

# Land Surface Data Assimilation of Satellite Derived Surface Soil Moisture

Towards an Integrated Representation of the Arctic Hydrological Cycle

---

Jostein Blyverket

Thesis for the degree of Philosophiae Doctor (PhD)  
University of Bergen, Norway  
2019

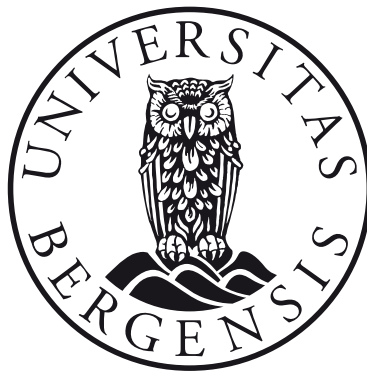
UNIVERSITY OF BERGEN



# Land Surface Data Assimilation of Satellite Derived Surface Soil Moisture

Towards an Integrated Representation of the Arctic  
Hydrological Cycle

Jostein Blyverket



Thesis for the degree of Philosophiae Doctor (PhD)  
at the University of Bergen

Date of defense: 21.10.2019

© Copyright Jostein Blyverket

The material in this publication is covered by the provisions of the Copyright Act.

Year: 2019

Title: Land Surface Data Assimilation of Satellite Derived Surface Soil Moisture

Name: Jostein Blyverket

Print: Skipnes Kommunikasjon / University of Bergen

# Acknowledgements

I would like to thank my supervisor and friend William A. Lahoz, who gave me the opportunity to work on this project in the first place. He was always very supportive of my ideas and responsive whenever I had questions regarding work or life in general. He will be missed.

I am grateful for the support I have received from friends and colleagues at NILU and in particular from Paul Hamer. Paul has provided support and guidance from the beginning of the project to the end. Thanks to Matthias Vogt, Philipp Schneider and Henrik Grythe for fruitful discussions, all your help and thrilling volleyball sessions. I am also grateful to NILU and Britt Ann Kåstad Høiskar for supporting my PhD work, even though I was running a bit late.

For technical help and discussions, I would like to thank Åsmund Bakketun and Trygve Aspelien at MET Norway, David Fairbairn at ECMWF and Sam-Erik Walker and Tove Svendby at NILU. I would like to thank Clément Albergel at Météo-France and Gabrielle De Lannoy at KU Leuven, who both have been very supportive and helpful with my papers and technical issues. To my supervisors in Bergen, Noel and Laurent, thanks for your guidance and help on my papers and the thesis. I am grateful to the Research Council of Norway for supporting my three months stay at Princeton University, where I met inspiring people, such as Eric Wood and Ming Pan. In addition, I am thankful to the Norwegian Space Centre and Hydrologirådet for travel support to conferences.

Thanks to my family and friends for taking my mind off work and their encouragements. Finally, a special thanks to my girlfriend Martine, for your support and patience.



# Abstract

The ability to accurately determine soil water content (soil moisture) over large areas of the Earth's surface has potential implications in meteorology, hydrology, water and natural hazards management. The advent of space-based microwave sensors, found to be sensitive to surface soil moisture, has allowed for long-term studies of soil moisture dynamics at the global scale. There are, however, areas where remote sensing of soil moisture is prone to errors because, e.g., complex topography, surface water, dense vegetation, frozen soil or snow cover affect the retrieval. This is particularly the case for the northern high latitudes, which is a region subject to more rapid warming than the global mean and also is identified as an important region for studying 21<sup>st</sup> century climate change.

Land surface models can help to close these observation gaps and provide high spatio-temporal coverage of the variables of interest. Models are only approximations of the real world and they can experience errors in, for example, their initialization and/or parameterization. In the past 20 years the research field of land surface data assimilation has undergone rapid developments, and it has provided a potential solution to the aforementioned problems. Land surface data assimilation offers a compromise between model and observations, and by minimization of their total errors it creates an analysis state which is superior to the model and observation alone. This thesis focuses on the implementation of a land surface data assimilation system, its applications and how to improve the separate elements that goes into such a framework. My ultimate goal is to improve the representation of soil moisture over northern high latitudes using land surface data assimilation.

In my three papers, I first show how soil moisture data assimilation can correct random errors in the precipitation fields used to drive the land surface model. A result which indicates

that a land surface model, driven by uncorrected precipitation, can have the same skill as a land surface model driven by bias-corrected precipitation. I show that passive microwave remote sensing can be utilized to monitor drought over regions of the world where this was thought to be impractical. I do this by creating a novel drought index based on passive microwave observations, and I validate the new index by comparing it with output from a land surface data assimilation system. Finally, I address knowledge gaps in the modelling of microwave emissions over northern high latitudes. In particular, I study the impact of neglecting multiple-scattering terms from vegetation in the radiative transfer models of microwave emission.

My three papers show that: (i) land surface data assimilation can improve surface soil moisture estimates at regional scales, (ii) passive microwave observations carries more information about the land surface over northern high latitudes than explored in the retrieval processing chain and (iii) including multiple-scattering terms in microwave radiative transfer models has the potential to increase the sensitivity for surface soil moisture below dense vegetation, and decrease biases between modelled and observed brightness temperature. In sum, my three papers lay the foundation for a land data assimilation system applicable to monitor the hydrological cycle over northern high latitudes.

# List of papers

I **Blyverket, J.**; Hamer, P.D.; Bertino, L.; Albergel, C.; Fairbairn, D.; Lahoz, W.A. An Evaluation of the EnKF vs. EnOI and the Assimilation of SMAP, SMOS and ESA CCI Soil Moisture Data over the Contiguous US. *Remote Sensing* 2019, 11. doi:10.3390/rs11050478.

II **Blyverket, J.**; Hamer, P.D.; Schneider, P.; Albergel, C.; Lahoz, W.A. Monitoring Soil Moisture Drought over Northern High Latitudes from Space. *Remote Sensing* 2019, 11(10). doi.org/10.3390/rs11101200

III **Blyverket, J.**; Hamer, P.D.; De Lannoy, G. Quantifying Higher Order Vegetation Scattering Effects in Passive Microwave Observations from SMAP over Northern Latitudes. *Manuscript in preparation*





# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Overview . . . . .	2
<b>2</b>	<b>Scientific Background</b>	<b>5</b>
2.1	Soil Moisture in the Earth System . . . . .	6
2.2	Soil Moisture Observations . . . . .	9
2.2.1	In Situ Soil Moisture Observations . . . . .	9
2.2.2	Microwave Remote Sensing of Soil Moisture . . . . .	9
2.2.3	Forward Modelling of Microwave Brightness Temperature . . . . .	10
2.2.4	Satellite Soil Moisture Retrieval . . . . .	12
2.3	Land Surface Modelling . . . . .	15
2.4	Land Surface Data Assimilation . . . . .	17
2.5	Monitoring of Hydrometeorological Extremes . . . . .	22
<b>3</b>	<b>Objectives</b>	<b>23</b>
<b>4</b>	<b>Summary of Papers</b>	<b>27</b>
<b>5</b>	<b>Discussion and Future Work</b>	<b>33</b>
5.1	Discussion . . . . .	34
5.2	Future Work . . . . .	37
5.2.1	Data Assimilation Developments . . . . .	37
5.2.2	Coupled Land-Atmosphere Data Assimilation . . . . .	38

**References** 41

**Scientific Papers** 57

# List of Figures

1.1	Land data assimilation schematic and an overview of my three papers . . . . .	3
2.1	Soil moisture control on 2m air temperature . . . . .	7
2.2	Emission pathways for the zeroth-order $\tau - \omega$ model . . . . .	11
2.3	Retrieval of soil moisture and vegetation optical depth (VOD) from SMAP using the Multi-Temporal Dual Channel Algorithm (MT-DCA) . . . . .	12
2.4	Illustration of the analysis step in a land data assimilation system. . . . .	18
2.5	Time-series of in situ, open loop and analysis over the Ford Dry Lake in situ station in the United States. . . . .	19
2.6	Rescaling of satellite observations to model climatology . . . . .	20
3.1	Monitoring drought from space . . . . .	24



# **Chapter 1**

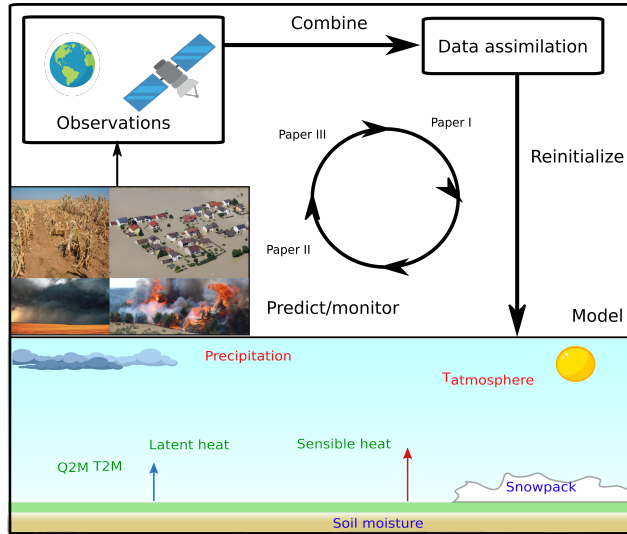
## **Introduction**

## 1.1 Overview

The land surface is a crucial component in the Earth system, as mass and energy fluxes to and from the land have an impact on the global water and energy cycles (Seneviratne et al., 2010; Trenberth et al., 2009, 2007). The atmosphere reacts rapidly and in a chaotic way to changes to the external forcing. While the land surface on the other hand, has a slower response. This inherent memory in land surface variables, such as snow, soil moisture and vegetation, can provide Earth system predictability at longer time-scales than the 10-day predictability limit in standard numerical weather prediction (Koster et al., 2003; Orsolini et al., 2013). For example, water stored as snow on land changes the land surface albedo, which in result can change large scale atmospheric circulation patterns (Henderson et al., 2018). Better representation of soil water content has been shown to increase the skill in the prediction of extreme events. As an example, more realistic land surface model physics was seen as one of the main factors for improving the predictability of the 2003 European heatwave (Weisheimer et al., 2011). Furthermore, extremes in the terrestrial water cycle, such as floods and droughts have large societal and economical impacts, which is exemplified by the billions of dollars spent annually by the European Union on drought mitigation and recovery (Gerber and Mirzabaev, 2017). Accurate forecasting and monitoring of the terrestrial water cycle components will help to minimize the negative impacts of an extreme event. For instance, early drought warning could allow farmers to sow their fields with crops that are more resilient against dry conditions.

The heterogeneity and complexity of the land surface makes it a difficult domain to both model and observe. The dawn of the satellite era has made it viable to monitor the water cycle and its different components from space (Balsamo et al., 2018; Sheffield et al., 2009). Satellites provide near-real-time observations at global scale, something in situ measurements never will do. Remote sensing sensors are, however, limited by inaccuracies in the instrumentation and representativeness errors of the observations. Representativeness errors are, e.g., mismatch in scale between the satellite observations and the model predictions. In addition, there are spatio-temporal gaps in the satellite observations, which are caused by the satellite orbit and/or the nature of the problem. It is, for example, not feasible to measure root-zone soil moisture because of the limited penetration depth of current microwave satellite platforms. Land surface

models allow us to close the spatio-temporal gaps in the observations, however as with all models, land surface models have their limitations. These limitations can be everything from errors in the parameterization of physical processes to errors in the atmospheric forcing.



**Figure 1.1:** Simple schematic of how land surface data assimilation works. Land surface models are combined with observations from satellite or in situ sensors using data assimilation, the new updated state is then applied as initial conditions for a new prediction. (Blue) example of prognostic variables that can be directly updated in offline land surface data assimilation, (green) diagnostic variables are indirectly updated through the update of the prognostics variables. The red atmospheric variables could potentially be updated if the data assimilation system is coupled with an atmospheric model. My three papers touch upon the implementation of a land surface data assimilation system (Paper I), the application of a land surface data assimilation system (Paper II) and remote sensing of soil moisture (Paper III).

Data assimilation (DA) offers a robust mathematical framework for combining information from land surface models and observations (Lahoz and De Lannoy, 2014; Lahoz et al., 2010; Reichle, 2008). As an example, given the vertical discretization of a land surface model and the (co-)variances in soil moisture content between the layers, DA will transfer information from the satellite derived surface soil moisture to the deeper layers of the model. In this way, it is possible to exploit the spatio-temporal coverage of the land surface model and at the same time correct model deficiencies with observational data. Numerous examples of successful land surface DA studies are available, applying a variety of DA techniques and observations. One of the most commonly used techniques in land DA is the ensemble Kalman



filter (EnKF) (Evensen, 1994). The EnKF updates prognostics variables on the fly and it is characterized by an ensemble of models states, which represents the background error covariance. Satellite-based microwave and optical observations are frequently used in land surface DA studies to update soil moisture (De Lannoy and Reichle, 2015; Lievens et al., 2017), snow (Aalstad et al., 2018; De Lannoy et al., 2009) and streamflow (López López et al., 2016).

There are nevertheless challenges related to land surface DA. For example, the skill of the analysis depends on how the model and observation errors are set, and how to represent the system errors in a realistic way is not always trivial. The background error covariances are particularly important, as they determine how to spread the observation information between model variables and in the horizontal and vertical direction.

I will in this thesis explain the concepts of land surface data assimilation, and how I combine information from land surface models and satellite observations to provide an improved estimate of the variables of interest. Figure 1.1 presents an overview of the different components of a land DA system and the main subjects of my three papers. Paper I focuses on data assimilation methods, Paper II focuses on the application of land DA and passive microwave remote sensing to study hydrometeorological extremes and Paper III focuses on remote sensing of soil moisture. I will start with a short introduction on the role of soil moisture in the Earth system in Section 2.1, I will then in Section 2.2 provide a background on how to measure soil moisture. In Section 2.3, I will give an introduction to land surface modelling, in Section 2.4 I will explain the mathematical framework that is used to combine model and observation data, and in Section 2.5 I will briefly explain the metrics I apply to monitor hydrometeorological extremes. After this general introduction, I will in Section 3 present the objectives and motivation for my three papers. In Section 4, I will summarize my papers and their main results. In Section 5, I will shortly discuss my results and potential future work. My three papers are attached at the end of the thesis.

## **Chapter 2**

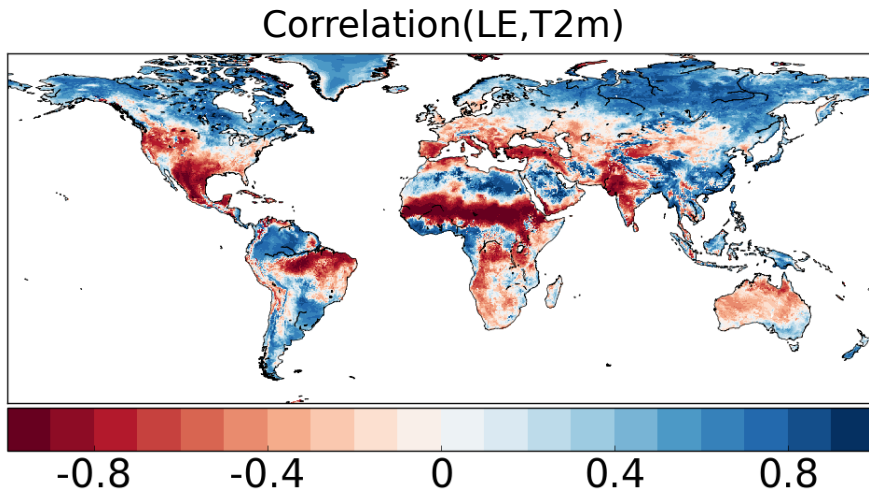
### **Scientific Background**

## 2.1 Soil Moisture in the Earth System

Soil moisture (SM) is here defined as the volumetric water content in the soil, with units  $\text{m}^3\text{m}^{-3}$ . SM is often separated into a surface and root-zone soil water content. Surface SM can be indirectly observed from satellite platforms, while models or other tools have to be used to give details about the water content in the deeper layers of the soil. In the Earth system, SM controls the partitioning of incoming solar radiation into outgoing latent and sensible heat fluxes (Seneviratne et al., 2010). Through this control on latent and sensible heat fluxes, SM influences the terrestrial energy cycle (Trenberth et al., 2009). A moist soil will cool the overlying air, because more of the incoming shortwave radiation is lost in the phase transition of water (liquid water to water vapour) compared to a drier soil, where the sensible heating will be larger. In this way, the SM influences the near surface air temperature. Figure 2.1 shows the correlation between latent heat flux (LE) and 2m air temperature (T2m), and it is based on 10 years of June, July and August monthly means, following the approach in (Jaeger et al., 2009; Seneviratne et al., 2006). Negative/positive correlation indicates strong/weak coupling between SM and air temperature. Areas with strong coupling are seen in the central United States, savanna regions of Brazil, Sahel, southern Africa, the Mediterranean basin, India and Australia. While weaker coupling is seen over, e.g., northern latitudes. The rationale behind this coupling metric is that areas with strong SM control on LE and T2m will show a negative correlation, because more of the incoming solar radiation goes to sensible heating. On the other hand, areas with positive correlations point to strong atmospheric control on LE, and the evaporation in these areas is energy limited.

In addition to the energy cycle, SM variations play a crucial role in the global carbon cycle, as it determines water availability for plants (D'Odorico et al., 2010) and influences long-term carbon uptake (Green et al., 2019). Extreme dry SM conditions, such as the 2003 heatwave, was found to decrease plant productivity and survival, eventually causing the European continent to be a source of  $\text{CO}_2$  (Ciais et al., 2005; Granier et al., 2007).

Several studies have investigated the relationship between SM and precipitation (Koster et al., 2004, 2003), and it is often hypothesized that dry SM anomalies lead to precipitation deficit and vice versa. This positive feedback mechanism is explained by the SM control on



**Figure 2.1:** Pearson correlation coefficient between latent heat flux (LE) and 2m air temperature (T2m) for June, July and August. Red/blue indicate strong/weak coupling. Negative correlation means that the latent heat flux is water limited, while positive correlation means that the latent heat flux is energy limited. The LE and T2m model data are extracted from the Copernicus Climate Change Service: <https://climate.copernicus.eu/>

partitioning of latent and sensible heat fluxes. Low precipitation can cause SM deficit, which in turn will increase the sensible heat flux and thereby the atmospheric temperature. Increased atmospheric temperature can change the planetary boundary layer and make the atmosphere less prone to precipitating deep convection, which again results in a reinforcement of the low precipitation (Berg et al., 2014; D’Odorico and Porporato, 2004). Precipitation sensitivity to early spring SM anomalies has been seen in General Circulation Models (GCMs) over, for example, the central United States (Koster et al., 2004, 2003). This large scale coupling is seen as a source for improved skill in sub-seasonal to seasonal prediction, in particular of near surface temperature (Koster et al., 2010, 2003). At shorter time-scales, Findell and Eltahir (2003) show that wet SM anomalies trigger deep convection more frequent than dry SM conditions, if the early morning atmosphere conditions are favourable.

Accurate initialization of SM is not only important for its potential control on long-term atmosphere predictability. Soil moisture deficiency causes agricultural droughts, which have large social and economic consequences (World Meteorological Organization, 2018). For the northern high latitudes, climate change is projected to on average increase wintertime precip-

itation (Greve et al., 2018), however, it is also projected to increase boreal summer aridity, with the likely effect of prolonging droughts and make them set in quicker (Samaniego et al., 2018; Trenberth et al., 2014). Thus, a way to monitor the terrestrial water cycle over northern areas is much needed. Through its combination of observation and model data, land surface data assimilation provides a comprehensive tool for monitoring and forecasting of SM extremes (Albergel et al., 2010, 2019, 2018; Luo and Wood, 2007).

I have highlighted the importance of accurate SM initialization, monitoring and forecasting for several aspects of the Earth system dynamics. In short, SM has a direct or indirect impact on: meteorology, hydrology, climate and biogeochemical cycles. I will in the next section briefly introduce how SM can be measured, both from in situ sensors and by using satellite remote sensing.

## 2.2 Soil Moisture Observations

### 2.2.1 In Situ Soil Moisture Observations

In situ SM observations are mostly based on either gravimetric and/or indirect observations (Robinson et al., 2008). Gravimetric observations are based on measurements of soil mass before and after drying (evaporating) all the liquid water in the soil. This is a time-consuming and destructive method, since the measurements cannot be repeated. The method is mostly used for calibration of indirect SM sensors. Indirect SM observations uses the dielectric properties of the soil and water to infer the water content. This method is therefore not destructive, however, as for the gravimetric method the spatial scale is limited to a point measurement. The limited spatial scale and cost of installation prohibits large scale monitoring of SM from in situ stations. Nevertheless, in situ SM observations are valuable for validation of satellite derived and modelled SM. The International Soil Moisture Network (ISMN) provides a database for in situ SM networks, and it has been extensively applied for validation of satellite derived and model derived SM (Dorigo et al., 2013, 2011).

### 2.2.2 Microwave Remote Sensing of Soil Moisture

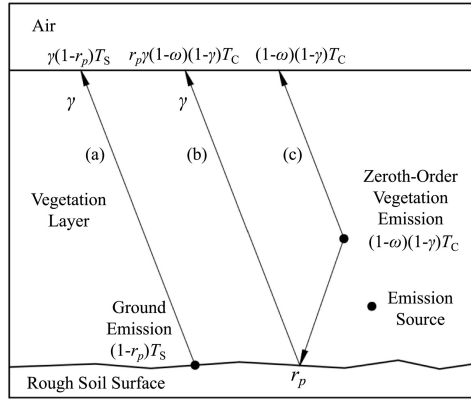
Satellite remote sensing of surface SM can be done using either active or passive instruments. Active instruments are based on radars or scatterometers, which illuminate the target and measure the amount of backscatter. Passive instruments, such as radiometers, are passively observing the electromagnetic radiation from the Earth's surface. Radiation in the microwave spectra, and in particular the X, C and L-band frequencies are sensitive to surface SM (Karthikeyan et al., 2017). An advantage with microwave observations in these bands is that there is little to no atmospheric attenuation, and the land surface can be measured through clouds and in the night-time. While active instruments can provide information at high spatial resolution they are also more sensitive to the surface roughness and vegetation water, and observations from passive instruments are often considered to be more sensitive to SM (Entekhabi et al., 2010). The rest of my thesis will therefore focus on passive microwave (PMW) observations.

