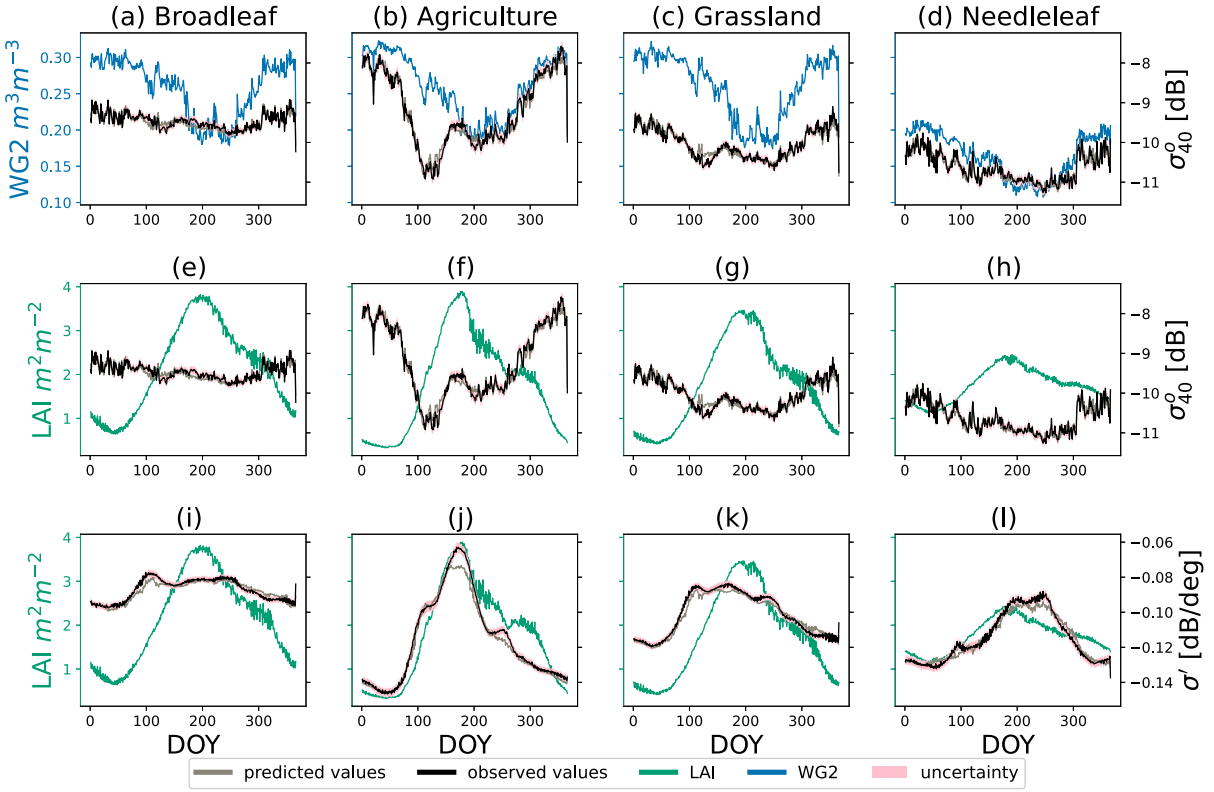

Fig. 6. Time series of WG2, LAI w.r.t. normalized backscatter and slope averaged from 2007 to 2019. The predicted value of normalized backscatter (slope) is the grey line and the observed normalized backscatter (slope) is the black line. The error estimate for each ASCAT observable is indicated by pink shading. Blue and green lines are LSVs of interest, e.g. WG2 (a - d) and LAI (e - l), respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

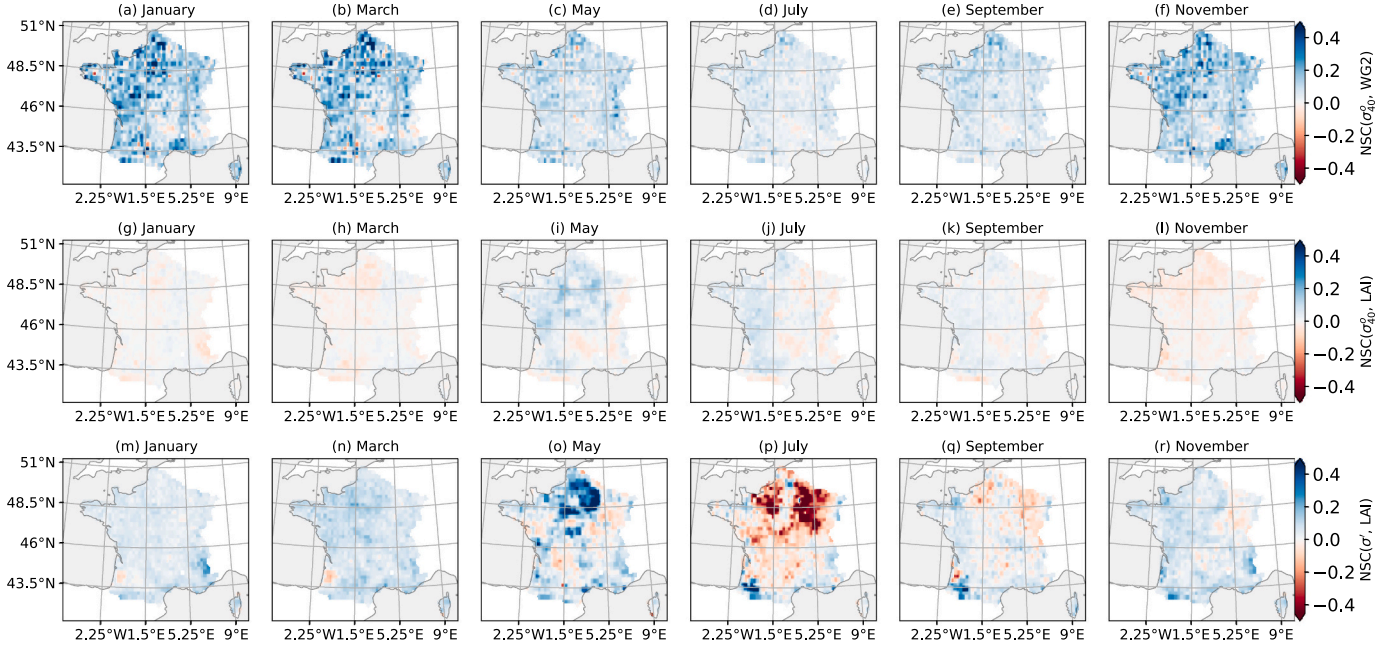

Fig. 7. Maps of NSC values of normalized backscatter and slope w.r.t WG2 and LAI. The columns show different months. The rows show different NSC, i.e. NSC(σ^o_{40}, WG2) (a - f), NSC(σ^o_{40}, LAI) (g - l) and NSC(σ' WG2) (g - l).

exception is in the Agriculture ROI during spring when the backscatter dynamics are primarily driven by the increase in biomass during the vegetative stages of the crops, resulting in NSC(σ^o_{40}, LAI) higher than NSC(σ^o_{40}, WG2).