Microwave observations are sensitive to SM because of the large difference in dielectric

constant between water ( $\sim 80$ ) and soil ( $\sim 4$ ) (Njoku and Entekhabi, 1996). The microwave emissivity of the land surface is therefore related to the water content in the soil. The observed microwave brightness temperature ( $T_b$ ) is proportional to the soil physical temperature, and the constant of proportionality is the soil microwave emissivity, which is influenced by SM variations (Karthikeyan et al., 2017; Wigneron et al., 2017). A wet soil has lower emissivity than a dry soil (high emissivity). A microwave radiometer is therefore effectively observing the dielectric constant of the soil - water medium. Soil water content is related to the dielectric constant via dielectric mixing models (Dobson et al., 1985; Wang and Schmugge, 1980). These models relate changes in dielectric constant of the soil - water medium to changes in volumetric water content. The dielectric mixing models are often based on empirical studies of in situ soil samples, thus assumptions have to be made about the mineral content of the soil within the satellite footprint. The upwelling microwave emission is sensitive to the overlying vegetation canopy and its vegetation water content (VWC). How this affects the microwave signal is related to the vegetation optical depth (VOD or  $\tau$ ) and the single-scattering albedo ( $\omega$ ). A zeroth-order radiative transfer model for microwave emission is the  $\tau - \omega$  model, and it describes how the microwave signal interacts with the vegetation (Mo et al., 1982). The  $\tau - \omega$  model has a central role in both SM retrieval algorithms and in the forward modelling of  $T_b$ . Passive microwave remote sensing of surface SM and forward modelling of  $T_b$  are both active research fields and has been for the last three to four decades (Karthikeyan et al., 2017; Njoku and Entekhabi, 1996; Schmugge, 2002).

### 2.2.3 Forward Modelling of Microwave Brightness Temperature

Forward modelling of the microwave brightness temperature is applied in both land DA applications and SM retrievals. Direct assimilation of  $T_b$  is desirable, as it in theory will introduce less errors in a land DA system (De Lannoy and Reichle, 2016). Less errors are introduced because the use of auxiliary data, that are potentially inconsistent with those used in the DA system, is avoided. For example, different inputs of land surface temperature, land cover and soil texture in the retrieval and the modelling system will ultimately lead to SM biases between the two datasets. In addition to biases caused by the use of different auxiliary data, there can



**Figure 2.2:** Emission pathways for the zeroth-order  $\tau - \omega$  model. (a) Ground emission attenuated by the vegetation, (b) vegetation emission reflected at the surface and (c) direct vegetation emission. Figure from [Feldman et al. \(2018\)](#).

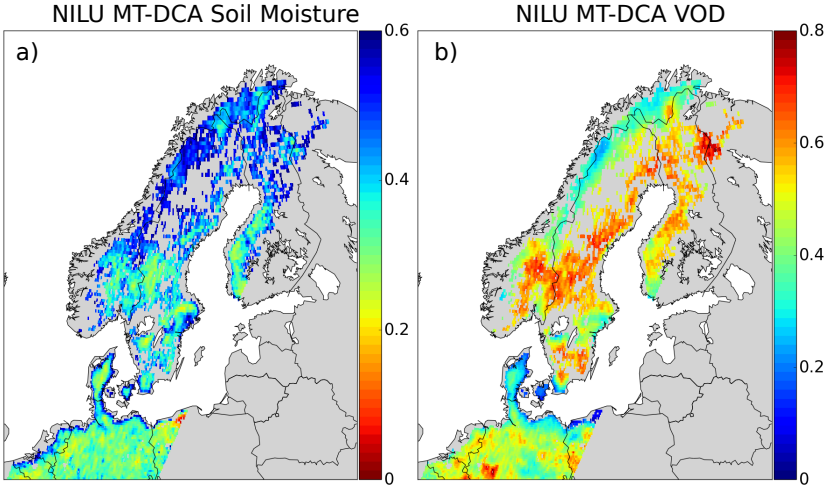
be biases because of limited knowledge of the problem at hand.

Even though PMW sensors are sensitive to surface SM, there are still limitations and challenges in the retrieval of SM from these platforms, see for example the review by [Karthikeyan et al. \(2017\)](#). Over northern areas the high open water fraction, complex topography, dense vegetation, soil freezing and snow cover make the SM retrieval difficult. The microwave emission from the soil is affected by the vegetation canopy, which can scatter and/or absorb the microwave signal ([Mo et al., 1982](#)). It is therefore crucial to take these vegetation effects into account when modelling the microwave emission from the land surface. In the original  $\tau - \omega$  model the total upwelling brightness temperature and how it is affected by the vegetation can be written as:

$$Tb_p^{0rh} = \underbrace{\gamma(1-r_p)T_s}_{\text{Canopy att.}} + \underbrace{\gamma r_p(1-\omega)(1-\gamma)T_c}_{\text{surf. att. refl. down. emiss.}} + \underbrace{(1-\omega)(1-\gamma)T_c}_{\text{upward veg. emission}}, \quad (2.1)$$

where  $p$  is the polarization (either horizontal H or vertical V). The transmissivity of the microwave emission through the canopy is given by  $\gamma$ ,  $r_p$  is the rough surface reflectivity and  $\omega$  is the single-scattering albedo of the vegetation layer.  $T_s$  is the soil temperature and  $T_c$  is the effective canopy temperature, see Fig. 2.2 for an illustration of the different terms. Failing to address the scattering and absorption within the canopy could lead to a too low/high modelled  $T_b$  signal, which consequently will lead to an overestimation/underestimation of SM and





**Figure 2.3:** (a) Soil moisture retrieval from the Soil Moisture Active Passive (SMAP) satellite using the 9 km enhanced product (Chan et al., 2018). Red/blue regions indicate low/high soil moisture. (b) Same as (a) but for vegetation optical depth (VOD), blue/red regions indicate low/high VOD. Retrieval from 3rd August 2017, descending overpass  $\sim$  06 UTC

thereby a wet/dry bias. Another source of bias is the water fraction within the satellite field of view, the microwave emissions from water bodies could cause non-negligible overestimation of SM and a resulting wet bias (Gouweleeuw et al., 2012; Loew, 2008). Research is ongoing to address the sensitivity of SM under a dense vegetation canopy (Feldman et al., 2018; Kurum, 2013; Kurum et al., 2012). In Paper III, I have added a first-order scattering term following Feldman et al. (2018), and the updated  $\tau - \omega$  model can be written as:

$$Tb_p^{Tot} = Tb_p^{0th} + Tb_p^{1st}. \quad (2.2)$$

Where the first-order scattering term  $Tb_p^{1st}$  accounts for multiple-scattering effects within the vegetation. This effectively means that the single-scattering term  $\omega$  in the  $\tau - \omega$  model is split into a zeroth ( $\omega_0$ ) and a first ( $\omega_1$ ) order scattering term.

## 2.2.4 Satellite Soil Moisture Retrieval

L-band radiometers are sensitive to SM in the top  $\sim$  5 cm of the soil, and for very dry conditions they are sensitive to SM in deeper layers (Kerr et al., 2012). Compared to the shorter

wavelengths of the X and C-band the longer wavelength of the L-band allows it to penetrate the vegetation canopy better. There are currently two satellite platforms dedicated to measure surface SM from space, and they both follow a sun-synchronous polar orbit with a revisit time of  $\sim 3$  days at the Equator. First, there is the Soil Moisture and Ocean Salinity (SMOS) satellite from the European Space Agency (ESA) (Kerr et al., 2012). SMOS was launched in November 2009, as an ESA Earth Explorer Mission. It is still functioning past its expected lifetime and therefore provides a ten year data record of PMW observations in the L-band. SMOS measures vertical and horizontal (V, H-polarization) brightness temperature at several incidence angles. The capability to measure  $T_b$  at several incidence angles has an advantage when retrieving more than one geophysical parameter from the observations. Although the observations made at the different angles have mutual information, there is still enough independent information to simultaneously retrieve, for example, SM and VOD (Fernandez-Moran et al., 2017).

Second, there is the NASA equivalent to SMOS, which is the Soil Moisture Active Passive (SMAP) satellite (Entekhabi et al., 2010). It was launched in 2015, and it carries both an L-band radiometer and a radar. However, the radar failed shortly after launch. SMAP measures  $T_b$  with both V and H-polarization, it is however limited to only measure  $T_b$  at one constant incidence angle. This constraints the number of parameters feasible to robustly retrieve using SMAP, as the information in the V and H-polarizations are correlated (Konings et al., 2015). Thus, for a single snapshot with two available polarizations the number of degrees of information is less than 2 (Konings et al., 2016). To retrieve both SM and VOD it is therefore necessary to utilize more than one overpass, and assume that the VOD is constant between two consecutive overpasses (separated by three days) (Konings et al., 2017, 2016). The Multi-Temporal Dual Channel Algorithm (MT-DCA) is a time-series approach which allows for robust retrieval of SM, VOD and vegetation scattering terms from the SMAP observations. An example is seen in Fig. 2.3, where I have simultaneously retrieved SM and VOD from the SMAP enhanced product (Chan et al., 2018). In Fig. 2.3, there are regions in eastern Norway, Sweden and Finland where the VOD is high (close to 0.8), which indicates that for this overpass there is a substantial attenuation by the vegetation over the Nordic countries. The

MT-DCA framework is advantageous as it permits for a retrieval of parameters without the need of auxiliary data, such as VOD, which would have been the case in parameter estimation through forward modelling of  $T_b$ .

It will be shown in Paper II and III that current state-of-the-art forward models and retrieval methods show deficiencies over northern areas. In general, the PMW remote sensing community suggest to study more in detail: (i) the selection of dielectric mixing models and their parameterizations, (ii) how to resolve the physical meaning of parameters, such as roughness constant and scattering albedo, and (iii) improved VOD retrieval to increase SM sensitivity in densely vegetated regions ([Karthikeyan et al., 2017](#)). In this thesis, I directly address issue number (ii) and (iii). I address point number (ii) by retrieving zeroth and first-order scattering terms over northern high latitudes. I implement the  $\tau - \omega$  model, and I modify it following [Feldman et al. \(2018\)](#), to allow for first-order scattering of the microwave signal in the vegetation. I address point number (iii) by applying the first-order  $\tau - \omega$  model within the MT-DCA framework.

## 2.3 Land Surface Modelling

Land surface models (LSMs) span a wide range of application areas, for example, they act as the lower boundary to the atmosphere over land in numerical weather prediction (NWP) and climate models (de Rosnay et al., 2014). Most of the LSMs solve for the water and energy budget, while equations for the carbon budget are included in some (Ek, 2018). Based on the computation of these prognostic (time-dependent) variables the LSMs also solve for the mass and energy fluxes to the atmosphere (diagnostic) variables. An advantage LSMs have over satellite observations and observations in general, is that they can provide output at high spatio-temporal resolution (Lahoz and De Lannoy, 2014).

The fundamentals of LSMs can be described by the flux budget equations for the energy:

$$R_n = H + LE + G, \quad (2.3)$$

and the water balance:

$$\Delta S = P - R - E. \quad (2.4)$$

Here  $R_n$  is the net radiation,  $H$  is the sensible heat,  $LE$  is the latent heat and  $G$  is the heat flux out or into the soil. For the water budget,  $\Delta S$  is the change in land surface water and includes soil moisture, ground water, snow and vegetation water. It is balanced by incoming precipitation  $P$ , runoff  $R$  and evapotranspiration  $E$ . Thus, SM has an indirect control on runoff and evapotranspiration. Furthermore, SM influences the heat fluxes to the atmosphere through the  $L$  in  $LE$ , where  $L$  is the latent heat of vaporization or sublimation. There is a vast number of available LSMs, and the basic underlying assumption in most of them are equations 2.3 and 2.4. How they solve the different physical processes is a different matter, and will not be covered in detail here, see for example Ek (2018) and references therein for a more detailed explanation of LSMs.

Since their infancy, LSMs have evolved and increased in complexity. Examples of developments are: (i) sub-grid variability, where the LSM is able to utilize auxiliary information at a higher resolution than the original grid (Chaney et al., 2016; Wood et al., 2011), (ii) transfer from simple bucket models to diffusion schemes (Boone et al., 2002; Noilhan

and Mahfouf, 1996), and (iii) evolving vegetation (Calvet et al., 1998). A few examples of LSMs are the Community Land Model (CLM) (Oleson et al., 2013), the Noah model (Ek, 2003), the Catchment LSM (Koster et al., 2000), the Variable Infiltration Capacity model (VIC) (Liang et al., 1994), the Hydrological-Tiled ECMWF Scheme for Surface Exchange over Land (H-TESSEL) (Balsamo et al., 2008) and the Interaction between Soil-Biosphere-Atmosphere diffusion (ISBA-DF) model in the SURFace EXternalisée (SURFEX) modelling framework (Masson et al., 2013).

In my work, I use the SURFEX LSM which is a state-of-the-art land surface model, and it has been extensively applied in other land DA studies (e.g., Albergel et al., 2019; Draper et al., 2009). It is also applied by the Nordic Met offices in the Applications of Research to Operations at Mesoscale (AROME) Meteorological Cooperation on Operational Numerical Weather Prediction (MetCoOp) system (Müller et al., 2017). Thus, knowledge from this thesis should be easy to transfer to their system for use in operational NWP over the Nordic region.

Despite the effort that goes into the development and calibration of LSMs, they still have limitations. Examples are parameterizations of physical processes, failure to represent sub-grid variability and the effect of errors in the atmospheric forcing (Lahoz and De Lannoy, 2014; Maggioni et al., 2011). In addition to this, model parameters are often related to auxiliary data, such as soil texture, land cover type and vegetation. Errors in these datasets will propagate into the LSM and introduce errors in the modelling framework.

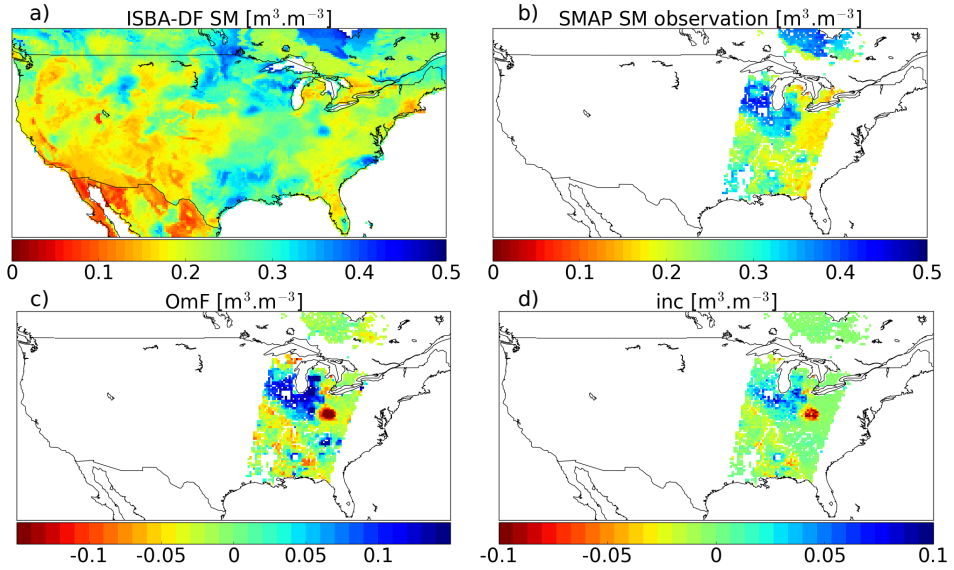
Studies show that combination of land surface models with satellite observations can correct for some of the aforementioned errors. For example, in Koster et al. (2018) the authors show that by tuning a model parameter using SMAP observations they are able to improve the model representation of water removal after a precipitation event (dry-down), when compared to in situ SM stations. Such simple model calibration methods based on increases in the correlation between observation and model data could work. However, this method neglects errors in both the model and observations. Thus, there should be a smarter way to combine the satellite observations with the land surface model, by for example, including their individual error characteristics and use this information to minimize the total error of the system. This method exists and it is referred to as data assimilation, and it will be detailed in the next chapter.

## 2.4 Land Surface Data Assimilation

Data assimilation (DA) is a mathematical framework for combining information from observations with numerical models. If the errors of the numerical model and the observations are known, DA can be used to minimize the total error of the system and estimate a new (analysis) value of the state variables. Data assimilation is applied daily in NWP, and it is identified as one of the reasons why today's 3-day forecast is as skillful as a 1-day forecast was less than 25 years ago (Lahoz et al., 2010). I will in this chapter focus on offline land surface DA of satellite derived surface SM. An offline system means that the LSM and atmospheric model are decoupled, thus there are no feedback between the two.

Land surface DA emerged 20 years ago, and there are two main reasons for the late blooming of this research field. First, most satellite missions were originally aimed at improving NWP, thus there was a lack of satellites dedicated to monitor the land surface (Houser et al., 2010). Second, the quality and complexity of land surface models started to improve, and they could now be applied at regional to global scale. One of the first studies to assimilate satellite derived surface SM observations at large scale were done by Reichle and Koster (2005). In their study, the authors assimilated (using the EnKF) global SM retrievals from the Scanning Multichannel Microwave Radiometer into the Catchment LSM. Other studies had until then focused on assimilation of screen level variables (Seuffert et al., 2003), field experiments (Crow and Wood, 2003; Reichle et al., 2001) and system development (Walker and Houser, 2001). After this, there have been numerous studies on land surface DA using sequential or variational methods, see for example the review by De Lannoy et al. (2018).

Data assimilation combines observations of the land surface and/or atmosphere to improve variables in the LSM (directly) or in the atmospheric model (indirectly). An example of this is provided in Fig. 2.4, where the model forecast (**a**) is combined with the satellite observation (**b**), to give the analyses increments seen in (**d**). The merging of model and observation data can either be done by the optimization of a cost function (variational methods) or by recursive update of the forecast state (sequential methods). Sequential methods in land DA are often based on the extended Kalman Filter (EKF) or the ensemble Kalman Filter (EnKF). The EKF and EnKF differ in the way they approximate nonlinearities in the land surface model. The



**Figure 2.4:** (a) Model forecast surface soil moisture from ISBA-DF, red/blue regions are low/high soil moisture content. (b) Same as (a) but for SMAP observed SM. (c) The Observation-minus-Forecast, blue/red regions indicate a wetter/drier SMAP observed SM. (d) Analyses increments from the EnKF, blue/red regions indicate where the analysis adds/removes water from the soil. Date: 15th June 2016, 12 UTC.