The strongest seasonal cycles in both NSC(σ^o_{40}, WG2) and NSC(σ^o_{40}, LAI) are observed in the Agriculture ROI. This is due to the contrast between bare soil in winter and a vegetated surface during the crop growing season. σ^o_{40} has a strong seasonal cycle (range of 3 dB) which is

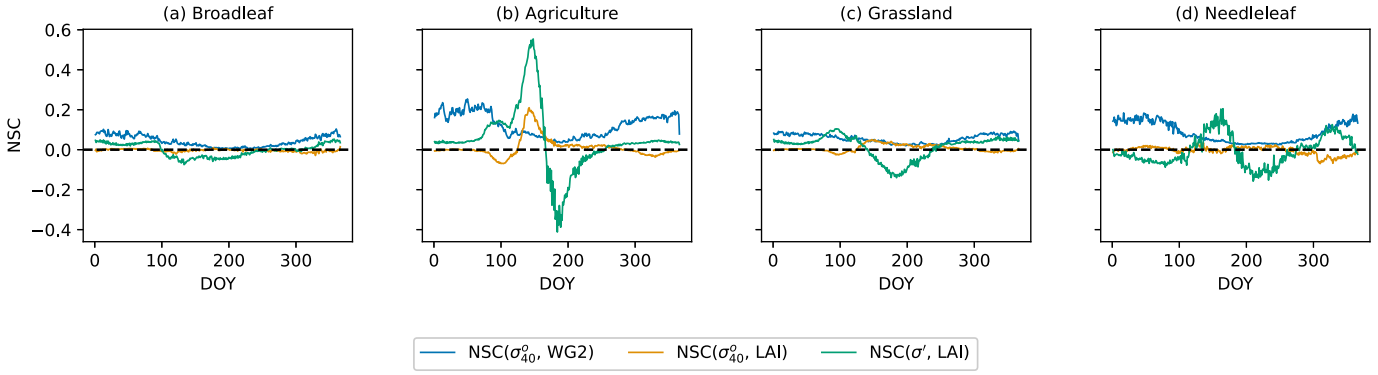

Fig. 8. Seasonal cycle of NSC values in Broadleaf (a), Agriculture (b), Grassland (c) and Needleleaf (d) regions of interest. The values correspond to the median value among the GPIs, and were calculated using all data from 2007 to 2019.

primarily influenced by WG2 in the bare period (DOY 0–100) and DOY (300–365), and a combination of WG2 and LAI in the growing season (DOY 100–300). NSC(σ^o_{40}, WG2) is higher in the Needleleaf ROI than in the Broadleaf ROI because the variations in σ^o_{40} are higher in the Needleleaf ROI and they closely follow WG2 variations (Fig. 6 (d)). This sensitivity to soil moisture under a needleleaf forest may be due to the multiple scattering between the trunks and soil during low LAI periods. This is similar to results found in L-band backscatter modeling studies in boreal forested areas (Tabatabaeenejad et al., 2012).

In Fig. 8 (a-d), NSC(σ', LAI) also show the sensitivity of slope to LAI. Fig. 8 (a) shows that NSC(σ', LAI) is very low in the Broadleaf ROI. This is because the slope is high throughout the year in the Broadleaf ROI and exhibits little seasonal variation despite the large change in LAI (Fig. 6 (i)). Conversely, a stronger seasonal cycle is observed in Needleleaf ROI 8 (a) because a relatively small change in LAI and stronger seasonal cycle in slope is observed in Needleleaf forest than in Broadleaf forest. The strongest seasonal cycle in (NSC(σ', LAI)) is observed in the Agriculture ROI where the biomass variations (reflected in LAI) during the crop growing season change the dominant scattering mechanism from surface scattering in winter to volumetric and multiple scattering in summer, resulting in a strong seasonal change in slope. Note that there are sign changes in (NSC(σ', LAI)) in all ROIs. From Fig. 6(i) to (l), these correspond to the difference in timing between the peaks in LAI and those in slope. Combined, Fig. 6(i) to (l) and Fig. 8 suggest that slope is sensitive to more than just LAI alone.

Results are presented in Fig. 9 to explain why the"spring peak" feature in Fig. 6 (i) is captured by the DNN despite the absence of internal vegetation water content in the DNN inputs. Fig. 9 (a) shows that the spring peak in slope coincides with a rapid decrease in soil moisture (WG) in all soil layers. The increase in slope from DOY 160 to 250 also coincides with a decrease in soil moisture throughout the profile. Fig. 9 (b) shows that the local maximum in slope on DOY 198 coincides with the peak in LAI, but that the spring peak coincides with a local peak in net radiation around DOY 105. Collectively, Fig. 9 (a) and (b) illustrate that slope variations reflect vegetation growth and are influenced by both water and energy availability in this Broadleaf ROI.

Comparing Fig. 9 (c) and (d), it is clear that the NSCs related to WG are larger in magnitude than those related to net radiation and LAI. From Fig. 9 (c), NSC(σ', WG2) is close to zero, but the magnitude of NSC with respect to soil moisture increases with depth reaching values of −0.2 for WG7. Similarity among the NSC values with respect to WG in different layers is due to the strong correlation in soil moisture in the various