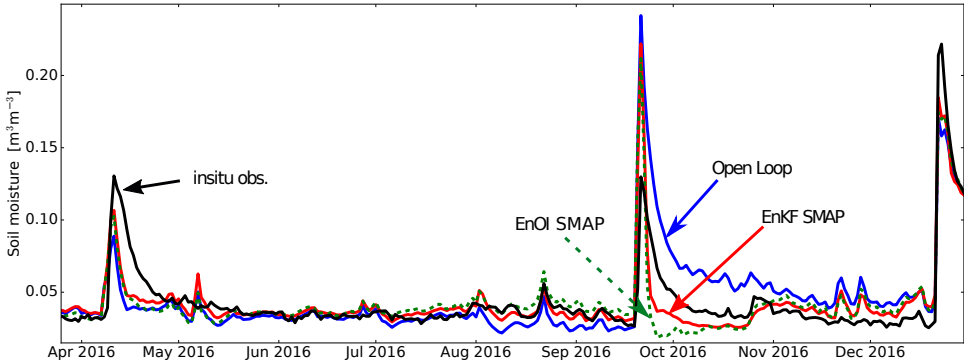
EKF uses a linearized propagation of the error covariance, while the EnKF nonlinearly propagates a finite sized ensemble forward in time. The EnKF is able to take into account errors in the atmospheric forcing, something which is difficult in the EKF. The EnKF is in addition easy to implement and it is flexible with regard to an increase in the number of state variables (Reichle et al., 2002). In my thesis, I focus on ensemble methods and to ease the conceptualization of the EnKF I write down its analysis equations:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_i^f), \quad (2.5)$$

where  $\mathbf{x}_i^a$  is the analysis for ensemble member  $i$ ;  $\mathbf{x}_i^f$  is the forecast state vector;  $\mathbf{y}$  is the vector of observations and  $\mathbf{H}$  linearly maps the state vector to observation space. The Kalman gain  $\mathbf{K}$  is given as:

$$\mathbf{K} = \mathbf{B}^f \mathbf{H}^T (\mathbf{H} \mathbf{B}^f \mathbf{H}^T + \mathbf{R})^{-1} \quad (2.6)$$

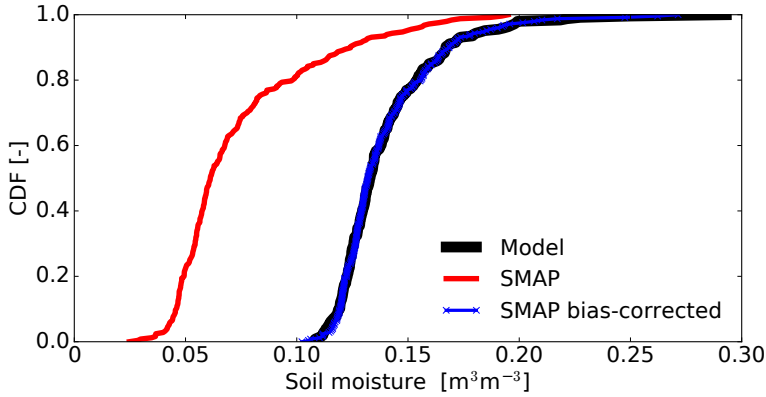
in which  $\mathbf{R}$  is the observation error-covariance matrix and  $\mathbf{B}^f$  is the forecast error-covariance



**Figure 2.5:** Time-series (1st April 2016 - 31st December 2016) of daily averaged surface soil moisture  $\text{m}^3\text{m}^{-3}$ , at the grid cell which contains the Ford Dry Lake in situ station in the United States. The in situ observation is the black line, land surface model open loop is blue line, EnOI of SMAP is green dashed line and the EnKF of SMAP is red. The mean simulated bias between the observation and model runs are removed for the ease of comparison.

matrix. The EnKF allows for a flow-dependent estimate of the background errors. This comes with an increase in the computational cost, since the model has to be run in an ensemble. In offline DA the LSM is forced externally by prescribed atmospheric forcing. Because there are no interactions between the atmosphere and the land surface in offline DA, and because the land surface model is not chaotic, perturbations in the land surface variables will be damped towards the original state quite fast. It is thus questionable whether ensemble methods (with flow dependency) are superior to variational or simpler Kalman filter methods in land surface DA of SM (Fairbairn et al., 2015; Reichle et al., 2002). A promising alternative to the EnKF is the ensemble Optimal Interpolation (EnOI) method (Counillon and Bertino, 2009; Evensen, 2003). The advantage of this method is that only one model run is needed. The analysis in the EnOI is computed solving an equation similar to Eq. 2.5, but now only for one model state (Evensen, 2003). In the EnOI, the  $\mathbf{B}^f$  is represented using either the ensemble open loop spread (which has errors of the day) or a climatological background error (no errors of the day and no flow dependency). The EnOI allows for a multi-variate analysis, just like the EnKF. In addition, computer time can be saved by only running the model once. The downside of the EnOI is that error estimates of the analysis are difficult to obtain, which is directly available through the analysis ensemble in the EnKF. Figure 2.5 shows an example of how the EnKF (red line) and EnOI (green dashed line) analyses differs from the open loop (blue line) at the





**Figure 2.6:** Cumulative Distribution Function (CDF) matching of satellite (red) data to the modelled derived CDF (black). The CDF of the bias-corrected observations is the blue crossed line. The estimated CDFs are based on a single grid cell over the United States for the period 1st April 2015 to 31st December 2016.

Ford Dry Lake in situ station in the United States. For this particular station, the Pearson correlation coefficient increased from 0.72 (open loop vs. in situ) to: 0.77 (EnOI vs. insitu) and 0.82 (EnKF vs. in situ) for the period 1st April 2015 to 31st December 2016.

The sequential filtering framework applied in this thesis updates the state variable at the current analysis time. An underlying assumption in sequential DA is that there are no systematic errors between the observations and the model. In other words, the system only corrects random errors between the observed and modelled quantity. As previously discussed, the use of different auxiliary data and lack of understanding of the PMW observations leads to biases between model derived and observed SM or Tb. The three most common ways to handle biases in land DA are: (i) a priori rescaling of the observations, i.e., anomaly assimilation, (ii) parameter tuning and (iii) online bias-correction (De Lannoy et al., 2018). In (i) the climatology (statistical moments) of the observations are matched with that of the LSM prior to the assimilation, which results in assimilation of anomalies. Method number (ii) is mostly used in assimilation of Tb by the tuning of parameters in the  $\tau - \omega$  model to remove long-term biases (De Lannoy et al., 2013). In method number (iii) the bias-correction is computed on the fly, either by assigning a forecast bias to the model or to the innovations (De Lannoy et al., 2013). In this work, I use method number (i), as it is the most frequently used method for assimilation of SM retrievals. To rescale the observations I use the cumulative distribution

function (CDF) matching method (Drusch et al., 2005; Reichle and Koster, 2004). The CDF-matching works by sorting the observed and modelled SM in increasing order, respectively, then I find the residual between the two sorted datasets and fit a polynomial to this function. By adding the fitted dataset to the original SM observations I am able to correct for the differences in the CDFs. An illustration of the CDF matching is shown in Fig. 3.1. The red line is the prior CDF of the SMAP observations (covering 1st April 2015 to 31st December 2016), while the black line is the model derived CDF. After bias-correction the new CDF of the observations follows the model derived CDF, thus most of the long-term differences have been removed.

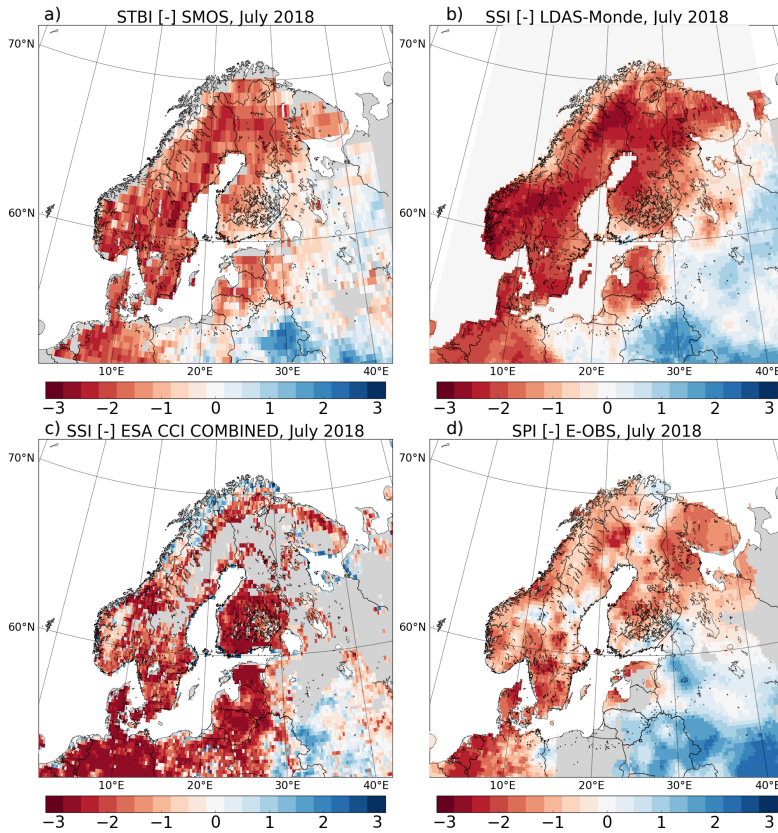
## 2.5 Monitoring of Hydrometeorological Extremes

Droughts and floods are hydrological/hydrometeorological extremes that have negative consequences for the human society and natural ecosystems (Wilhite, 2000). It is therefore important to monitor extreme events and the possible changes thereof, this is frequently done by computing standardized indices, which give information on how much (or how many standard deviations) an event deviates from normal conditions. To place an event in a historical perspective, it is therefore necessary to have a long time-series available of the variable of interest. The standardized indices are mostly based on precipitation, evapotranspiration, SM or a mix of several variables, depending on the phenomena they aim to describe (Svoboda, 2002). The underlying datasets used to compute these indices can be based on, for example, remote sensing observations (AghaKouchak et al., 2015; Sadri et al., 2018) or model data (Sheffield et al., 2004).

Given the data, there are two main methods to compute a standardized index. The first approach builds on the empirical distribution (non-parametric) of the data, while the second approach is based on fitting the data to a known distribution (parametric). In this work, I use a non-parametric approach for the standardization of the precipitation, while I use a parametric approach for the SM and Tb standardization. The non-parametric method is applied because the historical length of the precipitation data allows for a robust calculation of the empirical probability distribution function (PDF). For the SM and Tb standardizations, I am limited by the historical length of the data, thus a parametric approach where I fit the data to a known distribution is utilized. For SM this is the Beta-distribution following Sheffield et al. (2004), and for the Tb a normal distribution is assumed. The normality of the Tb data is tested using the Shapiro-Wilk test, and grid cells where this test fails are discarded from my analysis. A limitation of the SM and Tb indices from SMOS is the short record of the observations, i.e., 9 years. Ultimately, this means that the SM and Tb indices cannot describe an event in a historical perspective, but rather give a picture of the current conditions (Farahmand et al., 2015). The standardized precipitation index can be computed over a 69 year period, thus by comparing the SM and Tb indices to the precipitation index I am able to somewhat put them in a historical perspective.

# **Chapter 3**

## **Objectives**



**Figure 3.1:** (a) Drought index based on brightness temperature from SMOS. (b) Drought index based on surface soil moisture values from the Land Data Assimilation System Monde (LDAS-Monde). (c) Drought index based on the ESA CCI soil moisture product. (d) Drought index based the E-OBS in situ observed precipitation dataset. Red/blue values indicate drier/wetter than normal conditions.

My objectives in this thesis are twofold: (i) I want to implement a land DA system and (ii) I want to apply this system to study the hydrological cycle over northern latitudes. To address my objectives I split point (i) into two research questions:

- (a) What are the error characteristics of the model and observations in our land DA system?
- (b) How much improvement in skill does our land DA system provide when compared to independent reference data?

In my first point (i.a), I address the variability and biases of the LSM and satellite observations. I do this by comparing the spatio-temporal variability in the two SM datasets. In this way, I am

---

able to develop a methodology for rescaling satellite data prior to DA and address the biases between satellite and model data. I assess the quality of the land DA system and hence point (i.b) by using DA diagnostics and independent in situ SM data. DA diagnostics allows me to test the assumptions of a bias free system, which is crucial for the sequential filter performance. I also study the impact of land cover characteristics on the skill of our land DA analysis by dividing the SM stations into areas favourable and unfavourable to SM DA. Where favourable areas are defined as having less than  $5\text{kgm}^{-2}$  VWC, open water fraction less than 5% and little complex topography.

Observation characteristics and errors in PMW observations over northern latitudes are studied by implementing the  $\tau - \omega$  model in the MT-DCA framework. I am then able to calibrate model parameters and implement a potential model improvement (first-order scattering term), I do this to better understand the differences in modelled vs. observed Tb.

For my second objective (ii), I want to use a validated land DA system to study the hydrological cycle over northern high latitudes. I do this by assessing the benefit of my SM analyses for mapping the land surface in drought conditions. I create a drought index based on the SMOS PMW brightness temperature data, see Fig. 3.1 (a). I compare the SMOS index to drought metrics computed from the Land Data Assimilation System Monde (LDAS-Monde) (b) (Albergel et al., 2017), the ESA CCI SM product (c) (Dorigo et al., 2017, 2015) and in situ observed precipitation from the E-OBS dataset (d) (Haylock et al., 2008). My underlying objective is to create a drought metric which is able to provide improved spatio-temporal coverage over northern latitudes. A region where SM retrievals often are masked because of dense vegetation and high open water fraction.

To summarize, my main goals are to develop, validate and improve a land surface DA system and to use it to study the hydrological cycle over northern latitudes.



# **Chapter 4**

## **Summary of Papers**



In this chapter, I review the main results and conclusions from my three papers. I have in this thesis constructed a land surface DA system based on the EnKF and the SURFEX land surface modelling platform. Furthermore, I show how PMW remote sensing and the land DA system can be applied to monitor drought over northern high latitudes. Lessons learned from Paper I and II on the shortcomings of microwave emission modelling over northern latitudes lead me to study in more details the retrieval of geophysical parameters from PMW sensors. Paper III covers the study of multiple-scattering effects in PMW observations, and it also looks into the characteristics of PMW observations over northern latitudes.

**Paper I: An Evaluation of the EnKF vs. EnOI and the Assimilation of SMAP, SMOS and ESA CCI Soil Moisture Data over the Contiguous US.** Blyverket, J.; Hamer, P.D.; Bertino, L.; Albergel, C.; Fairbairn, D.; Lahoz, W.A. Published in: *Remote Sensing* 2019, 11. doi:10.3390/rs11050478.

Paper I addresses my objective number (i) and the two subsequent research questions (a) and (b) in Chapter 3. The paper focuses on the implementation and validation of the EnKF and the EnOI in SURFEX. Prior to the assimilation, the satellite observations are bias-corrected so that they share the same long-term climatology as the model. The sequential methods therefore correct random errors between my observations and model variables. I use the EnKF for which an ensemble of model runs are propagated forward in time by the SURFEX modelling platform. The spread of my model ensemble is taken as a direct measure of the model error, larger spread means larger uncertainty in the background forecast. I compare the EnKF with the computationally cheaper EnOI to investigate any flow dependencies of the background errors. Furthermore, I compare my land DA analysis with in situ SM stations. I am then able to assess the impact of the analyses on the representation of surface and root-zone SM. My key findings are:

- Using DA diagnostics I find that my land DA system is close to being bias-free.
- Comparison of the EnKF with the EnOI skill over in situ stations shows that the EnKF only has marginally higher skill than the EnOI.

- When I use land DA and uncorrected forcing my surface zone SM skill (compared to in situ stations) is close to the skill of an open loop using bias-corrected forcing.
- I find that analyses increments in the root-zone SM tend to decrease the skill of the model when compared to in situ SM data.
- I find that assimilation of the SMAP product outperforms the assimilation of the SMOS and ESA CCI data, respectively, in skill when compared to in situ stations over the CONUS.

The land DA system is implemented as a stand-alone-tool (outside the SURFEX source code), it can therefore easily be applied to other modelling frameworks besides SURFEX, and for longer/shorter assimilation windows. This makes the land DA framework flexible with respect to the assimilation of other prognostic variables such as snow, land surface temperature and leaf area index (LAI).

### **Paper II: Monitoring Soil Moisture Drought over Northern High Latitudes from Space.**

Blyverket, J.; Hamer, P.D.; Schneider, P.; Albergel, C.; Lahoz, W.A. Published in: *Remote Sensing* 2019, 11(10). doi.org/10.3390/rs11101200.

Paper II shows novel use of PMW observations and a land DA system over northern latitudes, and it directly addresses my objective number (ii) in Chapter 3. I do this by applying PMW remote sensing and a land surface DA system to monitor the 2018 Nordic drought. I create a novel drought index from the SMOS brightness temperature, which is named the Standardized Brightness Temperature Index (STBI). I validate the new drought index by comparison to a SM drought metric derived from the LDAS-Monde land DA system. The LDAS-Monde drought metric is based on SM analyses fields obtained from the assimilation of satellite derived surface SM and LAI. In Paper II the key findings are:

- The new drought index, the STBI, is able to capture the onset and extent of a drought over northern high latitudes when compared to other satellite data and the LDAS-Monde derived drought index.
- The STBI provides improved spatial coverage for drought monitoring over northern lat-

itudes when compared to SM retrievals, and in addition the same skill as the retrievals where there are data for both.

- The STBI does not capture the end of the 2018 Nordic drought.

While Paper II focuses on the STBI and its potential for drought monitoring, it should be emphasized that a drought index from a land DA system should in theory create a superior product compared to model or observations alone. Therefore, future studies will have to address the accuracy of a land DA system when compared to for instance, in situ observations of SM.

**Paper III: Quantifying Higher Order Vegetation Scattering Effects in Passive Microwave Observations from SMAP over Northern Latitudes.** Blyverket, J.; Hamer, P.D.; De Lannoy, G. *Manuscript in preparation.*

In Paper III, I study methods to improve the land DA system and I answer my research question number (i.a) in Chapter 3. To do this I implement the zeroth and first-order  $\tau - \omega$  radiative transfer model and the MT-DCA retrieval method. I apply this framework to retrieve SM, VOD and scattering albedo ( $\omega$ ) from SMAP observations over northern latitudes. In particular, I use the framework to study the effect of multiple-scattering terms ( $\omega_0$  and  $\omega_1$ ) in SMAP observations. I also compare two retrieval methods, the dual channel algorithm (DCA) and the MT-DCA. I do this to evaluate if a time-series approach is necessary for robust retrieval of more than two variables from SMAP. The key findings in Paper III are:

- My zeroth-order  $\tau - \omega$  retrieval of SM, VOD and  $\omega$  are close to values found in the literature.
- The DCA retrievals of SM and VOD are biased wet/high, respectively, compared to the MT-DCA method.
- Within the different land cover classes there is a substantial spread of  $\omega$  values.
- My results indicate that non-zero values of  $\omega_1$  comes from land covers with a high degree of regularity, i.e., mosaic and croplands, and not so much forested areas.

- Adding a first-order scattering term to the forward simulation of Tb reduced the model vs. observation bias to some extent.

Paper III allows me to study in detail PMW characteristics and errors over northern latitudes, and also how to potentially improve the forward modelling of Tb over this region. Improved Tb modelling would be of interest for Tb assimilation in the land DA system developed in Paper I.

In summary, my three papers cover aspects on: (i) the implementation of a land DA system, (ii) the application of a land DA system in conjunction with PMW observations to create a novel drought index, and (iii) calibration and development of a zeroth and first-order  $\tau - \omega$  model for use in a land DA system over northern latitudes.



## **Chapter 5**

### **Discussion and Future Work**

## 5.1 Discussion

In this chapter I will discuss my results, implications of these and potential future work. In Paper I, I use DA diagnostics to quantify properties of the analysis increments and any long-term observation-minus-forecast biases. I found that biases between the model and observations were effectively removed by the use of a CDF matching technique. It is still an open question whether the bias correction allows for maximum extraction of information from the satellite observations, or if some of the information is lost in the rescaling.

The land DA system presented in Paper I suggests that the EnKF only offers marginal improvement in skill compared to the EnOI. This result could imply that in land DA flow dependency of the background errors are negligible. Other studies have found the difference in skill between the EnKF and other sequential methods, such as the EKF, to be minor (Fairbairn et al., 2015; Reichle et al., 2002). I emphasize that the EnKF through its ensemble spread allows for a description of errors of the day, something which is not straightforward in the EKF and EnOI. Hence, an improved EnOI was obtained using the spread from the ensemble open loop as input to the background errors. When I take this into account, the EnOI computational cost is roughly half of the EnKF cost:  $N + 1$  ( $N$  is the ensemble size) model integrations compared to  $N + N$  integrations for the EnKF.

I find that a land DA system driven by uncorrected model forcing has skill close to that of an open loop driven by bias-corrected forcing. This means that in regions of the world without access to bias-corrected atmospheric forcing, a cheap alternative is to assimilate satellite derived SM. This result can be used to improve SM initialization in short-term NWP, and subsequently improved forecast of near surface variables through the SM-temperature coupling outlined in Chapter 2.1. Since land/vegetation processes have a longer memory than the weather time-scales, it should be expected that land DA has an impact on forecasting of land surface variables at longer time-scales, such as sub-seasonal to seasonal scales. However, results from the literature indicate that the sensitivity to land surface initial conditions in hydrological prediction is variable and largely depends on the region which is studied (Yuan et al., 2015).

While I find an improvement in surface SM skill for the DA analysis, no such improve-

ments are visible in the root-zone SM. Several studies have as well found that an improved surface zone skill not necessarily leads to an improvement in the root-zone (De Lannoy and Reichle, 2016; Kumar et al., 2008; Lievens et al., 2017). The explanation for this is twofold: (i) deficiencies in model physics, unrealistic weak/strong coupling strength between the soil layers, and (ii) spurious correlations between the soil layers in the DA system, leading to unphysical increments. In an ensemble DA system, the the explanation (i) can also be the main cause for (ii). For example, too strong model coupling between two soil layers could be reflected in unrealistic high covariances among the ensemble members.

Comparison of satellite skill in the DA analyses show that SMAP outperforms both the SMOS and the ESA CCI products. The ESA CCI data have undergone more extensive preprocessing and rescaling than the SMAP and SMOS datasets. This could lead to the introduction of additional errors when assimilating SM from the ESA CCI product. Furthermore, most of the assimilated ESA CCI data were from C-band radiometers and scatterometers. The C-band is not as sensitive to surface SM as the L-band, and this could lead to a smaller improvement in the analysis skill. Both SMAP and SMOS are L-band satellites designed to be sensitive to surface SM, so explanations for the difference in skill between the two are: (i) the SMAP radiometer has a lower instrument error than SMOS (1 K vs. 3 – 3.5 K) and it has radio frequency interference (RFI) mitigation strategies at instrumentation level, and (ii) differences in auxiliary data, such as land cover (Al-Yaari et al., 2017; De Lannoy et al., 2015).

There are room for improvements in several aspects of my land DA system. First, although an ensemble size of 12 is found to be sufficient for several applications in land DA (Yin et al., 2015), it may cause spurious correlations between the different soil layers. Especially in cases where the number of soil layers updated in the analysis is close to that of the ensemble size. However, as mentioned, other studies have also found it difficult to relate surface SM observations to improvement in simulated root-zone SM. Second, my local EnKF implementation is a 1D-filter, thus, I do not update surrounding (and often unobserved) grid cells, with the potential benefit this would have brought. In the literature, a 3D EnKF has been found to show better skill than a 1D-filter (Reichle and Koster, 2003), and could be implemented at the expense of higher memory demands. Third, I assimilate SM retrievals instead of direct PMW



observations, this could introduce errors from auxiliary datasets. However, studies show that this does not necessarily cause decrease in the analysis skill (De Lannoy and Reichle, 2016).

Paper II covers the application of PMW observations and a land DA system to monitor the 2018 Nordic drought. I show that satellite derived SM has limitations in spatio-temporal coverage over northern latitudes, and that this reduces its applicability for use in drought monitoring. A new index computed directly from the PMW Tb data shows increased spatio-temporal coverage and at the same time similar skill as the satellite derived SM, when compared to a land DA system. Although the new drought index was able to capture the drought onset and extent it failed to capture the drought end. This deficiency can be explained by looking at how the vegetation canopy changes the microwave emissions. Increased VWC leads to increased brightness temperature, which in turn is ambiguous with dry SM conditions. The PMW drought index should therefore be used as a supplementary tool and not as a replacement of other monitoring tools. One shortcoming of this study is that for an even more robust validation of the PMW and LDAS-Monde drought index, I could have included the validation against in situ SM stations. These stations are, however, limited to a few regions in Norway and Finland, and cannot provide a robust verification of the spatial skill of the two indices. Future studies should implement this drought metric over regions with denser SM networks and areas favourable to PMW observations, such as large areas in the CONUS.

The results from Paper II motivated my study of PMW observations in Paper III, and in particular the need to quantify the vegetation effects on the SM retrievals over northern latitudes. Comparison with the baseline SMAP SM product shows that my implementation of the zeroth-order  $\tau - \omega$  model within the DCA and MT-DCA frameworks have similar biases and correlations to those found in the literature (Konings et al., 2017). Retrievals of single-scattering albedo values per land cover class were also in line with other studies, which indicate that the zeroth-order model and the MT-DCA method works. My results indicate that non-zero values of  $\omega_1$  come from land covers with a high degree of regularity, and not so much from dense boreal forest. Thus, it is questionable to what degree the first-order scattering terms I find are solely a result of vegetation scattering. To further improve the retrieval of  $\omega_0$  and  $\omega_1$  parameters in the first-order model, it would be advantageous to refine the current brute

force model selection method, to an approach which takes observation and model errors into account.

Finally, adding a first-order scattering term to the forward simulation of Tb shows only a minor effect in reducing the model vs observation Tb bias. Especially, since the largest biases seems to be in areas with small (zero)  $\omega_1$  values, i.e., grasslands and open shrublands. Furthermore, I find that binning the  $\omega_1$  values into land cover classes creates too low values in many regions, and thereby a smaller first-order contribution to the total Tb signal. This results in a larger bias between the simulated and observed Tb than for the spatially distributed  $\omega_1$ . It would therefore be logical to use the  $\omega_1$  values retrieved per grid cell and not bin them per land cover class.

## 5.2 Future Work

### 5.2.1 Data Assimilation Developments

Future work could advantageously address the optimization of the land DA framework with respect to ensemble size, localization (1D vs. 3D EnKF), observation operator (Tb or SM assimilation) and variables to be assimilated. Assimilation of data from active microwave sensors, such as from Sentinel-1 should also be explored. The framework that has been developed in this thesis should be flexible enough for such upgrades.

In future work, it would be advantageous to also assimilate snow variables from microwave observations in the land DA system. For instance, snow cover has the potential to impact local and regional climate through its control on the surface energy balance. Driving factors for this control is the high albedo (and large albedo difference between snow and no snow conditions) and ground-insulating properties (Henderson et al., 2018). Snow DA is limited by the current microwave sensors sensitivity to snow variables and the use of optical sensors during winter-time (Aalstad et al., 2018).

A number of studies show that assimilation of satellite observed LAI has the potential to improve root-zone SM, and that it even has a bigger impact on SM than direct SM assimilation (Albergel et al., 2018, 2017). The increased spatio-temporal coverage of new LAI datasets

makes them applicable for land surface DA. However, one limitation is the use of optical remote sensing, which requires that the observed region must be cloud free, which is not too frequent at high latitudes. Therefore a path forward is to relate vegetation status to microwave observations.

## 5.2.2 Coupled Land-Atmosphere Data Assimilation

My land DA system provides a foundation for the development of coupled land-atmosphere DA, which is identified as an important future research area by the scientific community ([Carrassi et al., 2018](#); [Penny and Hamill, 2017](#)). Coupled land-atmosphere DA updates both the land and atmosphere domains, thus reduces potential imbalances at the land-atmosphere boundary. Coupled DA is broadly categorized into weakly (WCDA), and strongly CDA (SCDA). In WCDA, the coupling occurs through model physics in the forecast stage, using a coupled forecast model. The DA analysis is computed separately for the individual model domains, and the direct impact of the analysis is constrained to the domain in which the observations reside. The cross-domain information from the analysis step is propagated through the forward integration of the coupled model, which can be affected by post assimilation inconsistencies. In SCDA, coupled error covariances between the land-atmosphere variables are used to update simultaneously both domains. This approach allows observations in the land surface to instantaneously impact the atmosphere and vice versa. This has the benefit of producing a more consistent analysis of the fully coupled system.

Early attempts to couple DA systems have been reported, by e.g., ([Counillon et al., 2016](#); [Sluka et al., 2016](#)). However, the existing work has focused on CDA in the ocean-atmosphere system, and little work is done on the study of WCDA vs. SCDA in the land-atmosphere system.

Coupled DA has expected benefits in predictions across different time-scales, for example, NWP, sub-seasonal to seasonal, and multi-year and decadal (climate) predictions ([Penny and Hamill, 2017](#)). Other potential benefits include high-impact weather and regional prediction. An example of high-impact weather is the cold 2009/2010 European winter, linked to the negative phase of the North Atlantic Oscillation (NAO). Of particular interest is the coupling

---

between Eurasian snow cover and the NAO. Several studies have found that anomalies in the autumn Eurasian snow cover can influence the NAO in the following winter (see [Henderson et al. \(2018\)](#) and references therein). Thus, improved land-atmosphere initialization using CDA could benefit medium-range to monthly predictions of European winter conditions. The ability to improve the prediction of events such as the 2009/2010 winter in advance would have large socio-economic impacts.



# References

- Aalstad, K., Westermann, S., Schuler, T. V., Boike, J., and Bertino, L. (2018). Ensemble-based assimilation of fractional snow-covered area satellite retrievals to estimate the snow distribution at Arctic sites. *Cryosphere*, 12(1):247–270. [1.1](#), [5.2.1](#)
- AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., and Hain, C. R. (2015). Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*, 53(2):452–480. [2.5](#)
- Al-Yaari, A., Wigneron, J. P., Kerr, Y., Rodriguez-Fernandez, N., O’Neill, P. E., Jackson, T. J., De Lannoy, G. J., Al Bitar, A., Mialon, A., Richaume, P., Walker, J. P., Mahmoodi, A., and Yueh, S. (2017). Evaluating soil moisture retrievals from ESA’s SMOS and NASA’s SMAP brightness temperature datasets. *Remote Sensing of Environment*, 193:257–273. [5.1](#)
- Albergel, C., Calvet, J. C., Mahfouf, J. F., Rüdiger, C., Barbu, A. L., Lafont, S., Roujean, J. L., Walker, J. P., Crapeau, M., and Wigneron, J. P. (2010). Monitoring of water and carbon fluxes using a land data assimilation system: A case study for southwestern France. *Hydrology and Earth System Sciences*, 14(6):1109–1124. [2.1](#)
- Albergel, C., Dutra, E., Bonan, B., Zheng, Y., Munier, S., Balsamo, G., de Rosnay, P., Muñoz-Sabater, J., and Calvet, J.-C. (2019). Monitoring and Forecasting the Impact of the 2018 Summer Heatwave on Vegetation. *Remote Sensing*, 11(5). [2.1](#), [2.3](#)
- Albergel, C., Munier, S., Bocher, A., Bonan, B., Zheng, Y., Draper, C., Leroux, D. J., and 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 Sensing*, 10(10). [2.1](#), [5.2.1](#)

- Albergel, C., Munier, S., Jennifer Leroux, D., Dewaele, H., Fairbairn, D., Lavinia Barbu, A., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., Le Moigne, P., Decharme, B., Mahfouf, J. F., and 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. *Geoscientific Model Development*, 10(10):3889–3912. [3](#), [5.2.1](#)
- Balsamo, G., Agustì-Parareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Bousserez, N., Boussetta, S., Brown, A., Buizza, R., Buontempo, C., Chevallier, F., Choulga, M., Cloke, H., Cronin, M. F., Dahoui, M., Rosnay, P. D., Dirmeyer, P. A., Drusch, M., Dutra, E., Ek, M. B., Gentine, P., Hewitt, H., Keeley, S. P. E., Kerr, Y., Kumar, S., Lupu, C., Mahfouf, J.-F., McNorton, J., Mecklenburg, S., Mogensen, K., Muñoz-Sabater, J., Orth, R., Rabier, F., Reichle, R., Ruston, B., Pappenberger, F., Sandu, I., Seneviratne, S. I., Tietsche, S., Trigo, I. F., Uijlenhoet, R., Wedi, N., Woolway, R. I., and and Xubin Zeng (2018). Satellite and In Situ Observations for Advancing Global Earth Surface Modelling: A Review. *Remote Sensing*, 10. [1.1](#)
- Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K. (2008). A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System. *Journal of Hydrometeorology*, 10(3):623–643. [2.3](#)
- Berg, A., Lintner, B. R., Findell, K. L., Malyshev, S., Loikith, P. C., and Gentine, P. (2014). Impact of soil moisture-atmosphere interactions on surface temperature distribution. *Journal of Climate*, 27(21):7976–7993. [2.1](#)
- Boone, A., Masson, V., Meyers, T., and Noilhan, J. (2002). The Influence of the Inclusion of Soil Freezing on Simulations by a Soil–Vegetation–Atmosphere Transfer Scheme. *Journal of Applied Meteorology*, 39(9):1544–1569. [2.3](#)
- Calvet, J. C., Noilhan, J., Roujean, J. L., Bessemoulin, P., Cabelguenne, M., Olioso, A., and Wigneron, J. P. (1998). An interactive vegetation SVAT model tested against data from six contrasting sites. *Agricultural and Forest Meteorology*, 92(2):73–95. [2.3](#)

- Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G. (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5). [5.2.2](#)
- Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., Chaubell, J., Piepmeier, J., Yueh, S., Entekhabi, D., Colliander, A., Chen, F., Cosh, M. H., Caldwell, T., Walker, J., Berg, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., Uldall, F., Seyfried, M., Bosch, D., Starks, P., Holifield Collins, C., Prueger, J., van der Velde, R., Asanuma, J., Palecki, M., Small, E. E., Zreda, M., Calvet, J., Crow, W. T., and Kerr, Y. (2018). Development and assessment of the SMAP enhanced passive soil moisture product. *Remote Sensing of Environment*, 204:931–941. [2.3](#), [2.2.4](#)
- Chaney, N. W., Metcalfe, P., and Wood, E. F. (2016). HydroBlocks: a field-scale resolving land surface model for application over continental extents. *Hydrological Processes*, 30(20):3543–3559. [2.3](#)
- Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogee, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., De Noblet, N., Friend, A. D., Friedlingstein, P., Grünwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J. F., Sanz, M. J., Schulze, E. D., Vesala, T., and Valentini, R. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, 437(7058):529–533. [2.1](#)
- Counillon, F. and Bertino, L. (2009). Ensemble Optimal Interpolation: Multivariate properties in the Gulf of Mexico. *Tellus, Series A: Dynamic Meteorology and Oceanography*, 61(2):296–308. [2.4](#)
- Counillon, F., Keenlyside, N., Bethke, I., Wang, Y., Billeau, S., Shen, M. L., and Bentsen, M. (2016). Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the Norwegian Climate Prediction Model. *Tellus, Series A: Dynamic Meteorology and Oceanography*, 68(1). [5.2.2](#)



- Crow, W. T. and Wood, E. F. (2003). The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. *Advances in Water Resources*, 26(2):137–149. [2.4](#)
- De Lannoy, G. J., de Rosnay, P., and Reichle, R. H. (2018). Soil Moisture Data Assimilation. In *Handbook of Hydrometeorological Ensemble Forecasting*. Springer-Verlag Berlin Heidelberg. [2.4](#), [2.4](#)
- De Lannoy, G. J. and Reichle, R. H. (2015). Global Assimilation of Multiangle and Multipolarization SMOS Brightness Temperature Observations into the GEOS-5 Catchment Land Surface Model for Soil Moisture Estimation. *Journal of Hydrometeorology*, 17(2):669–691. [1.1](#)
- De Lannoy, G. J. and Reichle, R. H. (2016). Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. *Hydrology and Earth System Sciences*, 20(12):4895–4911. [2.2.3](#), [5.1](#)
- De Lannoy, G. J., Reichle, R. H., Houser, P. R., Arsenault, K. R., Verhoest, N. E. C., and Pauwels, V. R. N. (2009). Satellite-Scale Snow Water Equivalent Assimilation into a High-Resolution Land Surface Model. *Journal of Hydrometeorology*, 11(2):352–369. [1.1](#)
- De Lannoy, G. J., Reichle, R. H., and Pauwels, V. R. N. (2013). Global Calibration of the GEOS-5 L-Band Microwave Radiative Transfer Model over Nonfrozen Land Using SMOS Observations. *Journal of Hydrometeorology*, 14(3):765–785. [2.4](#)
- De Lannoy, G. J., Reichle, R. H., Peng, J., Kerr, Y., Castro, R., Kim, E. J., and Liu, Q. (2015). Converting between SMOS and SMAP Level-1 Brightness Temperature Observations over Nonfrozen Land. *IEEE Geoscience and Remote Sensing Letters*, 12(9):1908–1912. [5.1](#)
- de Rosnay, P., Balsamo, G., Albergel, C., Muñoz-Sabater, J., and Isaksen, L. (2014). Initialization of Land Surface Variables for Numerical Weather Prediction. *Surveys in Geophysics*, 35(3):607–621. [2.3](#)

- Dobson, M. C., Ulaby, F. T., Hallikainen, M. T., and El-Rayes, M. A. (1985). Microwave Dielectric Behavior of Wet Soil-Part II: Dielectric Mixing Models. *IEEE Transactions on Geoscience and Remote Sensing*, GE-23(1):35–46. [2.2.2](#)
- D’Odorico, P., Laio, F., Porporato, A., Ridolfi, L., Rinaldo, A., and Rodriguez-Iturbe, I. (2010). Ecohydrology of Terrestrial Ecosystems. *BioScience*, 60(11):898–907. [2.1](#)
- D’Odorico, P. and Porporato, A. (2004). Preferential states in soil moisture and climate dynamics. *Proceedings of the National Academy of Sciences*, 101(24):8848–8851. [2.1](#)
- Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, 203:185–215. [3](#)
- Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A., Zamojski, D., Cordes, C., Wagner, W., and Drusch, M. (2013). Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network. *Vadose Zone Journal*, 12(3):0. [2.2.1](#)
- Dorigo, W. A., Gruber, A., De Jeu, R. A., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R. (2015). Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sensing of Environment*, 162:380–395. [3](#)
- Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, A., and Jackson, T. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, 15(5):1675–1698. [2.2.1](#)
- Draper, C. S., Mahfouf, J. F., and Walker, J. P. (2009). An EKF assimilation of AMSR-E soil

- moisture into the ISBA land surface scheme. *Journal of Geophysical Research Atmospheres*, 114(20). [2.3](#)
- Drusch, M., Wood, E. F., and Gao, H. (2005). Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. *Geophysical Research Letters*, 32(15). [2.4](#)
- Ek, M. B. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research*, 108(D22):8851. [2.3](#)
- Ek, M. B. (2018). Land Surface Hydrological Models. In *Handbook of Hydrometeorological Ensemble Forecasting*, pages 1–42. Springer-Verlag Berlin Heidelberg. [2.3](#), [2.3](#)
- Entekhabi, D., Njoku, E. G., O’Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L., and Van Zyl, J. (2010). The Soil Moisture Active Passive (SMAP) Mission. *Proc. IEEE*, 98(5):704–716. [2.2.2](#), [2.2.4](#)
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics (part 1). *Journal of Geophysical Research*, 99(10):10,143–10,162. [1.1](#)
- Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367. [2.4](#)
- Fairbairn, D., Barbu, A. L., Mahfouf, J. F., Calvet, J. C., and Gelati, E. (2015). Comparing the ensemble and extended Kalman filters for in situ soil moisture assimilation with contrasting conditions. *Hydrology and Earth System Sciences*, 19(12):4811–4830. [2.4](#), [5.1](#)
- Farahmand, A., AghaKouchak, A., and Teixeira, J. (2015). A vantage from space can detect earlier drought onset: An approach using relative humidity. *Scientific Reports*, 5. [2.5](#)

- Feldman, A. F., Akbar, R., and Entekhabi, D. (2018). Characterization of higher-order scattering from vegetation with SMAP measurements. *Remote Sensing of Environment*, 219:324–338. [2.2](#), [2.2.3](#), [2.2.4](#)
- Fernandez-Moran, R., Al-Yaari, A., Mialon, A., Mahmoodi, A., Al Bitar, A., De Lannoy, G., Rodriguez-Fernandez, N., Lopez-Baeza, E., Kerr, Y., and Wigneron, J. P. (2017). SMOS-IC: An alternative SMOS soil moisture and vegetation optical depth product. *Remote Sensing*, 9(5). [2.2.4](#)
- Findell, K. L. and Eltahir, E. A. B. (2003). Atmospheric Controls on Soil Moisture–Boundary Layer Interactions. Part I: Framework Development. *Journal of Hydrometeorology*, 4(3):552–569. [2.1](#)
- Gerber, N. and Mirzabaev, A. (2017). Benefits of action and costs of inaction: Drought mitigation and preparedness – a literature review. Technical report, World Meteorological Organization (WMO) and Global Water Partnership (GWP). [1.1](#)
- Gouweleeuw, B. T., M. Van Dijk, A. I., Guerschman, J. P., Dyce, P., and Owe, M. (2012). Space-based passive microwave soil moisture retrievals and the correction for a dynamic open water fraction. *Hydrology and Earth System Sciences*, 16(6):1635–1645. [2.2.3](#)
- Granier, A., Reichstein, M., Bréda, N., Janssens, I. A., Falge, E., Ciais, P., Grünwald, T., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Facini, O., Grassi, G., Heinesch, B., Ilvesniemi, H., Keronen, P., Knohl, A., Köstner, B., Lagergren, F., Lindroth, A., Longdoz, B., Loustau, D., Mateus, J., Montagnani, L., Nys, C., Moors, E., Papale, D., Peiffer, M., Pilegaard, K., Pita, G., Pumpanen, J., Rambal, S., Rebmann, C., Rodrigues, A., Seufert, G., Tenhunen, J., Vesala, T., and Wang, Q. (2007). Evidence for soil water control on carbon and water dynamics in European forests during the extremely dry year: 2003. *Agricultural and Forest Meteorology*, 143(1-2):123–145. [2.1](#)
- Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., Lawrence, D. M., and Gentile, P. (2019). Large influence of soil moisture on long-term terrestrial carbon uptake. *Nature*, 565(7740):476–479. [2.1](#)

- Greve, P., Gudmundsson, L., and Seneviratne, S. I. (2018). Regional scaling of annual mean precipitation and water availability with global temperature change. *Earth System Dynamics*, 9(1):227–240. [2.1](#)
- Haylock, M. R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones, and New., M. (2008). A European daily high-resolution gridded dataset of surface temperature and precipitation. *Journal of Geophysical Research-Atmospheres*, 113:D20119. [3](#)
- Henderson, G. R., Peings, Y., Furtado, J. C., and Kushner, P. J. (2018). Snow–atmosphere coupling in the Northern Hemisphere. *Nature Climate Change*, 8(11):954–963. [1.1](#), [5.2.1](#), [5.2.2](#)
- Houser, P. R., De Lannoy, G. J., and Walker, J. P. (2010). Land surface data assimilation. In *Data Assimilation: Making Sense of Observations*, pages 549–597. Springer-Verlag Berlin Heidelberg. [2.4](#)
- Jaeger, E. B., Stöckli, R., and Seneviratne, S. I. (2009). Analysis of planetary boundary layer fluxes and land-atmosphere coupling in the regional climate model CLM. *Journal of Geophysical Research Atmospheres*, 114(17). [2.1](#)
- Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F. (2017). Four decades of microwave satellite soil moisture observations: Part 1. A review of retrieval algorithms. *Advances in Water Resources*, 109:106–120. [2.2.2](#), [2.2.3](#), [2.2.4](#)
- Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A., and Delwart, S. (2012). The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 50(5 PART 1):1384–1403. [2.2.4](#)
- Konings, A. G., McColl, K. A., Piles, M., and Entekhabi, D. (2015). How many parameters can be maximally estimated from a set of measurements? *IEEE Geoscience and Remote Sensing Letters*, 12(5):1081–1085. [2.2.4](#)
- Konings, A. G., Piles, M., Das, N., and Entekhabi, D. (2017). L-band vegetation optical depth