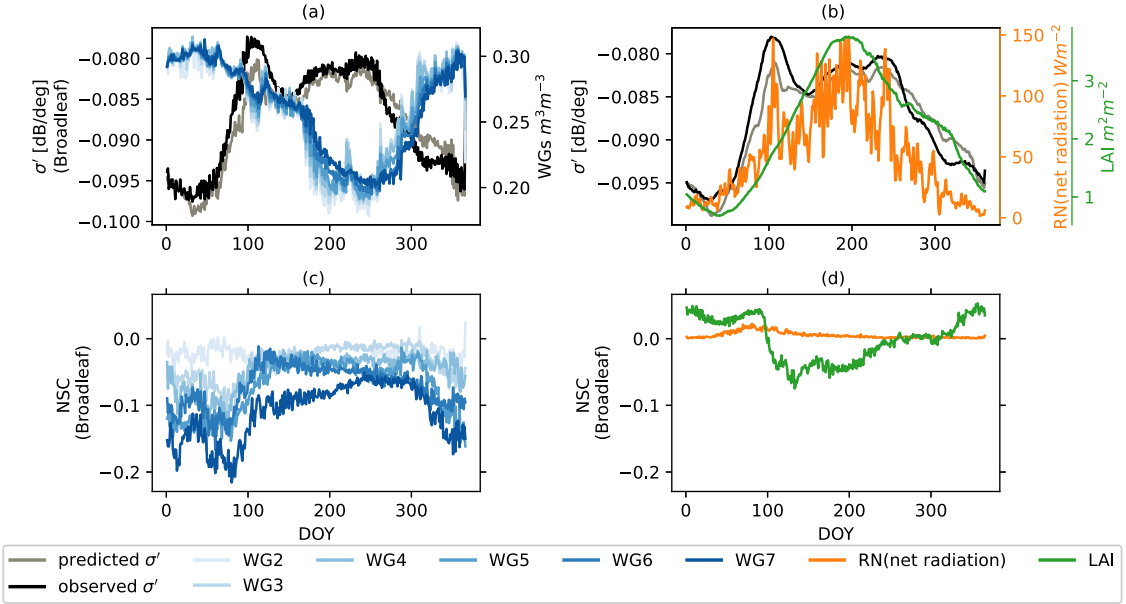

Fig. 9. Seasonal cycle of LSVs (including soil moisture from different depths (different WGs) (a) and vegetation related LSVs (LAI and RN) (b)) and their NSC values (c - d) of slope in Broadleaf regions of interest. The values correspond to the median value among the GPIs, and were calculated using all data from 2007 to 2019. The vegetation related LSVs have been smoothed by a 7-day window in order to compare their peaks with peak in σ'.

layers. The magnitude of of NSC(σ', LAI) is around half that of NSC with respect to soil moisture. While NSC(σ', RN) is non-zero and positive in spring, the magnitude is an order of magnitude smaller than that of NSC (σ', WG7). This suggests that net radiation may play a part in the dynamics of slope, but that the connection is indirect and weak. This connection between deep soil moisture and slope is compatible with the conclusion of Pfeil et al. (2020) that the slope is sensing the filling of branch and twigs prior to leaf-out. It is striking that the combination of ISBA and the DNN can predict the observed spring peak by capturing the soil moisture dynamics and mapping it directly to slope change without explicitly modeling the water storage in the vegetation.

4. Conclusions

A DNN was trained per grid point to relate ISBA LSVs to ASCAT normalized backscatter, slope and curvature for future use as the observation operator within a data assimilation framework. Agreement between the predicted and observed ASCAT observables during the independent validation period was very good, with better estimates obtained for backscatter and slope than for curvature. Given that we have a better understanding of the factors affecting backscatter and slope, and our expectation that curvature is influenced by vegetation geometry, this is perhaps not too surprising. Results also showed that land cover types have a large effect on model performance. Recall that an individual DNN model is obtained for every grid point. Therefore, smooth spatial patterns in the performance metrics, and the appearance of features related to dominant land cover and fraction of land cover provide an indication of robustness.

Performance could be improved through the inclusion of additional LSVs to capture controls on ASCAT observables that are currently missing. For example, the relatively poor performance in the Alps is partly because snow states were not included. Snow has a strong influence on backscatter and the seasonal variation in the backscatter-incidence angle relationship. The inclusion of temperature and radiation data provides information that there is a seasonal influence affecting the backscatter, but the explicit inclusion of snow states should lead to an improvement in performance in areas affected by snow cover. Performance could also be improved by adapting the ML implementation, or replacing the DNN with an alternative machine learning approach. The variability of predicted slope and curvature are generally larger than the observations. While this may partly be due to overfitting, it is primarily because the slope and curvature observations are obtained by aggregating local slope values in a 42-day window, i.e. an Epanechnikov kernel applied to the local slopes (Hahn et al., 2017). The time series of the observations are therefore, by definition, very smooth. In contrast, the DNN is trained with daily input data. One pragmatic solution to reduce the noise in the predicted observables would be to filter them a posteriori. Alternatively, one could consider approaches such as Long-Short Term Memory (LSTM) to handle the mapping from a temporal series of inputs to the ASCAT observables. Using LSTM or similar would also provide a means to explicitly handle the strong auto- and cross-correlations in the soil states. Furthermore, these correlations also mean that there is considerable redundancy in the list of input variables. While there is no requirement that the input variables are uncorrelated (Lu et al., 2017), performing dimensional reduction to provide a reduced set of inputs to the DNN would reduce the computational burden of the training step. One inherent limitation of DNN, and any data-driven approach, is that low probability events receive small weights and can be neglected during training and prediction. This was illustrated by the poor performance of the DNN in the wake of Cyclone Klaus, for example. In future research, attention must be paid to how low probability events can be flagged and handled in an assimilation system.

The NSC analysis indicates that the backscatter is very sensitive to surface soil moisture and LAI, and that these sensitivities vary during the year and among in different land covers. The observed spatial and temporal variations in the backscatter NSCs are consistent with our understanding of backscatter from theory, observations and modeling. This provides an important"sanity check" for our DNN, that it is capable of capturing known sensitivities. NSC values also indicate that the slope is sensitive to LAI but also to other quantities including root zone soil moisture variations or related quantities. Backscatter is sensitive to vegetation water content which is not directly accounted for in ISBA, but varies as water is transported from the root zone to the atmosphere through the vegetation as modeled in plant hydraulics model (Li et al., 2021). Note, however, that the NSC values are influenced by the choice of LSVs included as inputs to the DNN. In particular, our inability to explicitly include surface roughness and vegetation water content in the DNN inputs produces an increase in sensitivity of the DNN to other ISBA LSVs such as root zone soil moisture. This means that we are effectively using root zone soil moisture to compensate for the absence of vegetation water content data.

The NSC values together with the LSV variations provide new insights into the observables themselves. This was illustrated for the spring peak in slope in broadleaf deciduous forest. The DNN model was able to predict the spring peak in slope before the LAI peak. The NSC and LSV time series show that the spring peak is related to radiation availability and root water uptake, consistent with the hypothesis of Pfeil et al. (2020) that the spring peak could be attributed to increased water content in the branches prior to leaf-out. These results are significant because we can now describe the slope and curvature variations in terms of LSVs rather than scattering mechanisms. This brings us a step closer to being able to use them directly in applications or to constrain models.