- and effective scattering albedo estimation from SMAP. *Remote Sensing of Environment*, 198:460–470. [2.2.4](#), [5.1](#)
- Konings, A. G., Piles, M., Rötzer, K., McColl, K. A., Chan, S. K., and Entekhabi, D. (2016). Vegetation optical depth and scattering albedo retrieval using time series of dual-polarized L-band radiometer observations. *Remote Sensing of Environment*, 172:178–189. [2.2.4](#)
- Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C. T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C. H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y. C., Taylor, C. M., Verseghy, D., Vasic, R., Xue, Y., and Yamada, T. (2004). Regions of strong coupling between soil moisture and precipitation. *Science*, 305(5687):1138–1140. [2.1](#)
- Koster, R. D., Liu, Q., Mahanama, S. P. P., and Reichle, R. H. (2018). Improved Hydrological Simulation Using SMAP Data: Relative Impacts of Model Calibration and Data Assimilation. *Journal of Hydrometeorology*, pages JHM–D–17–0228.1. [2.3](#)
- Koster, R. D., Mahanama, S. P., Yamada, T. J., Balsamo, G., Berg, A. A., Boissierie, M., Dirmeyer, P. A., Doblas-Reyes, F. J., Drewitt, G., Gordon, C. T., Guo, Z., Jeong, J. H., Lawrence, D. M., Lee, W. S., Li, Z., Luo, L., Malyshev, S., Merryfield, W. J., Seneviratne, S. I., Stanelle, T., Van Den Hurk, B. J., Vitart, F., and Wood, E. F. (2010). Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters*, 37(2). [2.1](#)
- Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., and Kumar, P. (2000). A catchment-based approach to modeling land surface processes in a general circulation model 1. Model structure. *Journal of Geophysical Research Atmospheres*, 105(D20):24809–24822. [2.3](#)
- Koster, R. D., Suarez, M. J., Higgins, R. W., and Van den Dool, H. M. (2003). Observational evidence that soil moisture variations affect precipitation. *Geophysical Research Letters*, 30(5). [1.1](#), [2.1](#)
- Kumar, S. V., Reichle, R. H., Peters-Lidard, C. D., Koster, R. D., Zhan, X., Crow, W. T., Eylander, J. B., and Houser, P. R. (2008). A land surface data assimilation framework using

- the land information system: Description and applications. *Advances in Water Resources*, 31(11):1419–1432. [5.1](#)
- Kurum, M. (2013). Quantifying scattering albedo in microwave emission of vegetated terrain. *Remote Sensing of Environment*, 129:66–74. [2.2.3](#)
- Kurum, M., O'Neill, P. E., Lang, R. H., Joseph, A. T., Cosh, M. H., and Jackson, T. J. (2012). Effective tree scattering and opacity at L-band. *Remote Sensing of Environment*, 118:1–9. [2.2.3](#)
- Lahoz, W. A. and De Lannoy, G. J. (2014). Closing the Gaps in Our Knowledge of the Hydrological Cycle over Land: Conceptual Problems. *Surveys in Geophysics*, 35(3):623–660. [1.1](#), [2.3](#), [2.3](#)
- Lahoz, W. A., Khattatov, B., and Ménard, R. (2010). *Data assimilation: Making sense of observations*. Springer-Verlag Berlin Heidelberg. [1.1](#), [2.4](#)
- Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J. (1994). A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs. *J. Geophys. Res.*, 99(D7):14415–14428. [2.3](#)
- Lievens, H., Reichle, R. H., Liu, Q., De Lannoy, G. J., Dunbar, R. S., Kim, S. B., Das, N. N., Cosh, M., Walker, J. P., and Wagner, W. (2017). Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. *Geophysical Research Letters*, 44(12):6145–6153. [1.1](#), [5.1](#)
- Loew, A. (2008). Impact of surface heterogeneity on surface soil moisture retrievals from passive microwave data at the regional scale: The Upper Danube case. *Remote Sensing of Environment*, 112(1):231–248. [2.2.3](#)
- López López, P., Wanders, N., Schellekens, J., Renzullo, L. J., Sutanudjaja, E. H., and Bierkens, M. F. (2016). Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. *Hydrology and Earth System Sciences*, 20(7):3059–3076. [1.1](#)

- Luo, L. and Wood, E. F. (2007). Monitoring and predicting the 2007 U.S. drought. *Geophysical Research Letters*, 34(22). [2.1](#)
- Maggioni, V., Reichle, R. H., and Anagnostou, E. N. (2011). The Effect of Satellite Rainfall Error Modeling on Soil Moisture Prediction Uncertainty. *Journal of Hydrometeorology*, 12(3):413–428. [2.3](#)
- Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J. C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essauouini, K., Gibelin, A. L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J. F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A. (2013). The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geoscientific Model Development*, 6(4):929–960. [2.3](#)
- Mo, T., Choudhury, B. J., Schmugge, T. J., Wang, J. R., and Jackson, T. J. (1982). A model for microwave emission from vegetation-covered fields. *Journal of Geophysical Research*, 87(C13):11229–11237. [2.2.2](#), [2.2.3](#)
- Müller, M., Homleid, M., Ivarsson, K.-I., Køltzow, M. A. Ø., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., Dahlgren, P., Kristiansen, J., Randriamampianina, R., Ridal, M., and Vignes, O. (2017). AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction Model. *Weather and Forecasting*, 32(2):609–627. [2.3](#)
- Njoku, E. G. and Entekhabi, D. (1996). Passive microwave remote sensing of soil moisture. *Journal of Hydrology*, 184(1-2):101–129. [2.2.2](#)
- Noilhan, J. and Mahfouf, J. (1996). The ISBA land surface parameterization scheme. *Global and Planetary Change*, 13:145–159. [2.3](#)
- Oleson, K. W., Lawrence, D. M., Gordon, B., Flanner, M. G., Kluzek, E., Peter, J., Levis, S.,



- Swenson, S. C., Thornton, E., and Feddema, J. (2013). Technical description of version 4.0 of the Community Land Model (CLM). Technical Report NCAR July. [2.3](#)
- Orsolini, Y. J., Senan, R., Balsamo, G., Doblas-Reyes, F. J., Vitart, F., Weisheimer, A., Carrasco, A., and Benestad, R. E. (2013). Impact of snow initialization on sub-seasonal forecasts. *Climate Dynamics*, 41(7-8):1969–1982. [1.1](#)
- Penny, S. G. and Hamill, T. M. (2017). Coupled data assimilation for integrated earth system analysis and prediction. In *Bulletin of the American Meteorological Society*, volume 98, pages ES169–ES172. [5.2.2](#)
- Reichle, R. H. (2008). Data assimilation methods in the Earth sciences. *Advances in Water Resources*, 31(11):1411–1418. [1.1](#)
- Reichle, R. H. and Koster, R. D. (2003). Assessing the Impact of Horizontal Error Correlations in Background Fields on Soil Moisture Estimation. *Journal of Hydrometeorology*, 4(6):1229–1242. [5.1](#)
- Reichle, R. H. and Koster, R. D. (2004). Bias reduction in short records of satellite soil moisture. *Geophysical Research Letters*, 31(19). [2.4](#)
- Reichle, R. H. and Koster, R. D. (2005). Global assimilation of satellite surface soil moisture retrievals into the NASA catchment land surface model. *Geophysical Research Letters*, 32(2):1–4. [2.4](#)
- Reichle, R. H., McLaughlin, D. B., and Entekhabi, D. (2001). Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications. *IEEE Transactions on Geoscience and Remote Sensing*, 39(8):1708–1718. [2.4](#)
- Reichle, R. H., Walker, J. P., Koster, R. D., and Houser, P. R. (2002). Extended versus Ensemble Kalman Filtering for Land Data Assimilation. *Journal of Hydrometeorology*, 3(6):728–740. [2.4](#), [2.4](#), [5.1](#)
- Robinson, D. A., Campbell, C. S., Hopmans, J. W., Hornbuckle, B. K., Jones, S. B., Knight, R., Ogden, F., Selker, J., and Wendroth, O. (2008). Soil Moisture Measurement for Eco-

- logical and Hydrological Watershed-Scale Observatories: A Review. *Vadose Zone Journal*, 7(1):358. [2.2.1](#)
- Sadri, S., Wood, E. F., and Pan, M. (2018). A SMAP-Based Drought Monitoring Index for the United States. *Hydrology and Earth System Sciences Discussions*, pages 1–19. [2.5](#)
- Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E. F., and Marx, A. (2018). Anthropogenic warming exacerbates European soil moisture droughts. *Nature Climate Change*, 8(5):421–426. [2.1](#)
- Schmugge, T. (2002). Remote Sensing of Surface Soil Moisture. *Journal of Applied Meteorology*, 17(10):1549–1557. [2.2.2](#)
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J. (2010). Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3-4):125–161. [1.1](#), [2.1](#)
- Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C. (2006). Land-atmosphere coupling and climate change in Europe. *Nature*, 443(7108):205–209. [2.1](#)
- Seuffert, G., Wilker, H., Viterbo, P., Mahfouf, J.-F., Drusch, M., and Calvet, J.-C. (2003). Soil moisture analysis combining screen-level parameters and microwave brightness temperature: A test with field data. *Geophysical Research Letters*, 30(10). [2.4](#)
- Sheffield, J., Ferguson, C. R., Troy, T. J., Wood, E. F., and McCabe, M. F. (2009). Closing the terrestrial water budget from satellite remote sensing. *Geophysical Research Letters*, 36(7). [1.1](#)
- Sheffield, J., Goteti, G., Wen, F., and Wood, E. F. (2004). A simulated soil moisture based drought analysis for the United States. *Journal of Geophysical Research D: Atmospheres*, 109(24):1–19. [2.5](#)
- Sluka, T. C., Penny, S. G., Kalnay, E., and Miyoshi, T. (2016). Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophysical Research Letters*, 43(2):752–759. [5.2.2](#)

- Svoboda, M. (2002). The Drought Monitor. *Bulletin of the American Meteorological Society*, (April):1181–1190. [2.5](#)
- Trenberth, K. E., Dai, A., Van Der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., and Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, 4(1):17–22. [2.1](#)
- Trenberth, K. E., Fasullo, J. T., and Kiehl, J. (2009). Earth’s global energy budget. *Bulletin of the American Meteorological Society*, 90(3):311–323. [1.1](#), [2.1](#)
- Trenberth, K. E., Smith, L., Qian, T., Dai, A., and Fasullo, J. (2007). Estimates of the Global Water Budget and Its Annual Cycle Using Observational and Model Data. *Journal of Hydrometeorology*, 8(4):758–769. [1.1](#)
- Walker, J. P. and Houser, P. R. (2001). A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. *Journal of Geophysical Research Atmospheres*, 106(D11):11761–11774. [2.4](#)
- Wang, J. R. and Schmugge, T. J. (1980). An empirical model for the complex dielectric permittivity of soils as a function of water content. *IEEE Transactions on Geoscience and Remote Sensing*, GE-18(4):288–295. [2.2.2](#)
- Weisheimer, A., Doblas-Reyes, F. J., Jung, T., and Palmer, T. N. (2011). On the predictability of the extreme summer 2003 over Europe. *Geophysical Research Letters*, 38(5). [1.1](#)
- Wigneron, J. P., Jackson, T. J., O’Neill, P., De Lannoy, G., de Rosnay, P., Walker, J. P., Ferrazzoli, P., Mironov, V., Bircher, S., Grant, J. P., Kurum, M., Schwank, M., Munoz-Sabater, J., Das, N., Royer, A., Al-Yaari, A., Al Bitar, A., Fernandez-Moran, R., Lawrence, H., Mialon, A., Parrens, M., Richaume, P., Delwart, S., and Kerr, Y. (2017). Modelling the passive microwave signature from land surfaces: A review of recent results and application to the L-band SMOS & SMAP soil moisture retrieval algorithms. *Remote Sensing of Environment*, 192:238–262. [2.2.2](#)
- Wilhite, D. A. (2000). Chapter 1 Drought as a Natural Hazard. In *Drought: A Global Assessment*, pages 147–162. Routledge Publishers, London. [2.5](#)

- Wood, E. F., Roundy, J. K., Troy, T. J., Beek, L. P. H. V., Bierkens, M. F. P., Blyth, E., Roo, A. D., Döll, P., Ek, M., Famiglietti, J., Gochis, D., Giesen, N. V. D., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Lidard, C. P., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P. (2011). Hyperresolution global land surface modeling : Meeting a grand challenge for monitoring Earth ' s terrestrial water. *Water Resources Research*, 47:1–10. [2.3](#)
- World Meteorological Organization (2018). The State of the Global Climate in 2018. Technical report. <https://public.wmo.int/en/our-mandate/climate/wmo-statement-state-of-global-climate>. Accessed: 19th May 2019. [2.1](#)
- Yin, J., Zhan, X., Zheng, Y., Hain, C. R., Liu, J., and Fang, L. (2015). Optimal ensemble size of ensemble Kalman filter in sequential soil moisture data assimilation. *Geophysical Research Letters*, 42(16):6710–6715. [5.1](#)
- Yuan, X., Wood, E. F., and Ma, Z. (2015). A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development. *Wiley Interdisciplinary Reviews: Water*, 2(5):523–536. [5.1](#)



# **Scientific Papers**



# Paper I

## **An Evaluation of the EnKF vs. EnOI and the Assimilation of SMAP, SMOS and ESA CCI Soil Moisture Data over the Contiguous US**

**Blyverket, J.;** Hamer, P.D.; Bertino, L.; Albergel, C.; Fairbairn, D.; Lahoz, W.A.



*Remote Sensing*, **11**, (2019), doi:10.3390/rs11050478.





Article

# An Evaluation of the EnKF vs. EnOI and the Assimilation of SMAP, SMOS and ESA CCI Soil Moisture Data over the Contiguous US

Jostein Blyverket <sup>1,2,\*</sup> , Paul D. Hamer <sup>1</sup>, Laurent Bertino <sup>3</sup>, Clément Albergel <sup>4</sup> ,  
David Fairbairn <sup>5</sup> and William A. Lahoz <sup>1</sup>

<sup>1</sup> NILU—Norwegian Institute for Air Research, INBY, Instituttveien 18, 2007 Kjeller, Norway; pdh@nilu.no (P.D.H.); wal@nilu.no (W.A.L.)

<sup>2</sup> Geophysical Institute, University of Bergen UiB, Allegaten 70, 5020 Bergen, Norway

<sup>3</sup> Nansen Environmental and Remote Sensing Center, 5020 Bergen, Norway; laurent.bertino@nersc.no

<sup>4</sup> CNRM—Université de Toulouse, Météo-France, CNRS, 31057 Toulouse, France; clement.albergel@meteo.fr

<sup>5</sup> European Centre for Medium-range Weather Forecasts (ECMWF), Reading RG2 9AX, UK; David.Fairbairn@ecmwf.int

\* Correspondence: jb@nilu.no

Received: 21 January 2019; Accepted: 20 February 2019; Published: 26 February 2019



**Abstract:** A number of studies have shown that assimilation of satellite derived soil moisture using the ensemble Kalman Filter (EnKF) can improve soil moisture estimates, particularly for the surface zone. However, the EnKF is computationally expensive since an ensemble of model integrations have to be propagated forward in time. Here, assimilating satellite soil moisture data from the Soil Moisture Active Passive (SMAP) mission, we compare the EnKF with the computationally cheaper ensemble Optimal Interpolation (EnOI) method over the contiguous United States (CONUS). The background error-covariance in the EnOI is sampled in two ways: (i) by using the stochastic spread from an ensemble open-loop run, and (ii) sampling from the model spinup climatology. Our results indicate that the EnKF is only marginally superior to one version of the EnOI. Furthermore, the assimilation of SMAP data using the EnKF and EnOI is found to improve the surface zone correlation with in situ observations at a 95% significance level. The EnKF assimilation of SMAP data is also found to improve root-zone correlation with independent in situ data at the same significance level; however this improvement is dependent on which in situ network we are validating against. We evaluate how the quality of the atmospheric forcing affects the analysis results by prescribing the land surface data assimilation system with either observation corrected or model derived precipitation. Surface zone correlation skill increases for the analysis using both the corrected and model derived precipitation, but only the latter shows an improvement at the 95% significance level. The study also suggests that assimilation of satellite derived surface soil moisture using the EnOI can correct random errors in the atmospheric forcing and give an analysed surface soil moisture close to that of an open-loop run using observation derived precipitation. Importantly, this shows that estimates of soil moisture could be improved using a combination of assimilating SMAP using the computationally cheap EnOI while using model derived precipitation as forcing. Finally, we assimilate three different Level-2 satellite derived soil moisture products from the European Space Agency Climate Change Initiative (ESA CCI), SMAP and SMOS (Soil Moisture and Ocean Salinity) using the EnOI, and then compare the relative performance of the three resulting analyses against in situ soil moisture observations. In this comparison, we find that all three analyses offer improvements over an open-loop run when comparing to in situ observations. The assimilation of SMAP data is found to perform marginally better than the assimilation of SMOS data, while assimilation of the ESA CCI data shows the smallest improvement of the three analysis products.

**Keywords:** land data assimilation; EnKF; EnOI; SMAP; SMOS; ESA CCI for soil moisture

---

## 1. Introduction

Accurate knowledge of surface and root-zone soil moisture is important since it constrains the energy and water exchanges across the land–atmosphere interface. Soil moisture (SM) controls the surface infiltration vs. runoff process, water available for the biosphere through plant transpiration, and partitioning of incoming energy into sensible and latent heat fluxes [1,2]. Several studies have investigated if initialization of SM could help the skill of seasonal hydrological predictions [3–5], and also in which parts of the world such seasonal predictions are sensitive to the initial hydrological conditions [6]. In such regions, knowledge about the initial conditions of SM can provide a skilful prediction of subsequent SM conditions using historical atmospheric forcing or seasonal climate forecasts. One way of improving the SM state is through data assimilation (DA); this approach combines information from satellite observations with model data, using a mathematical framework, see, for example, [7–14].

Microwave satellite instruments allow us to monitor and map SM at the global scale, and provide a wealth of information that can be used to inform and improve Earth surface modelling [15]. However, satellite remote sensing of SM has limitations because of spatial and temporal gaps in the resulting data product. For example, current satellite instruments used to derive SM are operating in the X, C and L-band, which are mostly sensitive to SM in the upper soil layer (0–5 cm), and not the root-zone layer [16]. Spatial gaps are also found in the horizontal direction because not all parts of the globe will be covered every day, while temporal gaps are related to the revisit time of the satellites.

By merging land surface modelling and satellite observations using DA, we can create a product which combines the spatial and temporal coverage of a land surface model (LSM) with satellite derived surface SM (SSM). The LSM controls the partitioning of rainfall into runoff, evapotranspiration, drainage and soil moisture. Satellite observations can compensate for model parameterizations and precipitation events not described in the atmospheric forcing. For example, in Koster et al. [17], the authors found that, by assimilating and utilizing SM derived from the Soil Moisture Active Passive (SMAP) satellite, they were able to improve the model representation of dry-down events. By knowing the errors of the model and the observations, DA has the potential to create a superior product over and above either model or satellite SM alone. For instance, this approach has been taken in the development of the SMAP Level-4 product, where modelled SM is converted to brightness temperature and optimized towards the observed brightness temperature measured by SMAP [18].

One of the most applied and successful methods in land DA is the ensemble Kalman Filter (EnKF) [19]. This method solves the Kalman filter equations, while sampling the background error-covariance from an ensemble of model trajectories. The EnKF allows for flow dependent background errors and also gives the uncertainty of the analysis in the form of the ensemble spread. The input errors are set by perturbing the atmospheric forcing and/or the state variables and/or model parameters. How to correctly specify these errors on a larger domain is not trivial, and it is expected that an improved spatial description of model and observation errors will benefit land surface DA [12]. One of the major caveats with the EnKF is the high computational cost of the forward integration of an ensemble of model states [20] and references therein. This often make the EnKF unattractive for operational applications over larger domains.

To keep the computational cost at a minimum while still retaining some of the benefits from the EnKF, we suggest to use the ensemble Optimal Interpolation (EnOI) method [21,22]. The EnOI solves the same Kalman Filter equations as the EnKF; however, it relies on a prescribed background error-covariance. Land DA systems using the EnKF often do the ensemble forward integration twice, once for normalizing the satellite SM data to the model and the second integration for doing the analysis. Here, we suggest two methods for describing the background error-covariance in an EnOI

system: (i) using the ensemble open-loop forecast, which can be obtained from the model ensemble used to normalize the satellite SM data, and (ii) using a climatological background error-covariance from the model spinup. The major benefit with method (i) is that we only need to do the ensemble forward integration once, while, for method (ii), no ensemble integration is needed at all. More details about these methods are presented in Section 3.4.

The relative improvements of SM estimates between the analysis in a land DA system and the open-loop (i.e., no DA) run are affected by the quality of the atmospheric forcing used to drive the land surface model [23,24], whereby the usage of observation corrected or model derived precipitation tends to result in lower or higher impact of the land DA, respectively. A model derived precipitation dataset is expected to e.g., have more missed precipitation events, or represent precipitation events that never occurred, and these random errors in the precipitation field can be compensated for by land DA of satellite derived surface SM. In our study, we utilize two different atmospheric forcing datasets, (i) the Modern Era Retrospective-analysis for Research and Applications version 2 data (MERRA-2) from the Global Modeling and Assimilation Office (GMAO) [25] and (ii) data from the North American Land Data Assimilation System version 2 (NLDAS-2) [26]. By applying the observation corrected precipitation product from the NLDAS-2 dataset, and the model derived precipitation from the MERRA-2 dataset, we are able to evaluate how these atmospheric forcing datasets affect our land DA analysis results.

The main goal of this paper is to assess the EnOI versus the EnKF for soil moisture DA. The skill of the two filters is evaluated using DA diagnostics and in situ SM measurements. The DA diagnostics checks that the assumptions underlying the Kalman Filter equations are met, e.g., the whiteness of the observation-minus-forecast (O-F) residuals, and that the expected model and observation errors are close to the actual errors [27]. We do the model experiments over the contiguous United States (CONUS) because this region incorporates a wide range of biomes with different climatic conditions, while at the same time it has numerous in situ SM stations to use for validation. In situ observations act as an independent reference for testing the skill of the analysis vs the open-loop. Furthermore, the comparison to in situ SM observations is utilized to: (i) evaluate the impact of the quality of the atmospheric forcing on the land DA results, (ii) assess whether a land DA analysis driven by model derived precipitation can have comparable skill to an open-loop run using observation corrected precipitation, and (iii) assess the skill of the land DA analysis using three different SM satellite products from the European Space Agency Climate Change Initiative (ESA CCI), Soil Moisture and Ocean Salinity (SMOS) and SMAP, respectively. Section 2 presents the different satellite SM products, in situ observations and model auxiliary data. In Section 3, we describe the bias correction, modelling system and data assimilation algorithms. Section 4 describes the experiments and validation metrics. Section 5 presents the results.