One of our primary motivations was to develop a DNN model to use as a measurement operator for assimilation. Results suggest that the DNN model is capable of predicting the observations very well. The NSC analysis provides insight into the information content of the slope and curvature, and which LSVs should be included in the state vector for update. The direct mapping of LSVs to observables circumvents two key challenges related to the use of radiative transfer models in assimilation. First, it circumvents the challenge of translating the LSVs provided by the physical process model to the set of inputs required by the radiative transfer model. Second, the direct mapping between the model and observables ensures that there is a consistent climatology between the predicted and actual observables, obviating the need for CDF-matching or bias correction. Furthermore, this approach lends itself to multi-observation assimilation as the list of LSVs and observables can easily be expanded to produce a measurement operator to ingest data from multiple sources. Future research will investigate the performance of the DNN as a measurement operator in an assimilation context. In particular, additional research is needed to determine the degree to which assimilation using a DNN-based measurement operator can outperform an existing approach (e.g. Water Cloud Model) in terms of accuracy and practicality.

Credit author statement

X.S, S.S-D and M.H. were responsible for the conceptualization, methodology, formal analysis, investigation, visualization and writing (original draft preparation). S.H. and W.W. provided resources (ASCAT data). B.B., C.A. and J-C.C. provided resources (ISBA model simulations). O.K. and S.G. contributed to the methodology and investigation. All authors contributed to the analysis of results and writing (review and editing).

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.

Acknowledgement

X.S and S·S-D are supported by Dutch Research Foundation (NWO) - User Support Programme Space Research [grant number ALWGO.2018.036] with support from the Netherlands eScience Center [Grant number NLeSC C19.007]. Ths work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. X.S, S·S-D and M.H. were responsible for the conceptualization, methodology, formal analysis, investigation, visualization and writing (original draft preparation). S.H. and W.W. provided resources (ASCAT data). B.B., C. A. and J-C.C. provided resources (ISBA model simulations). O.K. and S. G. contributed to the methodology and investigation. All authors contributed to the analysis of results and writing (review and editing). The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2022.113116.

References

Aires, F., Weston, P., Rosnay, P., Fairbairn, D., 2021. Statistical approaches to assimilate ASCAT soil moisture information—i. methodologies and first assessment. Q. J. R. Meteorol. Soc. https://doi.org/10.1002/qj.3997.

Albergel, C., Munier, S., Leroux, D.J., Dewaele, H., Fairbairn, D., Barbu, A.L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., Moigne, P.L., Decharme, B., Mahfouf, J.F., Calvet, J.C., 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. 10, 3889–3912. https://doi.org/10.5194/gmd-10-3889-2017.

Albergel, C., Munier, S., Bocher, A., Bonan, B., Zheng, Y., Draper, C., Leroux, D., Calvet, J.C., 2018. LDAS-monde sequential assimilation of satellite derived observations applied to the contiguous US: an ERA-5 driven reanalysis of the land surface variables. Remote Sens. 10, 1627. https://doi.org/10.3390/rs10101627.

Albergel, C., Dutra, E., Bonan, B., Zheng, Y., Munier, S., Balsamo, G., de Rosnay, P., Muñoz-Sabater, J., Calvet, J.C., 2019. Monitoring and forecasting the impact of the 2018 summer heatwave on vegetation. Remote Sens. 11, 520. https://doi.org/10.3390/rs11050520.

Anderegg, W.R.L., Wolf, A., Arango-Velez, A., Choat, B., Chmura, D.J., Jansen, S., Kolb, T., Li, S., Meinzer, F., Pita, P., de Dios, V.R., Sperry, J.S., Wolfe, B.T., Pacala, S., 2017. Plant water potential improves prediction of empirical stomatal models. PLoS One 12, e0185481. https://doi.org/10.1371/journal.pone.0185481.

Ball, J.T., Woodrow, I.E., Berry, J.A., 1987. A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In: Progress in Photosynthesis Research. Springer, Netherlands, pp. 221–224. https://doi.org/10.1007/978-94-017-0519-6_48.

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

Barbu, A.L., 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, 18, pp. 173–192. https://doi.org/10.5194/hess-18-173-2014.

Bengio, Y., 2009. Learning deep architectures for AI. Foundations Trend Mach. Learn. 2, 1–127. https://doi.org/10.1561/2200000006.

Bianchini, M., Scarselli, F., 2014. On the complexity of neural network classifiers: a comparison between shallow and deep architectures. IEEE Trans. Neural Netw. Learn. Syst. 25, 1553–1565. https://doi.org/10.1109/tnnls.2013.2293637.

Bittner, S., Janott, M., Ritter, D., Köcher, P., Beese, F., Priesack, E., 2012. Functional–structural water flow model reveals differences between diffuse- and ring-porous tree species. Agric. For. Meteorol. 158-159, 80–89. https://doi.org/10.1016/j.agrformet.2012.02.005.

Bohrer, G., Mourad, H., Laursen, T.A., Drewry, D., Avissar, R., Poggi, D., Oren, R., Katul, G.G., 2005. Finite element tree crown hydrodynamics model (FETCH) using porous media flow within branching elements: a new representation of tree hydrodynamics. Water Resour. Res. 41 https://doi.org/10.1029/2005wr004181.

Bonan, G.B., Williams, M., Fisher, R.A., Oleson, K.W., 2014. Modeling stomatal conductance in the earth system: linking leaf water-use efficiency and water transport along the soil-plant-atmosphere continuum. Geosci. Model Dev. Discuss. 7, 3085–3159. https://doi.org/10.5194/gmdd-7-3085-2014.

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

Brocca, L., Crow, W.T., Ciabatta, L., Massari, C., de Rosnay, P., Enenkel, M., Hahn, S., Amarnath, G., Camici, S., Tarpanelli, A., Wagner, W., 2017. A review of the applications of ASCAT soil moisture products. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2285–2306. https://doi.org/10.1109/jstars.2017.2651140.

Calvet, J.C., Soussana, J.F., 2001. Modelling CO2-enrichment effects using an interactive vegetation SVAT scheme. Agric. For. Meteorol. 108, 129–152. https://doi.org/10.1016/s0168-1923(01)00235-0.

Calvet, J.C., Noilhan, J., Roujean, J.L., Bessemoulin, P., Cabelguenne, M., Olioso, A., Wigneron, J.P., 1998. An interactive vegetation SVAT model tested against data from six contrasting sites. Agric. For. Meteorol. 92, 73–95. https://doi.org/10.1016/s0168-1923(98)00091-4.

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

Calvet, J.C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., 2007. In situ soil moisture observations for the CAL/VAL of SMOS: The SMOSMANIA network. In: 2007 IEEE International Geoscience and Remote Sensing Symposium. IEEE. https://doi.org/10.1109/igarss.2007.4423019.

Castruccio, S., McInerney, D.J., Stein, M.L., Crouch, F.L., Jacob, R.L., Moyer, E.J., 2014. Statistical emulation of climate model projections based on precomputed gcm runs. J. Clim. 27, 1829–1844. https://doi.org/10.1175/jcli-d-13-00099.1.

Chevallier, F., Mahfouf, J.F., 2001. Evaluation of the Jacobians of Infrared Radiation Models for Variational Data Assimilation, 40, pp. 1445–1461. https://doi.org/10.1175/1520-0450(2001)040<1445:eotjoi>2.0.co;2.

Chevallier, F., Chéruy, F., Scott, N.A., Chédin, A., 1998. A neural network approach for a fast and accurate computation of a longwave radiative budget. J. Appl. Meteorol. 37, 1385–1397. https://doi.org/10.1175/1520-0450(1998)037<1385:annafa>2.0.co;2.

Damour, G., Simonneau, T., Cochard, H., Urban, L., 2010. An overview of models of stomatal conductance at the leaf level. Plant Cell Environ. https://doi.org/10.1111/j.1365-3040.2010.02181.x.

Decharme, B., Martin, E., Faroux, S., 2013. Reconciling soil thermal and hydrological lower boundary conditions in land surface models. J. Geophys. Res.-Atmos. 118, 7819–7834.

Dewaele, H., Munier, S., Albergel, C., Planque, C., Laanaia, N., Carrer, D., Calvet, J.C., 2017. Parameter optimisation for a better representation of drought by LSMs: inverse modelling vs. sequential data assimilation. Hydrol. Earth Syst. Sci. 21, 4861–4878. https://doi.org/10.5194/hess-21-4861-2017.

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

Entekhabi, D., Yueh, S., O'Neill, P.E., Kellogg, K.H., Allen, A., Bindlish, R., Brown, M., Chan, S., Colliander, A., Crow, W.T., et al., 2014. Smap Handbook–Soil Moisture Active Passive: Mapping Soil Moisture and Freeze/Thaw from Space.

Fairbairn, D., Barbu, A.L., Napoly, A., Albergel, C., Mahfouf, J.F., Calvet, J.C., 2017. The effect of satellite-derived surface soil moisture and leaf area index land data assimilation on streamflow simulations over France. Hydrol. Earth Syst. Sci. 21, 2015–2033. https://doi.org/10.5194/hess-21-2015-2017.

Faroux, S., Tchuenté, A.T.K., Roujean, J.L., Masson, V., Martin, E., Moigne, P.L., 2013. ECOCLIMAP-II/europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. Geosci. Model Dev. 6, 563–582. https://doi.org/10.5194/gmd-6-563-2013.

Fer, I., Kelly, R., Moorcroft, P.R., Richardson, A.D., Cowdery, E.M., Dietze, M.C., 2018. Linking big models to big data: efficient ecosystem model calibration through bayesian model emulation. Biogeosciences 15, 5801–5830. https://doi.org/10.5194/bg-15-5801-2018.

Ferrazzoli, P., Guerriero, L., Schiavon, G., 1997. A vegetation classification scheme validated by model simulations. In: IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development. IEEE. https://doi.org/10.1109/igarss.1997.608984.

Fillion, L., Mahfouf, J.F., 2003. Jacobians of an Operational Prognostic Cloud Scheme, 131, pp. 2838–2856. https://doi.org/10.1175/1520-0493(2003)131<2838:joaopc>2.0.co;2.

Fisher, R.A., Koven, C.D., Anderegg, W.R.L., Christoffersen, B.O., Dietze, M.C., Farrior, C. E., Holm, J.A., Hurtt, G.C., Knox, R.G., Lawrence, P.J., Lichstein, J.W., Longo, M., Matheny, A.M., Medvigy, D., Muller-Landau, H.C., Powell, T.L., Serbin, S.P., Sato, H., Shuman, J.K., Smith, B., Trugman, A.T., Viskari, T., Verbeeck, H., Weng, E., Xu, C., Xu, X., Zhang, T., Moorcroft, P.R., 2017. Vegetation demographics in earth system models: a review of progress and priorities. Glob. Chang. Biol. 24, 35–54. https://doi.org/10.1111/gcb.13910.

Fontanelli, G., Paloscia, S., Zribi, M., Chahbi, A., 2013. Sensitivity analysis of x-band SAR to wheat and barley leaf area index in the merguellil basin. Remote Sens. Lett. 4, 1107–1116. https://doi.org/10.1080/2150704x.2013.842285.

Forman, B.A., Reichle, R.H., 2015. Using a support vector machine and a land surface model to estimate large-scale passive microwave brightness temperatures over snow-covered land in north america. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 4431–4441. https://doi.org/10.1109/jstars.2014.2325780.

Forman, B.A., Xue, Y., 2016. Machine learning predictions of passive microwave brightness temperature over snow-covered land using the special sensor microwave imager (SSM/I). Phys. Geogr. 38, 176–196. https://doi.org/10.1080/02723646.2016.1236606.

Forman, B.A., Reichle, R.H., Derksen, C., 2014. Estimating passive microwave brightness temperature over snow-covered land in north america using a land surface model and an artificial neural network. IEEE Trans. Geosci. Remote Sens. 52, 235–248. https://doi.org/10.1109/tgrs.2013.2237913.

Friesen, J.C., 2008. Regional Vegetation Water Effects on Satellite Soil Moisture Estimations for West Africa.

Friesen, J., Steele-Dunne, S.C., van de Giesen, N., 2012. Diurnal differences in global ERS scatterometer backscatter observations of the land surface. IEEE Trans. Geosci. Remote Sens. 50, 2595–2602. https://doi.org/10.1109/tgrs.2012.2193889.

Frison, P., Mougin, E., Hiernaux, P., 1998. Observations and interpretation of seasonal ERS-1 wind scatterometer data over northern Sahel (Mali). Remote Sens. Environ. 63, 233–242. https://doi.org/10.1016/s0034-4257(97)00137-5.