## 2. Data

### 2.1. Satellite Derived SSM

#### 2.1.1. ESA CCI ACTIVE and PASSIVE SSM Product

The ESA CCI for soil moisture project has been ongoing since 2011; the goal of the project is to create a global, consistent SSM product using active and passive satellite sensors [28–31]. As of 18 January 2019, the dataset spans from 1978 until 30 June 2018. The ESA CCI SSM product is divided into three different datasets. First, the ACTIVE product, which consists of daily SSM values from active sensors, using backscatter as a measure of SSM. Second, the PASSIVE product, which uses passively observed brightness temperature as a measure of SSM. Finally, the COMBINED product, fusing the ACTIVE and PASSIVE product into one, for improved spatial and temporal coverage.

The ACTIVE SSM retrieval product is based on the change detection method developed at the Vienna University of Technology (TU Wien) [32–34]. The change detection method is derived for C-band scatterometers. Surface soil moisture is retrieved directly using scatterometer measurements

without the need for an iterative adjustment process. It is assumed that the relationship between the SSM and the backscatter coefficient  $\sigma^0$  is linear. For the time-period 31 March 2015–31 December 2016 used in our study, the ACTIVE SSM observations stem from the METOP-A and B satellites (ASCAT-A and ASCAT-B).

The PASSIVE SSM dataset is derived using the land surface brightness temperature (Tb) in the microwave range, which is related to the soil dielectric constant and hence the SSM content. The conversion from observed Tb to SSM values is handled by the Land Parameter Retrieval Model (LPRM) [35]. The LPRM is a radiative transfer model that simultaneously retrieves vegetation density, SSM and surface temperature.

Several studies have assimilated the different ESA CCI products, e.g., [13,36]. Draper and Reichle [37] found that the largest improvements in SSM were seen when assimilating both the active and passive data, compared to assimilating active/passive data only. For this reason, we simultaneously assimilate the separate ACTIVE and PASSIVE products, which are given in the range of the change detection algorithm and the land parameter retrieval model, respectively. The different overpass times of the satellites in the ACTIVE and PASSIVE products allow for assimilation of both products, without any further pre-processing. In this work, we use the ESA CCI SSM v04.2, which became public in January 2018. The ESA CCI ACTIVE and PASSIVE products are delivered on a regular longitude/latitude grid. We remap the ESA CCI observations to the model grid (with a resolution of 25 km) using a nearest neighbour approach. This ensures that the observed SSM values are not smoothed values from an interpolation scheme. Further pre-processing of the satellite data is explained in Section 3.1. More information about the ESA CCI SSM product, including quality assessment and range of uses is available in Refs. [29,30].

### 2.1.2. SMAP Level-2 SSM

The Soil Moisture Active Passive (SMAP) satellite measures passive microwave emissions from the Earth's surface [38]. It was launched in 2015 by NASA, carrying a passive instrument (L-band radiometer) and an active instrument (L-band radar). The active instrument failed shortly after launch, hence we only use data from the radiometer. The satellite derived SSM is an average over the whole footprint area, which is around 40 km for L-band radiometers. The SMAP Level-2 (L2) SSM product is extracted from the National Snow and Ice Data Center; we use the v5 R16010 product covering the period 31 March 2015–31 December 2016. We only use SMAP-L2 data where the SSM uncertainty is less than  $0.1 \text{ m}^3 \text{ m}^{-3}$ , the SSM is in a realistic range ( $0.0\text{--}0.6 \text{ m}^3 \text{ m}^{-3}$ ), and where the LSM observation layer (layer-2) has a temperature above  $2^\circ \text{C}$  [12]. The SMAP-L2 swath data are gridded to the LSM grid using a nearest neighbour approach.

### 2.1.3. SMOS Level-2 SSM

The Soil Moisture and Ocean Salinity (SMOS) satellite, was launched in 2010 by ESA, and similarly to SMAP, SMOS uses a radiometer instrument operating in the L-band [39,40]. In contrast to the SMAP satellite, which has a constant observation incidence angle, SMOS measures brightness temperature for a range of observation incidence angles. This angular information can be used to separate soil and vegetation signals over land [41]. The SMOS-L2 SSM data are extracted from the SMUDP2 v650 reprocessed product obtained from the ESA SMOS dissemination service. To match the experiment period, we use SMOS-L2 data spanning from 31 March 2015–31 December 2016, which is the overlapping time-period between the SMAP-L2, SMOS-L2 and the ESA CCI v4.2 products. The quality controls applied before bias correction consist of (i) removing observations outside a realistic SM range ( $0.0\text{--}0.6 \text{ m}^3 \text{ m}^{-3}$ ), (ii) choosing SM values with uncertainties less than  $0.1 \text{ m}^3 \text{ m}^{-3}$ , (iii) choosing SM values for instances where the radio frequency interference (RFI) probability is less than 0.3, and (iv) choosing SM values where the retrieval flag is not raised for snow and ice. Finally, the observations are discarded if the soil temperature in the LSM is below  $2^\circ \text{C}$ . The SMOS-L2 swath values are gridded to the model grid using the same approach as for the SMAP-L2 SSM product.

## 2.2. In Situ Data

The high number of in situ stations over the United States allows us to compare the land DA results to independent reference data. This allows for a more robust comparison and validation of the analysis results. We use data from the International Soil Moisture Network (ISMN) [42,43], and, more specifically, data from the Soil Climate Analysis Network (SCAN) [44] and the U.S. Climate Reference Network (USCRN) [45,46]. A promising way to validate the land surface data assimilation system is by using cosmic-ray neutron measurements from the COsmic-ray Soil Moisture Observing System (COSMOS) [47,48]. This network is not applied in our study because we wanted to facilitate the comparison of average skill metrics across the SCAN and USCRN networks as done in other land DA studies. There are several limitations when validating satellite and model SM data using sparse in situ networks. These limitations include: (i) spatial mismatch (representativeness error) between what the station (i.e., a point) measures and what the model and satellite measure (i.e., domain averaged SM), and (ii) in situ station measurements have instrumental errors. The in situ data are pre-processed following Reichle et al. [49], this means that: (i) we only use data flagged as good in the ISMN data product, (ii) all values are discarded if the measured soil temperature at that level is below 2 °C, and (iii) we only use SM values in the range between 0 and  $\theta_{\text{saturation}}$  for that station. The  $\theta_{\text{saturation}}$  information for a given station is obtained from the ISMN auxiliary data product. After this quality control of the data, we create a temporal average using the preceding and succeeding hourly values of a 3 h window. Surface zone soil moisture (sfzsm) and root-zone soil moisture (rzsm) are calculated for the individual in situ stations following the approach in Ref. [12], and excluding the in situ SM sensor at 0.5 m depth following [50].

## 2.3. Atmospheric Forcing

We use atmospheric forcing from NLDAS-2 [26]; this dataset provides hourly input of precipitation (PRECIP), air temperature (Tair), specific humidity (SH), longwave downward (LW) and shortwave downward radiation (SW). The NLDAS-2 dataset has observation corrected precipitation and also bias corrections for incoming SW radiation; this means that the forcing dataset error will be small compared to the errors associated with, e.g., a forecast product. The NLDAS-2 forcing is upscaled from the native 12.5 km grid to the model 25 km grid using bilinear interpolation. The NLDAS-2 dataset is a high-resolution dataset focused on North America, which allows it to ingest high-quality precipitation observations at a fine scale.

The second atmospheric forcing dataset is from the Modern Era Retrospective-analysis for Research and Applications version 2 (MERRA-2) [25]. The MERRA-2 product is a state-of-the-art global atmospheric reanalysis with a resolution of  $0.5^\circ \times 0.625^\circ$ . From MERRA-2, we utilize hourly input of PRECIP, Tair, SH, LW and SW. To evaluate the impact of the atmospheric forcing on the land DA analysis, we apply the MERRA-2 PRECIP without any observation corrections. The MERRA-2 product was downloaded from the Goddard Earth Sciences Data Information Services Center (GES DISC). We use bilinear interpolation to remap the MERRA-2 product from the native grid to the LSM grid.

## 3. Methods

### 3.1. Bias Correction of the ESA CCI ACTIVE and PASSIVE, SMAP-L2 and SMOS-L2 SSM Products

In sequential DA algorithms, it is assumed that there is no bias between the model and the observations. Sequential DA methods only correct random errors, hence systematic errors in the land DA system need to be removed. In land DA, this is often done by matching the statistical moments of the observations to that of the model, e.g., [51,52]. The ACTIVE and PASSIVE products are delivered as daily products, and they are a composite of different satellite overpasses at a given date. To avoid additional bias between the model and observations, the timing of the forecast (model) and observation (satellite retrieval) need to coincide as much as possible. Following De Lannoy and Reichle [12], we

use a temporal gap of  $\pm 1.5$  h centred at 00, 03..., and 21 UTC; satellite overpasses within this time window are mapped to the centre of the time window.

The range of the ACTIVE product is given in percent, 0–100%. The ACTIVE data are converted to volumetric SM using the model minimum and maximum SM for a given grid-cell. For the conversion to volumetric SM, we only use model timestamps where we also have observational data. The PASSIVE product is given in volumetric SM and does not need any additional unit conversion.

To handle the bias between the model and observations, we rescale the observations using cumulative distribution function matching (CDF-matching) as in [52,53]. We note that different bias correction methods could have an impact on the DA results as seen in, e.g., [54,55]. In this study, we assume that the biases between the satellite derived SSM and modelled SSM are close to being stationary [12], which means that we can use a lumped CDF-approach. If the biases are seasonally varying, this assumption could lead to remaining biases and dampening of short-term variability in the satellite derived SSM, as pointed out by, e.g., [9].

On the technical side, the CDF-matching is performed for each individual grid-cell independently, and requires that a given grid-cell has more than 200 observations over the whole time-period. Otherwise, all the values at that particular grid-cell are discarded. The CDF-matching works by ranking the observations and model values for an individual grid-cell at a specific time e.g., 0900 UTC. By taking the difference between the modelled and observed ranked datasets and then fitting a 5th order polynomial to these points, we can find the new observed value by adding this difference to the old observed value. The CDF-matched observations will then have a CDF which is matched with the CDF of the model for that grid-cell and time of the day.

### 3.2. Modelling System

In our study, we use the SURFEX v.8.0 (SURFace EXternalisée [56]) modelling framework. Within the SURFEX land surface modelling platform, we use the Interaction between Soil Biosphere and Atmosphere diffusion scheme (ISBA-DF) [57–59]. Vertical transport of water is solved using the mixed form of Richard's equations while the soil temperature is solved using the one-dimensional Fourier law. The ISBA model also includes soil freezing [60] and an explicit snow scheme [61]. Fourteen vertical layers are used over a depth of 12 m depending on grid-cell characteristics. The depth of the different layers are the same as in [13]. Layer one is a skin layer, we use layer two (between 0.01 and 0.04 m) as the model equivalent of the observation layer. The time-step for the LSM is set to 30 min, output is saved at 3 h intervals. In this work, we only use one sub-grid patch where the parameters for the mass and energy balance are aggregated from the different land covers within the grid-cell. More information about the model can be found in Decharme et al. [59].

Land cover is extracted from ECOCLIMAP, a global database at 1 km resolution for land surface parameters [62]. For nature grid-cells in the model, surface parameters are computed using the fraction of 12 vegetation types from the 1 km resolution land cover map. The leaf area index (LAI) is derived from the ECOCLIMAP database and given as a climatology for each calendar month. The clay and sand fractions are extracted from the Harmonized World Soil Database (HWSD) [63]. In SURFEX, several soil parameters are derived from the clay and sand fraction using formulas from Noilhan and Lacarrere [64]. The hydraulic conductivity and soil water potential are related to the liquid soil water content through the Clapp and Hornberger relations. Grid-cell orography is computed in SURFEX by aggregating the GTOPO 1 km resolution elevation dataset to the 25 km model grid.

### 3.3. Data Assimilation Using the Ensemble Kalman Filter

There are several different versions of the EnKF; in this work, we use the Ensemble Square Root Filter (ESRF) from Sakov and Oke [65]. This algorithm is applied both in the EnKF and the EnOI methods, for simplicity we will refer to the ESRF as the EnKF in the rest of this paper. The EnKF interface with SURFEX in this work has been developed at the Norwegian Institute for Air Research (NILU) and it is different from the EnKF developed at the National Centre for Meteorological

Research (CRNM) [20,66]. The state vector  $\mathbf{x}$  consists of the top four model layers, i.e.,  $\theta_1$  (0.001–0.01 m),  $\theta_2$  (0.01–0.04 m),  $\theta_3$  (0.04–0.10 m) and  $\theta_4$  (0.10–0.20 m) for all of the  $N$  ensemble members. Owing to spurious correlations found between the model equivalent observation layer (layer-2) and layers 5 to 14, we choose to only update the top four model layers. Hence, the root-zone is not explicitly updated in the DA analysis. An observation vector  $\mathbf{y}$  holds the satellite observation ( $\theta_{\text{obs}}$ ); the EnKF combines an ensemble of model states ( $\mathbf{x}_N$ ) with the observations  $\mathbf{y}$ . For both the EnKF and EnOI, we apply a three-hourly assimilation window, and this is to ensure that the model forecast SSM is as close as possible to the satellite derived SSM. The EnKF analysis equation is written as:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_i^f),$$

where  $\mathbf{x}_i^a$  is the analysis for ensemble member  $i$ ;  $\mathbf{x}_i^f$  is the forecast state;  $\mathbf{y}$  is the vector of observations and  $\mathbf{H}$  linearly maps the state vector to observation space. The Kalman gain  $\mathbf{K}$  is given as:

$$\mathbf{K} = \mathbf{B}^f \mathbf{H}^T (\mathbf{H} \mathbf{B}^f \mathbf{H}^T + \mathbf{R})^{-1}$$

in which  $\mathbf{R}$  is the observation error-covariance matrix and  $\mathbf{B}^f$  is the forecast error-covariance matrix. The analysis error-covariance is updated according to:

$$\mathbf{B}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}^f,$$

for the EnKF, where  $\mathbf{I}$  is the identity matrix. In the EnKF, the model error-covariance matrix  $\mathbf{B}^f$  is stored and manipulated implicitly via an ensemble  $\mathbf{x}_N$  of  $N$  model states, and is given by:

$$\mathbf{B}^f = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T = \frac{1}{N-1} \mathbf{A}\mathbf{A}^T,$$

where  $\mathbf{x}$  is the ensemble mean defined as

$$\mathbf{x} = \frac{1}{N} \sum_i^N \mathbf{x}_i.$$

In the above equations,  $\mathbf{A}$  represents the deviations from the ensemble mean. The EnKF used here is currently a 1-D scheme with independent soil columns in the  $x$ - and  $y$ -directions. The observation operator is set to:  $\mathbf{H} = (0, 1, 0, 0)$ , so that it only selects the cross-correlations between the observation layer and layers 1,3 and 4. We apply 12 ensemble members for the EnKF and the EnOI, this choice is based on: (i) the state vector is relatively small (four variables), (ii) other studies have found that this number is sufficient and there is little to gain when increasing it further [37,67,68], and (iii) as an effort to minimize the computational cost.

### 3.4. Ensemble Optimal Interpolation (EnOI)

In the EnKF equations above, the background error-covariance is modified in the analysis step, and this allows for a flow dependent background error. However, this comes with an added computational cost as we need  $N$  model integrations ( $N$  being the number of ensembles). Here we argue that an ensemble Optimal Interpolation (or statistical interpolation) method could perform well in a land surface DA system [21,22]. The main advantage of the EnOI is the low computational cost since we only need one ensemble member. This is especially useful for operational systems and for reanalyses where we run an ensemble open-loop for the period of interest to normalize the satellite data (as done here). Then, we can use the background error-covariance from the ensemble open-loop for “free” in the EnOI. It is also worth noting that, when moving towards coupled land–atmosphere DA systems, the EnOI could be even more beneficial since we only need one model integration of the coupled land–atmosphere system. The caveat of the EnOI is how to determine a realistic background error;



in this study, we use the forecast spread from the ensemble open-loop run, and a climatologically sampled  $\mathbf{A}$  of the background error-covariance. The EnOI equations are given as:

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}'(\mathbf{y} - \mathbf{H}\mathbf{x}^f),$$

where

$$\mathbf{K}' = \mathbf{B}'^f \mathbf{H}^T (\mathbf{H} \mathbf{B}'^f \mathbf{H}^T + \mathbf{R})^{-1}$$

and the background error-covariance is:

$$\mathbf{B}'^f = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}'_i - \mathbf{x}') (\mathbf{x}'_i - \mathbf{x}')^T = \frac{1}{N-1} \mathbf{A}' \mathbf{A}'^T,$$

where  $\mathbf{A}'$  is the deviation from the ensemble mean of the prescribed model spread. The number of ensembles  $N$  used to calculate the anomalies is 12 for the EnOI. In the EnOI, we only update a single model run instead of updating the whole ensemble as in the EnKF. This means that there is no ensemble spread after the analysis and hence no information about the uncertainty.

The EnOI with a climatological background error (EnOI-Clim) solves the same equations as outlined above. The differences between the EnOI and EnOI-Clim arise from the way the ensemble anomalies are computed. We calculate the EnOI-Clim anomalies by sampling from the model spinup (single model run). First, we create a climatology for each of the assimilation windows, and then  $N$  values from that time of the day are subtracted from the climatology, and hence creating the ensemble anomalies. Following Counillon and Bertino [69], we introduce a scaling factor  $\alpha \in (0, 1]$ , which in this work is set to  $\alpha = 0.3$  to dampen the ensemble variance. For the EnOI-Clim, we choose  $N = 122$  (days) to represent the number of ensembles. Our ensemble was created using March, April and May data from the years 2012, 2013 and 2014; we choose to use a stationary sampled climatology in the representation of the ensemble anomalies. Later work will address whether a seasonally varying climatology would have a positive impact on the EnOI-Clim analysis results. The downside with the EnOI-Clim method is that it will not represent the errors of the day; however, it is expected that, over time, useful information will be transferred from the satellite observation and into the model via this analysis.

### 3.5. Model and Observation Errors

Models are simplified representations of the real world and have different sources of errors. For a land surface model, some of these errors are: (i) error of representativeness, (ii) error in model parameters/parameterization of physical processes, and (iii) external errors such as atmospheric forcing, land cover classification and sand and clay fraction. These different errors can in theory be represented in the DA system. Here we represent errors in the forcing (by perturbing the different forcing fields) and model parameter errors (by perturbing the model parameters, which partially encompass the sub-grid scale representativeness errors).

The perturbations done to the atmospheric forcing follow what is done in [23,49,67,70]. Time-series of cross-correlated forcing fields are generated using an autoregressive lag-1 model (AR(1)). These perturbations in the atmospheric forcing allow for an ensemble of model runs, where the spread represents the model uncertainty. Precipitation and shortwave radiation have a lower bound of zero, the perturbations of these variables are therefore multiplicative. Perturbations in longwave radiation are additive. To avoid bias in the forcing files, the perturbations are constrained to zero for the additive variables and to one for the multiplicative variables. The pseudorandom field  $q_t$  is given by:

$$q_t = \alpha q_{t-1} + \beta w_t,$$

where  $\alpha = \frac{\Delta t}{\tau}$  and  $\beta = \sqrt{(1 - \alpha^2)}$  where  $\tau = 24$  h and  $\Delta t = 1$  h. The  $w_t$  is a normally distributed random field. To ensure physical consistency in the perturbation parameters (e.g., increase in longwave radiation gives decrease in shortwave radiation), we impose cross-correlations on the pseudorandom fields using the correlations given in Table 1. The ensemble is kept unbiased centred at the original variable, which implies that the original forcing variable is an ensemble member and at the same time is the ensemble mean. Note that, since the land surface model is nonlinear, this needs not to be true for the prognostic model variables [71].

Errors in the model parameters are based on perturbation of the clay and sand fractions. In SURFEX, the clay and sand fraction acts as proxy variables for the saturated, wilting point and field capacity volumetric water content. Thus, these changes in the sand and clay fractions will ultimately lead to different model physics for the different ensemble members. Since these fractions are bounded by 0 and 1, we use a Logit-normal transformation following the approach in [72]. The parameters are first logit transformed; then, we add  $N$  ensemble members of Gaussian white noise to the logit transformed variable before we do the inverse transformation. This is done for each individual model grid-cell using a uniform spatial random variable and a standard deviation of  $\pm 10\%$ .

**Table 1.** Perturbation parameters for the atmospheric variables precipitation (PRECIP), shortwave (SW) and longwave (LW) downward radiation and the model parameters (clay and sand fractions). The perturbations are multiplicative (M), additive (A) and Logit-normal. The cross-correlation between the variables are given in the Cross-corr columns.