Garand, L., Turner, D.S., Larocque, M., Bates, J., Boukabara, S., Brunel, P., Chevallier, F., Deblonde, G., Engelen, R., Hollingshead, M., Jackson, D., Jedlovec, G., Joiner, J., Kleespies, T., McKague, D.S., McMillin, L., Moncet, J.L., Pardo, J.R., Rayer, P.J., Salathe, E., Saunders, R., Scott, N.A., Delst, P.V., Woolf, H., 2001. Radiance and Jacobian Intercomparison of Radiative Transfer Models Applied to HIRS and AMSU Channels, 106, pp. 24017–24031. https://doi.org/10.1029/2000jd000184.

Gibelin, A.L., Calvet, J.C., Roujean, J.L., Jarlan, L., Los, S.O., 2006. Ability of the land surface model ISBA-a-gs to simulate leaf area index at the global scale: comparison with satellites products. J. Geophys. Res. 111 https://doi.org/10.1029/2005jd006691.

Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., 2016. Deep Learning, vol. 1. MIT press, Cambridge.

Hahn, S., Reimer, C., Vreugdenhil, M., Melzer, T., Wagner, W., 2017. Dynamic characterization of the incidence angle dependence of backscatter using metop ASCAT. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2348–2359. https://doi.org/10.1109/jstars.2016.2628523.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.

Hong, Y., Hsu, K.L., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. J. Appl. Meteorol. 43, 1834–1853. https://doi.org/10.1175/jam2173.1.

Jacobs, C., van den Hurk, B., de Bruin, H., 1996. Stomatal behaviour and photosynthetic rate of unstressed grapevines in semi-arid conditions. Agric. For. Meteorol. 80, 111–134. https://doi.org/10.1016/0168-1923(95)02295-3.

Janott, M., Gayler, S., Gessler, A., Javaux, M., Klier, C., Priesack, E., 2010. A one-dimensional model of water flow in soil-plant systems based on plant architecture. Plant Soil 341, 233–256. https://doi.org/10.1007/s11104-010-0639-0.

Jarlan, L., Mougin, E., Frison, P., Mazzega, P., Hiernaux, P., 2002. Analysis of ERS wind scatterometer time series over Sahel (Mali). Remote Sens. Environ. 81, 404–415. https://doi.org/10.1016/s0034-4257(02)00015-9.

Khabbazan, S., Steele-Dunne, S., Vermunt, P., Judge, J., Vreugdenhil, M., Gao, G., 2022. The influence of surface canopy water on the relationship between l-band backscatter and biophysical variables in agricultural monitoring. Remote Sens. Environ. 268, 112789. https://doi.org/10.1016/j.rse.2021.112789.

Kim, S.B., Moghaddam, M., Tsang, L., Burgin, M., Xu, X., Njoku, E.G., 2014. Models of l-Band Radar Backscattering Coefficients over Global Terrain for Soil Moisture Retrieval, 52, pp. 1381–1396. https://doi.org/10.1109/tgrs.2013.2250980.

Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization arXiv preprint arXiv:1412.6980.

Konings, A.G., Piles, M., Das, N., Entekhabi, D., 2017a. L-band vegetation optical depth and effective scattering albedo estimation from SMAP. Remote Sens. Environ. 198, 460–470. https://doi.org/10.1016/j.rse.2017.06.037.

Konings, A.G., Yu, Y., Xu, L., Yang, Y., Schimel, D.S., Saatchi, S.S., 2017b. Active microwave observations of diurnal and seasonal variations of canopy water content across the humid african tropical forests. Geophys. Res. Lett. 44, 2290–2299. https://doi.org/10.1002/2016gl072388.

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.

Konings, A.G., Saatchi, S.S., Frankenberg, C., Keller, M., Leshyk, V., Anderegg, W.R.L., Humphrey, V., Matheny, A.M., Trugman, A., Sack, L., Agee, E., Barnes, M.L., Binks, O., Cawse-Nicholson, K., Christoffersen, B.O., Entekhabi, D., Gentine, P., Holtzman, N.M., Katul, G.G., Liu, Y., Longo, M., Martinez-Vilalta, J., McDowell, N., Meir, P., Mencuccini, M., Mrad, A., Novick, K.A., Oliveira, R.S., Siqueira, P., Steele-Dunne, S.C., Thompson, D.R., Wang, Y., Wehr, R., Wood, J.D., Xu, X., Zuidema, P.A., 2021. Detecting Forest Response to Droughts with Global Observations of Vegetation Water Content, 27, pp. 6005–6024. https://doi.org/10.1111/gcb.15872.

Lafont, S., Zhao, Y., Calvet, J.C., Peylin, P., Ciais, P., Maignan, F., Weiss, M., 2012. Modelling LAI, surface water and carbon fluxes at high-resolution over France: comparison of ISBA-a-gs and ORCHIDEE. Biogeosciences 9, 439–456. https://doi.org/10.5194/bg-9-439-2012.

Lannoy, G.J.M.D., Reichle, R.H., 2016. 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. https://doi.org/10.1175/jhm-d-15-0037.1.

Lemordant, L., Gentine, P., Swann, A.S., Cook, B.I., Scheff, J., 2018. Critical impact of vegetation physiology on the continental hydrologic cycle in response to increasing CO2. Proc. Natl. Acad. Sci. 115, 4093–4098. https://doi.org/10.1073/pnas.1720712115.

Leroux, D., Calvet, J.C., Munier, S., Albergel, C., 2018. Using satellite-derived vegetation products to evaluate LDAS-monde over the euro-mediterranean area. Remote Sens. 10, 1199. https://doi.org/10.3390/rs10081199.

Li, L., Yang, Z.L., Matheny, A.M., Zheng, H., Swenson, S.C., Lawrence, D.M., Barlage, M., Yan, B., McDowell, N.G., Leung, L.R., 2021. Representation of plant hydraulics in the Noah-MP land surface model: model development and multiscale evaluation. J. Adv. Model. Earth Syst. 13 https://doi.org/10.1029/2020ms002214.

Lievens, H., Martens, B., Verhoest, N., Hahn, S., Reichle, R., Miralles, D., 2017. Assimilation of global radar backscatter and radiometer brightness temperature observations to improve soil moisture and land evaporation estimates. Remote Sens. Environ. 189, 194–210. https://doi.org/10.1016/j.rse.2016.11.022.

Liu, Y.Y., de Jeu, R.A.M., McCabe, M.F., Evans, J.P., van Dijk, A.I.J.M., 2011. Global long-term passive microwave satellite-based retrievals of vegetation optical depth. Geophys. Res. Lett. 38 https://doi.org/10.1029/2011gl048684 n/a–n/a.

Liu, Y., Kumar, M., Katul, G.G., Feng, X., Konings, A.G., 2020. Plant hydraulics accentuates the effect of atmospheric moisture stress on transpiration. Nat. Clim. Chang. 10, 691–695. https://doi.org/10.1038/s41558-020-0781-5.

Liu, Y., Konings, A.G., Kennedy, D., Gentine, P., 2021. Global Coordination in Plant Physiological and Rooting Strategies in Response to Water Stress, 35. https://doi.org/10.1029/2020gb006758.

Lu, Z., Pu, H., Wang, F., Hu, Z., Wang, L., 2017. The expressive power of neural networks: a view from the width. Adv. Neural Inf. Proces. Syst. 30.

Macelloni, G., Paloscia, S., Pampaloni, P., Marliani, F., Gai, M., 2001. The relationship between the backscattering coefficient and the biomass of narrow and broad leaf crops. IEEE Trans. Geosci. Remote Sens. 39, 873–884. https://doi.org/10.1109/36.917914.

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

Matheny, A.M., Bohrer, G., Garrity, S.R., Morin, T.H., Howard, C.J., Vogel, C.S., 2015. Observations of stem water storage in trees of opposing hydraulic strategies. Ecosphere 6. https://doi.org/10.1890/es15-00170.1 art165.

Matheny, A.M., Fiorella, R.P., Bohrer, G., Poulsen, C.J., Morin, T.H., Wunderlich, A., Vogel, C.S., Curtis, P.S., 2016. Contrasting strategies of hydraulic control in two codominant temperate tree species. Ecohydrology 10. https://doi.org/10.1002/eco.1815.

Matheny, A.M., Mirfenderesgi, G., Bohrer, G., 2017. Trait-based representation of hydrological functional properties of plants in weather and ecosystem models. Plant Diversity 39, 1–12. https://doi.org/10.1016/j.pld.2016.10.001.

Mattia, F., Toan, T.L., Picard, G., Posa, F., D'Alessio, A., Notarnicola, C., Gatti, A., Rinaldi, M., Satalino, G., Pasquariello, G., 2003. Multitemporal c-band radar measurements on wheat fields. IEEE Trans. Geosci. Remote Sens. 41, 1551–1560. https://doi.org/10.1109/tgrs.2003.813531.

McCuen, R.H., et al., 2005. Hydrologic Analysis and Design, vol. 888. Pearson prentice hall, Upper Saddle River, NJ.

Melzer, T., 2013. Vegetation modelling in warp 6.0, in. In: Proceedings of the EUMETSAT Meteorological Satellite Conference, Vienna, Austria, pp. 16–20.

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.

Noilhan, J., Mahfouf, J.F., 1996. The ISBA land surface parameterisation scheme. Glob. Planet. Chang. 13, 145–159. https://doi.org/10.1016/0921-8181(95)00043-7.

Noilhan, J., Planton, S., 1989. A simple parameterization of land surface processes for meteorological models. Mon. Weather Rev. 117, 536–549. https://doi.org/10.1175/1520-0493(1989)117<0536:aspols>2.0.co;2.

Petchiappan, A., Steele-Dunne, S.C., Vreugdenhil, M., Hahn, S., Wagner, W., Oliveira, R., 2021. The Influence of Vegetation Water Dynamics on the ASCAT Backscatter-Incidence Angle Relationship in the Amazon. https://doi.org/10.5194/hess-2021-406.

Pfeil, I., Wagner, W., Forkel, M., Dorigo, W., Vreugdenhil, M., 2020. Does ASCAT observe the spring reactivation in temperate deciduous broadleaf forests? Remote Sens. Environ. 250, 112042 https://doi.org/10.1016/j.rse.2020.112042.

Powell, T.L., Galbraith, D.R., Christoffersen, B.O., Harper, A., Imbuzeiro, H.M.A., Rowland, L., Almeida, S., Brando, P.M., da Costa, A.C.L., Costa, M.H., Levine, N.M., Malhi, Y., Saleska, S.R., Sotta, E., Williams, M., Meir, P., Moorcroft, P.R., 2013. Confronting model predictions of carbon fluxes with measurements of Amazon forests subjected to experimental drought. New Phytol. 200, 350–365. https://doi.org/10.1111/nph.12390.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven earth system science. Nature 566, 195–204. https://doi.org/10.1038/s41586-019-0912-1.

Rodríguez-Fernández, N., de Rosnay, P., Albergel, C., Richaume, P., Aires, F., Prigent, C., Kerr, Y., 2019. SMOS neural network soil moisture data assimilation in a land surface model and atmospheric impact. Remote Sens. 11, 1334. https://doi.org/10.3390/rs11111334.

Rüdiger, C., Albergel, C., Mahfouf, J.F., Calvet, J.C., Walker, J.P., 2010. Evaluation of the observation operator jacobian for leaf area index data assimilation with an extended kalman filter. J. Geophys. Res. 115. https://doi.org/10.1029/2009jd012912.

Schmidhuber, J., 2015. Deep learning in neural networks: an overview. Neural Netw. 61, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003.

Schroeder, R., McDonald, K.C., Azarderakhsh, M., Zimmermann, R., 2016. ASCAT MetOp-a diurnal backscatter observations of recent vegetation drought patterns over