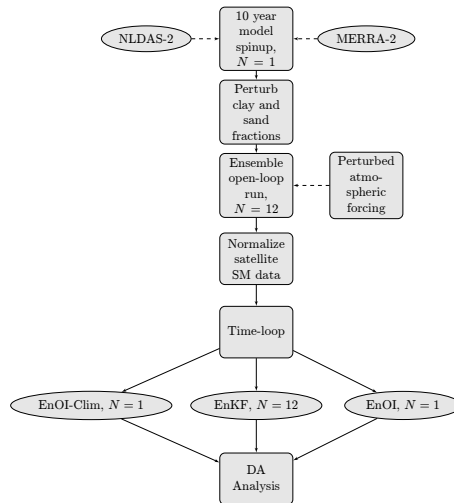
Variable	Type	Std Dev	Cross-Corr with Perturbations		
			PRECIP	SW	LW
PRECIP	M	0.5	1	−0.8	0.5
SW	M	0.3	−0.8	1	−0.5
LW	A	30 W/m <sup>2</sup>	0.5	−0.5	1
clay	Logit normal	0.1	-	-	-
sand	Logit normal	0.1	-	-	-

The setting of observation errors is based on the assumption of a linear relationship between the dynamic range of SM values and the errors [66,73]. This relationship is given as  $(\theta_{fc} - \theta_{wilt})$ , where  $\theta_{fc}$  is the field capacity and  $\theta_{wilt}$  is the wilting point, they both depend on the soil texture and vegetation type for each grid-cell. This relationship is scaled with a dimensionless scaling coefficient  $\lambda^o$ , resulting in an observation error given by:  $\lambda^o(\theta_{fc} - \theta_{wilt}) \text{ m}^3\text{m}^{-3}$ . Setting  $\lambda^o = 0.3$ , results in a spatial average observation error of  $0.026 \text{ m}^3\text{m}^{-3}$ , which is slightly larger than the spatial average reported by De Lannoy and Reichle [12] in their study over CONUS. We assume a diagonal  $\mathbf{R}$  matrix with no cross-correlated observation errors. In this work, we choose to use the same observation error for the three different satellite products. This is because the current study is on of the first that compares DA analyses of the three Level-2 soil moisture products. Hence, we would like to compare them on equal ground. Further tuning of the observation error would have to be addressed in future work.

#### 4. Experiments and Skill Metrics

The land surface DA experiments are designed as follows. Perturbations are applied to the model parameters (clay and sand fractions), in order to represent model parameter errors. The atmospheric forcing is perturbed according to the process defined in Section 3. This creates an ensemble of model configurations, and the ensembles are propagated 3 h forward in time by the ISBA LSM. Given the model background error from the ensemble and the observational error, the EnKF weights each individual ensemble and updates the state vector. The updated state vector is the initial condition for the next model integration. Given correct specification of background and observation errors the updated state vector should not be worse and should in general be superior to the model or satellite

alone. Information from the observation layer (layer two in the state vector) is propagated downward in the analysis update by  $\mathbf{K}$  and subsequently by the model integration. In the EnOI experiments, we use either the ensemble open-loop model integration or the climatological sampled background covariance matrix in the analysis step—see Figure 1. The full list of model experiments is found in Table 2.



**Figure 1.** The experiments are setup using a 10 year spinup for the land surface initial conditions. From the initial state, we perform three different experiments: (i) the EnKF with evolving background error; (ii) the EnOI using the background error from the ensemble open-loop run, and (iii) the EnOI using a climatological background error as the ensemble spread.

The output SM is given for the eight top soil layers, which represent the first metre of soil. We create a surface zone soil moisture (sfzsm, layers 1 to 2) and a root-zone soil moisture (rzsm, layers 1 to 8) product, from weighting the different layers according to their depth. We use two sets of skill metrics to evaluate our experiments; one is the internal DA diagnostics, which check that the conditions for applying the different DA algorithms are met and also that the output is reasonable given the errors set in the system. In addition, we make sure that the biases in the system are removed and that the expected model and observation errors correspond to the actual errors. To diagnose the settings of the model and observation errors, we compute the standard deviation of the normalized innovations. This metric is given as:

$$\text{Std}\left(\frac{\mathbf{O} - \mathbf{F}}{\sqrt{\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}}}\right) \sim 1, \quad (1)$$

where  $\mathbf{O}$  is the observations and  $\mathbf{F}$  is the model forecast, and the  $\mathbf{O} - \mathbf{F}$  is the observation-minus-forecast or innovations. If this metric is larger than 1, the actual errors are said to be underestimated, and overestimated if it is smaller than 1 [18,27]. In addition, we compute the lagged auto-correlation between the  $\mathbf{O} - \mathbf{F}$  pairs for each assimilation window, then take the average over the assimilation windows and the domain to create an auto-correlation function. The  $\mathbf{O} - \mathbf{F}$  sequence is assumed to behave as white noise in the Kalman Filter equations; if there are persistent biases, the auto-correlation function will differ from that of a white noise process. Hence, we can diagnose whether or not the system is bias free and close to behaving optimally [18].

The other metrics we use are the Pearson correlation coefficient ( $R$ ), and the unbiased root mean square difference (ubrmsd), computed between the model outputs and the individual in situ stations. The 95% confidence interval (CI) for the correlation coefficient is computed using the Fisher

Z-transformation. The metrics are computed using all three hourly values as long as the in situ station has more than 150 measurements. The domain average metrics are computed by averaging over all the in situ stations, without any weighting depending of the clustering of the stations as in De Lannoy and Reichle [70]. This could mean that we are weighting densely sampled regions too much in our domain average metrics, and in the worse case lead to an overoptimistic CI.

**Table 2.** List of experiments performed. The open-loop run using NLDAS-2 forcing is given as eOL\_NLDAS, while the open-loop run using MERRA-2 forcing is given as eOL\_MERRA. The data assimilation runs are either the EnKF, EnOI or EnOI-Clim with the corresponding atmospheric forcing prescribed to the LSM. For the satellite skill experiments we also add M (MERRA-2) or N (NLDAS-2) to the experiment names depending on atmospheric forcing used. CPU cost is based on the total number of runs needed to do the analysis, for the EnKF this is 12 + 12 (eOL + EnKF).

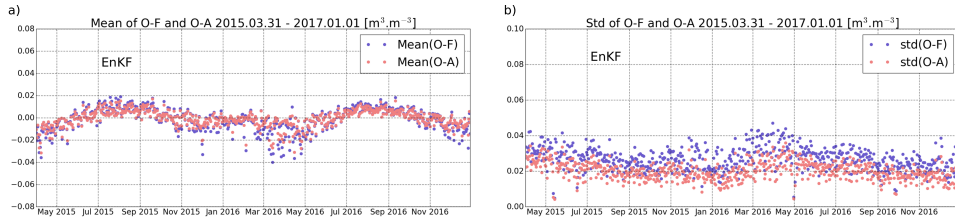
Exp.	$N_{\text{ens}}$	Param. Pert.	Rel. Obs. Error $\lambda$	Satellite Data	Forcing	CPU Cost
eOL_MERRA2	12	yes	-	-	MERRA-2	12
EnKF_MERRA2	12	yes	0.3	SMAP	MERRA-2	24
EnOI_MERRA2	1	yes	0.3	SMAP	MERRA-2	13
EnOI_MERRA2-Clim	1	no	0.3	SMAP	MERRA-2	2
eOL_NLDAS	12	yes	-	-	NLDAS-2	12
EnKF_NLDAS	12	yes	0.3	SMAP	NLDAS-2	24
EnOI_NLDAS	1	yes	0.3	SMAP	NLDAS-2	13
EnOI_NLDAS-Clim	1	no	0.3	SMAP	NLDAS-2	2
Satellite skill experiments						
EnOI_ESA_M	1	yes	0.3	ESA CCI	MERRA-2	13
EnOI_SMAP_M	1	yes	0.3	SMAP	MERRA-2	13
EnOI_SMOS_M	1	yes	0.3	SMOS	MERRA-2	13
EnOI_ESA_N	1	yes	0.3	ESA CCI	NLDAS-2	13
EnOI_SMAP_N	1	yes	0.3	SMAP	NLDAS-2	13
EnOI_SMOS_N	1	yes	0.3	SMOS	NLDAS-2	13

## 5. Data Assimilation Results and Discussion

### 5.1. Filter Performance EnKF vs. EnOI Using Data Assimilation Diagnostics

In this section, we compare the EnKF vs. EnOI using output information from the DA system, these skill metrics indicate how and if the DA system is performing according to its underlying assumptions. First, we compute the domain average filter performance, this is done separately for each of the assimilation windows; then, a final aggregation is applied to give the domain average for a given day. Figure 2a shows the mean O-F and O-A difference for the EnKF; plots for the EnOI and EnOI-Clim are not shown as the results are very similar to that of the EnKF. We notice that there is a slight wet bias of  $0.01 \pm 0.01 \text{ m}^3\text{m}^{-3}$  during the two summer seasons for the EnKF, this is also seen for the EnOI and EnOI-Clim (not shown). In general, the filters are moving the model forecast closer to the observations (the O-F difference is on average larger than the O-A difference). From Figure 2a, we interpret the system as close to bias free (long term mean close to zero), thus the CDF-matching has removed most of the systematic errors between the model and satellite data.

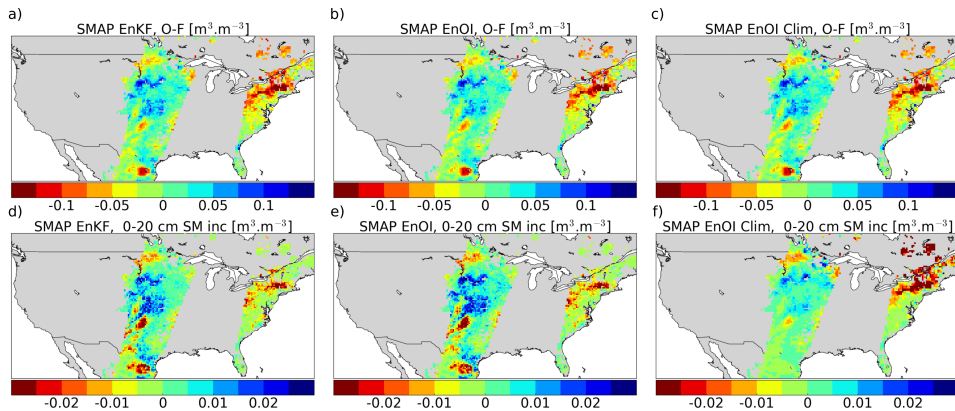
Figure 2b shows the standard deviation of the O-F and O-A differences; this illustrates the typical size of the O-F/A difference. The long-term means of the O-F and O-A spatial standard deviations are 0.0136/0.0105, 0.0139/0.0102 and 0.0139/0.011  $\text{m}^3\text{m}^{-3}$  for the EnKF, EnOI and EnOI-Clim, respectively. The largest differences in standard deviations of O-F vs. O-A are seen for the EnOI. An explanation for this is that in the EnOI the ensemble spread does not change after the analysis, while, for the EnKF, the ensemble spread decreases after the analysis. The smaller spread will in turn result in a larger weight to the model forecast, hence the corrections towards the satellite observations will not be as large as in the EnOI.



**Figure 2.** (a) domain average Observation-minus-Forecast (O-F) soil moisture ( $\text{m}^3\text{m}^{-3}$ ) (blue) and Observation-minus-Analysis soil moisture ( $\text{m}^3\text{m}^{-3}$ ) (red) residual for the EnKF; (b) domain standard deviation of O-F ( $\text{m}^3\text{m}^{-3}$ ) (blue) and O-A soil moisture ( $\text{m}^3\text{m}^{-3}$ ) (red) for the EnKF.

In Figure 3a–c, the instantaneous O-F residuals are given for the EnKF, EnOI and EnOI-Clim. The differences between the O-F residuals are small for the three filters, indicating that the model runs are quite similar, and that none of the filters introduces biases in their analysis. The observation layer increments (analysis-minus-forecast, not shown), show minor differences for the EnKF and the EnOI, which means that the background error-covariances in the two filters are close to identical. This suggests that the updates in the background error-covariance of the EnKF in the analysis step are minor (at least until 16 May 2015).

Figure 3d–f show the instantaneous increments for the top four layers; for the EnKF and the EnOI, the magnitude of the increments are close to those in the observation layer (plot not shown), while, for the EnOI-Clim, they are slightly smaller. By comparing Figure 3a–c with Figure 3d–f, we see that the three filters add/remove water to/from the top four layers when the observations are wetter/drier than the model forecast. For the EnKF and EnOI, an O-F residual of  $\pm 0.1 \text{ m}^3\text{m}^{-3}$  is translated into an increment of  $\pm 0.02 \text{ m}^3\text{m}^{-3}$  in the top four soil layers. The EnOI-Clim has the smallest increments; this is likely because the climatological background error underestimates the actual background error for this analysis step.

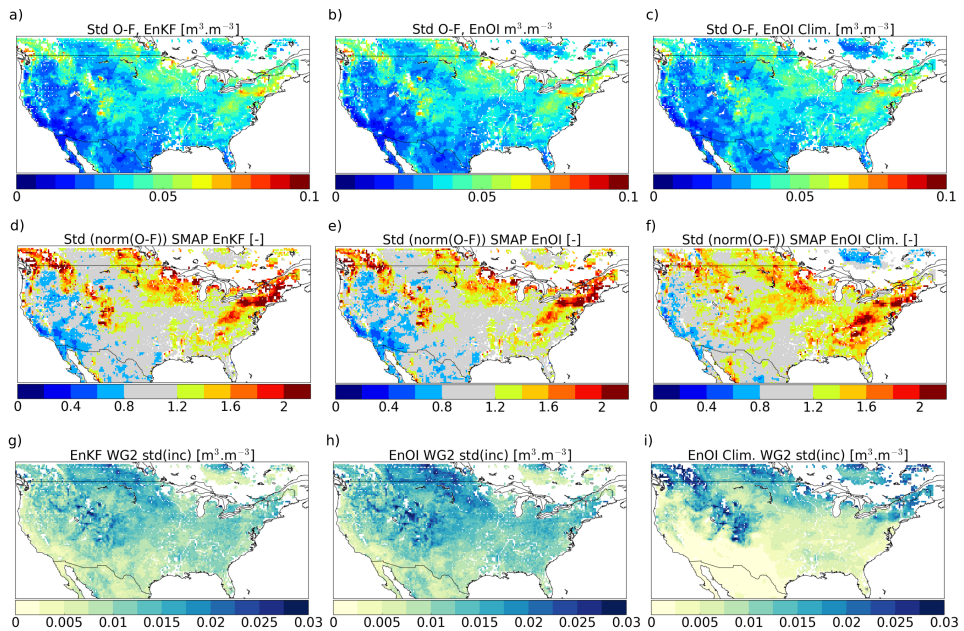


**Figure 3.** Instantaneous analysis output from 16 May 2015 at 1200 UTC for the EnKF, EnOI and EnOI-Clim; (a–c) observation-minus-Forecast (O-F) soil moisture residual ( $\text{m}^3\text{m}^{-3}$ ) for the model observation layer; (d–f) soil moisture increments ( $\text{m}^3\text{m}^{-3}$ ) for the top four soil layers 0–20 cm. Blue/red colours indicate positive/negative O-F residuals and increments.

A soil-layer vs. correlation plot (not shown) indicated that the correlation between the different soil layers decreased to zero at layer five, before it increased again. This was likely due to spurious correlations between the observation layer and the deeper layers. It was difficult to judge how physically consistent the updates in the lower layers were, compared to the update in the observation layer. For this reason, we only update the top four layers in the ISBA model, instead of updating the

top eight layers (representing the first metre of soil). We therefore found it better to let the model physics transfer the analysis information further down in the root-zone.

Spatial maps of the O-F standard deviations are shown in Figure 4a–c, the figures illustrate the typical spatial residual between the SMAP observations and the model forecast. The domain averages for the three filters are 0.0354 (EnKF), 0.0355 (EnOI) and 0.0354  $\text{m}^3\text{m}^{-3}$  (EnOI-Clim), which is close to what De Lannoy and Reichle [70] found for SMOS SM innovations. Typically, dry/wet regions (west/east) have small/large O-F differences, respectively. Large O-F differences are especially found in the Rocky and Appalachian mountain ranges. Soil moisture variability is expected to be higher in the mountainous regions, which might be a cause for the larger O-F differences seen here. The differences in the spatial patterns between Figure 4a,b are minor, indicating that the EnKF (flow dependent background error) and the EnOI (with the ensemble open-loop static background error) are very similar.



**Figure 4.** Assimilation diagnostics for the period from 31 March 2015 to 31 December 2016, for the EnKF, EnOI and EnOI-Clim. (a–c) standard deviation of the Observation-minus-Forecast (O-F) soil moisture residual ( $\text{m}^3\text{m}^{-3}$ ), blue/red colours indicate low/high O-F residuals; (d–f) standard deviation of normalized innovations (dimensionless), blue/red colours indicate too large/small assumed errors; (g–i) standard deviation of the observation layer soil moisture increments ( $\text{m}^3\text{m}^{-3}$ ), yellow/blue colours indicate low/high typical increments.

Figure 4d–f show spatial maps of the standard deviation of the normalized innovations—see Equation (1) for a detailed description. A value close to 1 (grey regions) means that the assumed errors are close to the actual errors for this grid-cell. Values smaller than 1 (blue regions) have assumed errors that are typically too large compared to the actual errors, while red regions (values larger than 1) have assumed errors that are typically smaller than the actual errors. Domain average values for the standard deviation of the normalized innovations are 1.18, 1.15 and 1.18 (dimensionless), for the EnKF, EnOI and EnOI-Clim, respectively. The spatial patterns of the normalized innovations are close to identical for the EnKF and EnOI; the EnOI-Clim has a domain average very close to that of the EnOI and EnKF, even though the spatial pattern differs. The blue regions in the west (too large assumed

errors) are not as prominent in the EnOI-Clim filter. This is likely because of a low ensemble spread in the dry regions for the climatologically sampled background error in the EnOI-Clim. Regions where the assumed errors are too small (red) cover larger parts of the domain for the EnOI-Clim compared to the EnKF and EnOI. A likely reason for this is that the computation of the climatology averages out precipitation events, leading to a small spread in the ensemble.

Comparing Figure 4a–c to Figure 4d–f, we notice that regions with large O-F residuals tend to be regions where the model and observation errors are underestimated, i.e., too low spread or too low observation error, or both, while regions with low O-F residual have model and observation errors that are too large (blue regions). Again, we see that there is not much difference between an evolving background error in the EnKF and a static ensemble background error in the EnOI.

Figure 4g–i show the standard deviations of the observation layer increments. The domain average of a typical observation layer increment for the EnKF, EnOI and EnOI-Clim are 0.008, 0.009 and  $0.007 \text{ m}^3 \text{ m}^{-3}$ . The EnKF and EnOI spatial patterns are very similar, with a lower magnitude for the EnKF than the EnOI, which is likely due to the analysis update of the background error. Dry regions in the south and southwest have smaller increments than the vegetated wet regions in the mountains and to the east. Typical increments for the EnOI-Clim are small in the south and southwest, most likely because the climatology is not able to create a large enough ensemble spread, while the mountainous regions have larger increments because of a greater variability in SM, which is likely increasing the ensemble spread.

The final assessment of the EnKF, EnOI and EnOI-Clim filter performance is done by looking at the lagged auto-correlation between pairs of O-F residuals. From DA theory, the assumption is that the innovations should be white noise for the filter to perform close to “optimal”. Lag is given in days and averaged over the different analysis times and over the spatial domain. The plot (not shown) of the lagged autocorrelation yields similar results for all three filters. The auto-correlation drops towards zero (values between 0.05 and 0.2) but does not follow the same pattern as white noise. This could be an indication of a remaining bias between the model and the observations. However, it is small and in alignment with what other studies have found for SMAP DA [18].

Overall, the DA statistics for the EnKF and EnOI show that, from the perspective of the DA diagnostics, they perform similarly. For the EnOI-Clim, further study is needed to understand how best to capture the errors of the day in a climatologically sampled background error-covariance.

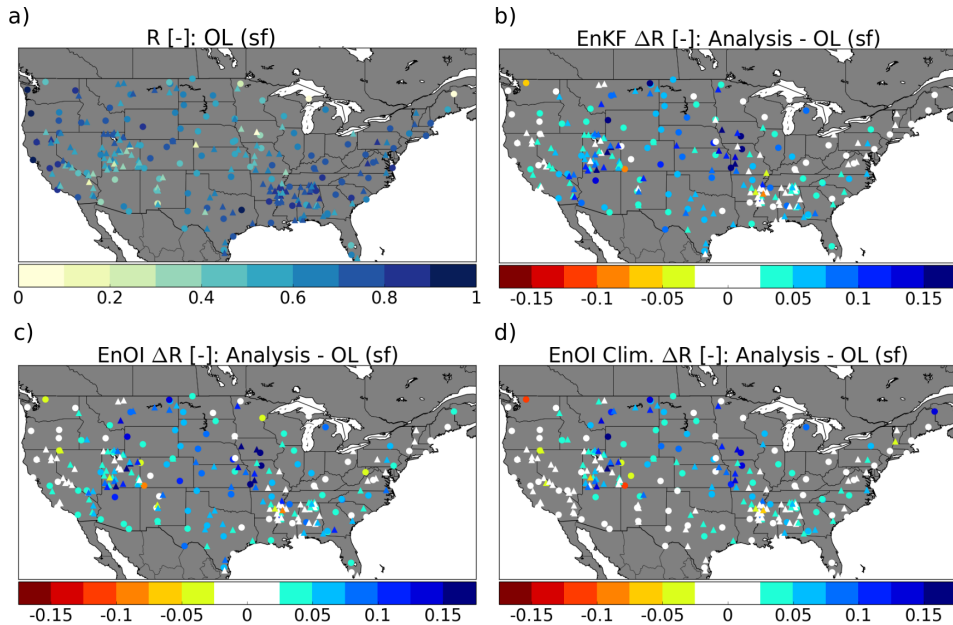
## 5.2. EnKF vs. EnOI; Comparison with In Situ SM Data

### 5.2.1. Correlation Skill

While internal DA diagnostics are useful for evaluating the land DA system’s consistency with regard to DA assumptions, it does not convey any information about the system performance relative to independent data. For such an evaluation, we use the SCAN and USCRN in situ networks. Figure 5a shows the Pearson R correlation coefficient between the ensemble open-loop run using MERRA-2 forcing (eOL\_MERRA2) and the SCAN (triangles) and the USCRN (circles) stations. The eOL\_MERRA2 run alone clearly has good skill in the southeast, while, for example, in the mountainous regions (Rockies), it shows lower correlation values. The domain average R between the eOL\_MERRA2 and SCAN is 0.62, while it is 0.66 between the eOL\_MERRA2 and USCRN; the domain average surface zone statistics are summarized in Table 3. For the surface zone correlation, both the EnKF (0.67/0.7) and the EnOI (0.66/0.7) are found to improve the correlation with the SCAN and USCRN in situ stations at the 95% significance level, respectively.