the contiguous u.s.: an assessment of spatial extent and relationship with precipitation and crop yield. Remote Sens. Environ. 177, 153–159. https://doi.org/10.1016/j.rse.2016.01.008.
Scott, C.A., Bastiaanssen, W.G., Ahmad, M.U.D., 2003. Mapping root zone soil moisture using remotely sensed optical imagery. J. Irrig. Drain. Eng. 129, 326–335.
Seabold, S., Perktold, J., 2010. statsmodels: Econometric and statistical modeling with python. In: 9th Python in Science Conference.
Shamambo, D.C., 2020. Assimilation de données satellitaires pour le suivi des ressources en eau dans la zone Euro-Méditerranée. Ph.D. thesis. Toulouse 3.
Shamambo, D., Bonan, B., Calvet, J.C., Albergel, C., Hahn, S., 2019. Interpretation of ASCAT radar scatterometer observations over land: a case study over southwestern France. Remote Sens. 11, 2842. https://doi.org/10.3390/rs11232842.
Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical bayesian optimization of machine learning algorithms. In: Advances in Neural Information Processing Systems, pp. 2951–2959.
Sperry, J.S., Love, D.M., 2015. What plant hydraulics can tell us about responses to climate-change droughts. New Phytol. 207, 14–27. https://doi.org/10.1111/nph.13354.
Sperry, J.S., Wang, Y., Wolfe, B.T., Mackay, D.S., Anderegg, W.R.L., McDowell, N.G., Pockman, W.T., 2016. Pragmatic hydraulic theory predicts stomatal responses to climatic water deficits. New Phytol. 212, 577–589. https://doi.org/10.1111/nph.14059.
Steele-Dunne, S.C., Friesen, J., van de Giesen, N., 2012. Using diurnal variation in backscatter to detect vegetation water stress. IEEE Trans. Geosci. Remote Sens. 50, 2618–2629. https://doi.org/10.1109/tgrs.2012.2194156.
Steele-Dunne, S.C., Hahn, S., Wagner, W., Vreugdenhil, M., 2019. Investigating vegetation water dynamics and drought using metop ASCAT over the north american grasslands. Remote Sens. Environ. 224, 219–235. https://doi.org/10.1016/j.rse.2019.01.004.
Steele-Dunne, S.C., Hahn, S., Wagner, W., Vreugdenhil, M., 2021. Towards including dynamic vegetation parameters in the EUMETSAT H SAF ASCAT soil moisture products. Remote Sens. 13, 1463. https://doi.org/10.3390/rs13081463.
Stiles, J., Sarabandi, K., Ulaby, F., 2000. Electromagnetic scattering from grassland. II. Measurement and modeling results. IEEE Trans. Geosci. Remote Sens. 38, 349–356. https://doi.org/10.1109/36.823930.
Tabatabaeenejad, A., Burgin, M., Moghaddam, M., 2012. Potential of L-band radar for retrieval of canopy and subcanopy parameters of boreal forests. IEEE Trans. Geosci. Remote Sens. 50, 2150–2160. https://doi.org/10.1109/tgrs.2011.2173349.
Tang, G., Long, D., Behrangi, A., Wang, C., Hong, Y., 2018. Exploring deep neural networks to retrieve rain and snow in high latitudes using multisensor and reanalysis data. Water Resour. Res. 54, 8253–8278. https://doi.org/10.1029/2018wr023830.
Tao, Y., Gao, X., Hsu, K., Sorooshian, S., Ihler, A., 2016. A deep neural network modeling framework to reduce bias in satellite precipitation products. J. Hydrometeorol. 17, 931–945. https://doi.org/10.1175/jhm-d-15-0075.1.
Teubner, I.E., Forkel, M., Jung, M., Liu, Y.Y., Miralles, D.G., Parinussa, R., van der Schalie, R., Vreugdenhil, M., Schwalm, C.R., Tramontana, G., Camps-Valls, G., Dorigo, W.A., 2018. Assessing the relationship between microwave vegetation optical depth and gross primary production. Int. J. Appl. Earth Obs. Geoinf. 65, 79–91. https://doi.org/10.1016/j.jag.2017.10.006.
Teuling, A.J., Taylor, C.M., Meirink, J.F., Melsen, L.A., Miralles, D.G., van Heerwaarden, C.C., Vautard, R., Stegehuis, A.I., Nabuurs, G.J., de Arellano, J.V.G., 2017. Observational evidence for cloud cover enhancement over western European forests. Nat. Commun. 8 https://doi.org/10.1038/ncomms14065.
Ulaby, F.T., Moore, R.K., Fung, A.K., 1986. Microwave Remote Sensing: Active and Passive. Volume 3-from Theory to Applications.
Ulaby, F.T., Sarabandi, K., Mcdonald, K., Whitt, M., Dobson, M.C., 1990. Michigan microwave canopy scattering model. Int. J. Remote Sens. 11, 1223–1253. https://doi.org/10.1080/01431169008955090.
van den Hurk, B.J., Viterbo, P., Beljaars, A., Betts, A., 2000. Offline Validation of the ERA40 Surface Scheme.
van Emmerik, T., Steele-Dunne, S.C., Judge, J., van de Giesen, N., 2015. Impact of diurnal variation in vegetation water content on radar backscatter from maize during water stress. IEEE Trans. Geosci. Remote Sens. 53, 3855–3869. https://doi.org/10.1109/tgrs.2014.2386142.
van Emmerik, T., Steele-Dunne, S., Paget, A., Oliveira, R.S., Bittencourt, P.R.L., Barros, De V.F., van de Giesen, N., 2017. Water stress detection in the amazon using radar. Geophys. Res. Lett. 44, 6841–6849. https://doi.org/10.1002/2017gl073747.
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, 59, pp. 7322–7337. https://doi.org/10.1109/tgrs.2020.3035881.
Verrelst, J., Sabater, N., Rivera, J., Muñoz-Marí, J., Vicent, J., Camps-Valls, G., Moreno, J., 2016. Emulation of leaf, canopy and atmosphere radiative transfer models for fast global sensitivity analysis. Remote Sens. 8, 673. https://doi.org/10.3390/rs8080673.
Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., Strauss, P., 2018. Sensitivity of Sentinel-1 backscatter to vegetation dynamics: an Austrian case study. Remote Sens. 10, 1396. https://doi.org/10.3390/rs10091396.
Wagner, W., Lemoine, G., Rott, H., 1999a. A method for estimating soil moisture from ERS scatterometer and soil data. Remote Sens. Environ. 70, 191–207. https://doi.org/10.1016/s0034-4257(99)00036-x.
Wagner, W., Noll, J., Borgeaud, M., Rott, H., 1999b. Monitoring soil moisture over the Canadian prairies with the ERS scatterometer. IEEE Trans. Geosci. Remote Sens. 37, 206–216. https://doi.org/10.1109/36.739155.
Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., Steinnocher, K., 2013. The ASCAT soil moisture product: a review of its specifications, validation results, and emerging applications. Meteorol. Z. 22, 5–33. https://doi.org/10.1127/0941-2948/2013/0399.
Willis, R., Yeh, W.W.G., 1987. Groundwater Systems Planning and Management.
Xue, Y., Forman, B.A., 2015. Comparison of passive microwave brightness temperature prediction sensitivities over snow-covered land in north america using machine learning algorithms and the advanced microwave scanning radiometer. Remote Sens. Environ. 170, 153–165. https://doi.org/10.1016/j.rse.2015.09.009.
Xue, Y., Forman, B.A., Reichle, R.H., 2018. Estimating snow mass in north america through assimilation of advanced microwave scanning radiometer brightness temperature observations using the catchment land surface model and support vector machines. Water Resour. Res. 54, 6488–6509. https://doi.org/10.1029/2017wr022219.