A difference plot of the EnKF\_MERRA2-eOL\_MERRA2 ( $\Delta R$ ) correlations is given in Figure 5b. Here we note that regions in Figure 5a showing lower R-values, tend to have a positive  $\Delta R$ , while for regions in the southeast, east and west (California), there is little change in  $\Delta R$ . The largest positive impact from the land DA is seen in the Midwest. The high number of in situ stations with improved

correlations make us confident that the large scale error settings (for model and observation errors) in the EnKF are reasonable.



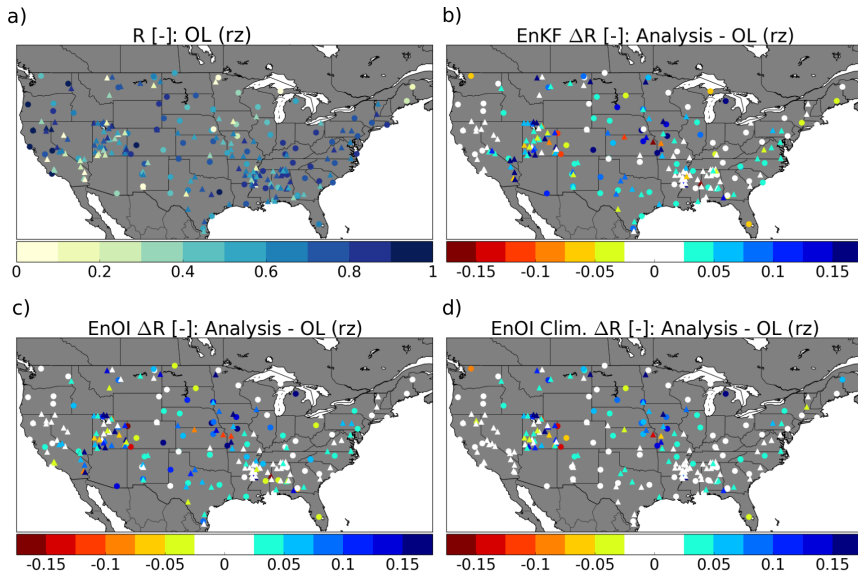
**Figure 5.** (a) Pearson R correlation coefficient between the ensemble open-loop (eOL) surface zone soil moisture and the SCAN (triangles) and USCRN (circles) stations. Yellow/blue indicate low/high correlation; (b)  $\Delta R$  (EnKF correlation minus eOL correlation), blue/red colours indicate positive/negative impact (226 of 259 stations improved); (c)  $\Delta R$  EnOI (190 of 259 stations improved); (d)  $\Delta R$  EnOI-Clim (183 of 259 stations improved).

For the EnOI\_MERRA2, in Figure 5c, much of the spatial pattern is the same as for the EnKF\_MERRA2 in Figure 5b; however, more stations show a negative impact. As for the EnKF\_MERRA2, the largest positive impact is seen in the Midwest. We emphasize that, in the difference plots, we have not taken into account if the individual station improvements are statistically significant or not in the colour coding; we only compute the difference in correlation and count the number of positive vs. negative values. The significance of the improvements is taken into account when we calculate the domain average statistics found in Table 3.

The EnOI\_MERRA2-Clim in Figure 5d shows a band going from the northwest to the southeast where there is little change in R; and this region is also the region with the smallest standard deviation of the increments in Figure 4. For the EnOI\_MERRA2-Clim, there are more stations with negative values for the  $\Delta R$  than for the EnKF and EnOI.

The root-zone eOL\_MERRA2 SM correlations with the in situ networks are given in Figure 6a; and, again, we notice high correlations in the southeast and east. Lower correlations are found in the Midwest and in the mountainous region (Rockies). Computing the domain average for the eOL\_MERRA2 yields an R of 0.63 and 0.68 for the SCAN and USCRN stations, domain average statistics for the root-zone SM can be found in Table 3.

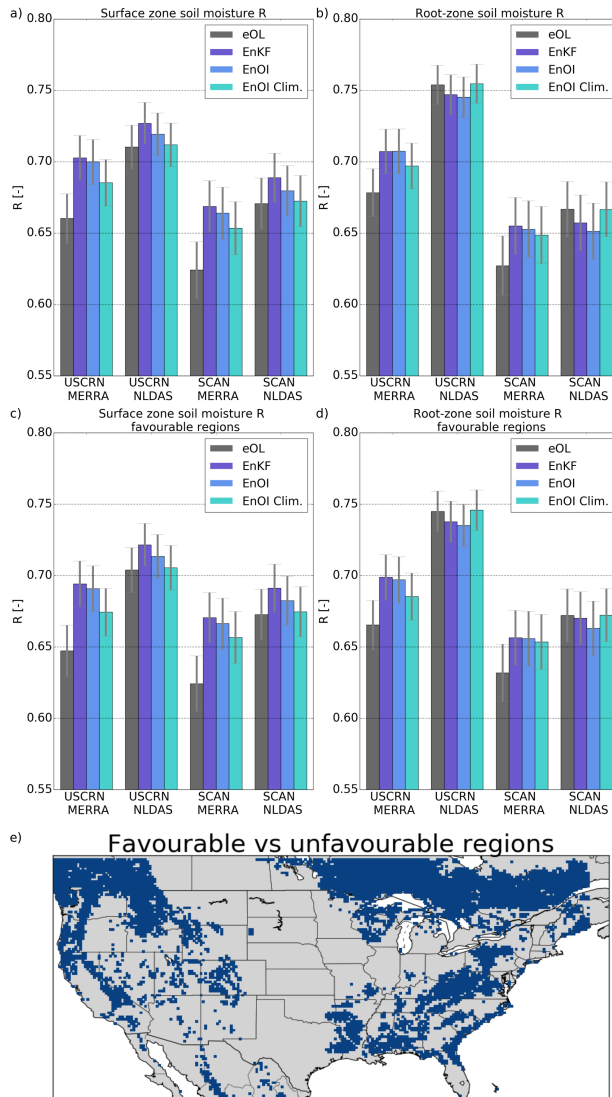




**Figure 6.** (a) Pearson R correlation coefficient between the ensemble open-loop (eOL) root-zone soil moisture and the SCAN (triangles) and USCRN (circles) stations. Yellow/blue indicate low/high correlation. (b)  $\Delta R$  (EnKF correlation minus eOL correlation), blue/red colours indicate positive/negative impact (153 of 223 stations improved). (c)  $\Delta R$  EnOI (135 of 223 stations improved), (d)  $\Delta R$  EnOI-Clim (145 of 223 stations improved).

The EnKF\_MERRA2  $\Delta R$  in Figure 6b does not show as much improvement as for the surface zone. There are more stations with a negative impact and the size of the negative impact is larger than for the surface zone. For the root-zone, there are fewer stations with enough data for the intercomparison, thus the total number of stations is smaller. Although the stations with positive impact from the land DA are more scattered than for the surface zone, we find that most of the stations with the largest improvements in the EnKF\_MERRA2 analysis are seen in the Midwest. The largest positive impacts from the EnOI\_MERRA2 DA analysis are also in the Midwest (see Figure 6c). For the EnOI\_MERRA2-Clim in Figure 6d, the impact of the DA analysis is lower (more stations with white colour coding, showing little positive or negative impact), and again there is a band from west–southeast with little change, and this is most likely due to small increments in the analysis. Although the changes overall in correlations are smaller than for the EnOI\_MERRA2, the total number of stations with improved correlations are larger for the EnOI\_MERRA2-Clim.

The summary statistics of the eOL, EnKF, EnOI and EnOI-Clim assimilation of SMAP data are given in Figure 7 and Table 3. In Figure 7a, the EnKF and EnOI driven by MERRA-2 forcing show a significant improvement in the surface zone (non-overlapping confidence intervals) for both the SCAN and USCRN networks, when compared to the eOL\_MERRA2. In the same figure, only for EnKF, EnOI and eOL driven by NLDAS forcing, we see that there is an improvement, but it is not statistically significant (i.e., overlapping confidence intervals). One reason for this might be the already high correlation between the eOL\_NLDAS and the in situ stations. The NLDAS precipitation correction filters out random errors in the forcing data, which leaves little room for improvement in the land DA analysis. We also note that the EnKF using MERRA-2 forcing has almost the same skill as the eOL\_NLDAS. This surface zone improvement in the EnKF\_MERRA2 analysis towards an eOL\_NLDAS run driven by observation corrected precipitation shows that assimilation of SMAP-L2 data could be a promising way to improve the land surface state in regions of the world where high-quality meteorological data are unavailable.



**Figure 7.** Summary statistics for the Pearson R correlation coefficient between the model and in situ data, a 95% confidence interval for the domain averaged correlation values are given as grey error bars. (a) domain average surface zone R, eOL (black), EnKF (dark blue), EnOI (light blue) and EnOI-Clim (turquoise); (b) domain average root-zone R; (c) domain average R “favourable” regions surface zone; (d) domain average R “favourable” regions root-zone; (e) favourable (grey land pixels) vs. unfavourable (blue land pixels) regions for soil moisture DA. See text for description of “favourable” vs. “unfavourable” regions.

**Table 3.** Surface and root-zone Pearson R correlation coefficient and unbiased root mean square difference (ubrmsd) with SCAN and USCRN in situ networks. Confidence intervals (CI) are given in half values of the 95% interval for the individual experiments. The experiments with highest skill metrics are highlighted in bold.

Exp.	SCAN <sup>a,b</sup>			USCRN <sup>c,d</sup>		
	R	CI	Ubrmsd	R	CI	Ubrmsd
sfzsm (0 – 5) cm, (m <sup>3</sup> m <sup>-3</sup> )						
eOL_MERRA	0.62	±0.020	0.057	0.66	±0.017	0.052
EnKF_MERRA2	<b>0.67</b>	<b>±0.018</b>	<b>0.053</b>	<b>0.70</b>	<b>±0.016</b>	<b>0.049</b>
EnOI_MERRA2	0.66	±0.018	0.054	0.70	±0.016	0.049
EnOI_MERRA2-Clim	0.65	±0.019	0.054	0.68	±0.016	0.050
eOL_NLDAS	0.67	±0.018	0.053	0.71	±0.015	0.048
EnKF_NLDAS	<b>0.69</b>	<b>±0.017</b>	<b>0.052</b>	<b>0.73</b>	<b>±0.015</b>	<b>0.047</b>
EnOI_NLDAS	0.68	±0.018	0.052	0.72	±0.015	0.048
EnOI_NLDAS Clim	0.67	±0.018	0.053	0.71	±0.015	0.048
rzsm (0 – 100) cm, (m <sup>3</sup> m <sup>-3</sup> )						
eOL_MERRA	0.63	±0.021	0.041	0.68	±0.017	0.042
EnKF_MERRA2	<b>0.66</b>	<b>±0.020</b>	<b>0.040</b>	<b>0.71</b>	<b>±0.016</b>	<b>0.040</b>
EnOI_MERRA2	0.65	±0.020	0.040	0.71	±0.016	0.040
EnOI_MERRA2-Clim	0.65	±0.020	0.040	0.70	±0.016	0.041
eOL_NLDAS	<b>0.67</b>	<b>±0.019</b>	<b>0.041</b>	<b>0.75</b>	<b>±0.014</b>	<b>0.041</b>
EnKF_NLDAS	0.66	±0.019	0.041	0.75	±0.014	0.041
EnOI_NLDAS	0.65	±0.020	0.041	0.75	±0.014	0.041
EnOI_NLDAS Clim	0.67	±0.019	0.041	0.75	±0.014	0.041

Number of stations: <sup>a</sup>  $N_{sfzsm} = 151$ , <sup>b</sup>  $N_{rzsm} = 138$ , <sup>c</sup>  $N_{sfzsm} = 108$ , <sup>d</sup>  $N_{rzsm} = 85$ .

Figure 7b and Table 3 show the domain average root-zone correlation between the EnKF, EnOI, EnOI-Clim and the eOL with the in situ stations. The EnKF and EnOI driven by MERRA-2 forcing have equal skill ( $R = 0.71$  for both DA algorithms) for the comparison with the USCRN network, and the improvement relative to the eOL ( $R = 0.68$ ) is close to being statistically significant at the 95% level. The EnOI-Clim also shows an improvement when compared to the USCRN ( $R = 0.7$ ); however, the improvement is not significant at the 95% level. For SCAN using MERRA-2 forcing, all of the filters show improvements compared to the eOL\_MERRA2, but none are statistically significant. Using the NLDAS forcing, both the USCRN and SCAN comparisons show that the filters have a lower skill in the root-zone than the eOL\_NLDAS; however, the difference is not statistically significant. Here, the EnOI\_NLDAS Clim has the best skill, most likely because of the small corrections done by the filter in some regions of the domain (see discussion in Section 5.1).

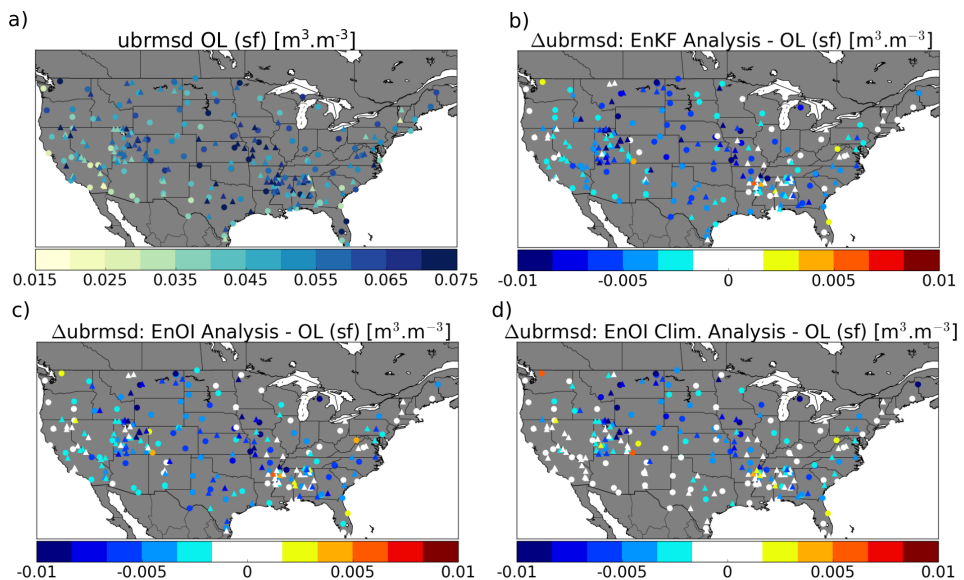
Figure 7c,d are the same as Figure 7a,b, except they show the statistics for the regions favourable for soil moisture DA. These are regions where the water fraction in the grid-cell is below 5%, (i.e., excluding lakes), there is no permanent snow or ice, and there is little dense vegetation, i.e., forest, and/or no strong topography in the grid-cell. The only statistically significant improvement in root-zone SM is seen for the EnKF over USCRN stations located in regions favourable for soil moisture DA. Here, the correlation value increases from 0.67 (eOL\_MERRA2) to 0.7 (EnKF). The overall model performance (eOL) is however found to be lower when computed for favourable areas than for the whole domain. A likely reason for this is that the stations in the southeast with high initial correlations are left out of the overall statistics, since they are in a region unfavourable for soil moisture DA. It is expected that the DA analysis will have very little impact over this region, hence it is not included in the overall statistics. The favourable vs. unfavourable regions are in this way not reflecting regions with high initial eOL skill, but regions where we expect that the DA can have a positive impact, regardless of the initial skill of the eOL.

Overall, there were small differences between the EnKF and EnOI filters, while the EnOI-Clim had the lowest correlation skill (Table 3). The small differences between the EnKF and EnOI can be explained by two things: (i) the relatively long time-window between observations in the same

grid-cell allows the spread in the EnKF to approach that of the EnOI, and (ii) the non-chaotic behaviour of the system suppresses the importance of the initial conditions in the forward model integrations. Points (i) and (ii) imply that it is limited how different a forward integration of an ensemble (EnKF) can be from that of a single model trajectory (EnOI). Skill improvements were smaller when utilizing observation corrected forcing (NLDAS); however, we did see that, when utilizing a model derived forcing (MERRA-2) and land DA, we were able to get comparable skill between the land DA analysis and the ensemble open-loop run with observation corrected forcing. We still saw these improvements using the EnOI highlighting the potential use of this computationally cheaper method for estimating soil moisture.

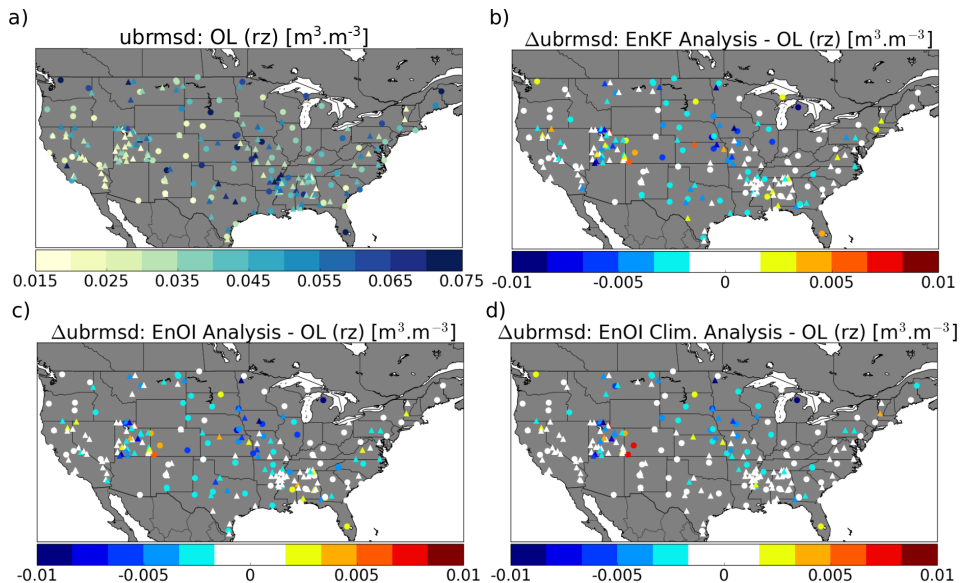
### 5.2.2. Unbiased Root Mean Square Difference

In Figure 8a, we have plotted the surface zone unbiased root mean square difference (ubrmsd), between the eOL (using MERRA-2 forcing) and the SCAN (triangles) and USCRN (circles) networks. Low/high ubrmsd values are depicted with yellow/blue values, respectively. The lowest ubrmsd values are found in regions where the SM variability is naturally limited (dry regions to the west), while higher values are found where the SM variability is high (the wetter east). The  $\Delta$ ubrmsd for the EnKF\_MERRA2 minus the eOL\_MERRA2 metric is plotted in Figure 8b, blue regions indicate that the EnKF\_MERRA2 has improved the ubrmsd skill and the opposite for red regions. Both the EnKF\_MERRA2 and EnOI\_MERRA2 show the largest increase in ubrmsd skill in the Midwest, while the west coast and mountainous region have lower to no increase in ubrmsd skill. For the EnOI\_MERRA2-Clim, little ubrmsd change is seen in the west (along the Pacific coast) and in the south (Arizona, New Mexico and Texas), which is likely linked to the smaller analysis increments found in these regions—see Section 5.1.



**Figure 8.** (a) unbiased root mean square difference (ubrmsd,  $\text{m}^3\text{m}^{-3}$ ) between the ensemble open-loop (eOL) surface zone soil moisture and the SCAN (triangles) and USCRN (circles) stations. Yellow/blue colours indicate low/high ubrmsd; (b)  $\Delta$ ubrmsd (EnKF minus eOL); blue/red colours indicate positive/negative impact of the analysis (229 of 259 stations improved); (c)  $\Delta$ ubrmsd (EnOI minus eOL) (201 of 259 stations improved); (d)  $\Delta$ ubrmsd (EnOI-Clim minus eOL) (197 of 259 stations improved).

The root-zone eOL\_MERRA2 ubrmsd for the SCAN (triangles) and USCRN (circles) networks are plotted in Figure 9a, low/high values are depicted in yellow/blue colours, respectively. The natural variability of root-zone SM is lower than that of the surface, hence more stations have lower values in this figure than in Figure 8a. Figure 9b is the root-zone  $\Delta$ ubrmsd between the EnKF\_MERRA2 and the eOL\_MERRA2. Improvements (blue colours) are mostly found in the Midwest, while most negative EnKF\_MERRA2 analysis impacts are found in the southeast, east and in the mountain ranges to the west. Figure 9c, shows the  $\Delta$ ubrmsd for the EnOI\_MERRA2, as for the EnKF\_MERRA2 most improvements are found in the Midwest. As seen for the surface zone, the domain average absolute value is smaller for the EnOI\_MERRA2-Clim than for the EnKF\_MERRA2 and the EnOI\_MERRA2. A likely reason for the EnOI\_MERRA2-Clim to have more stations with positive improvements is that the analysis increments are smaller for this filter.

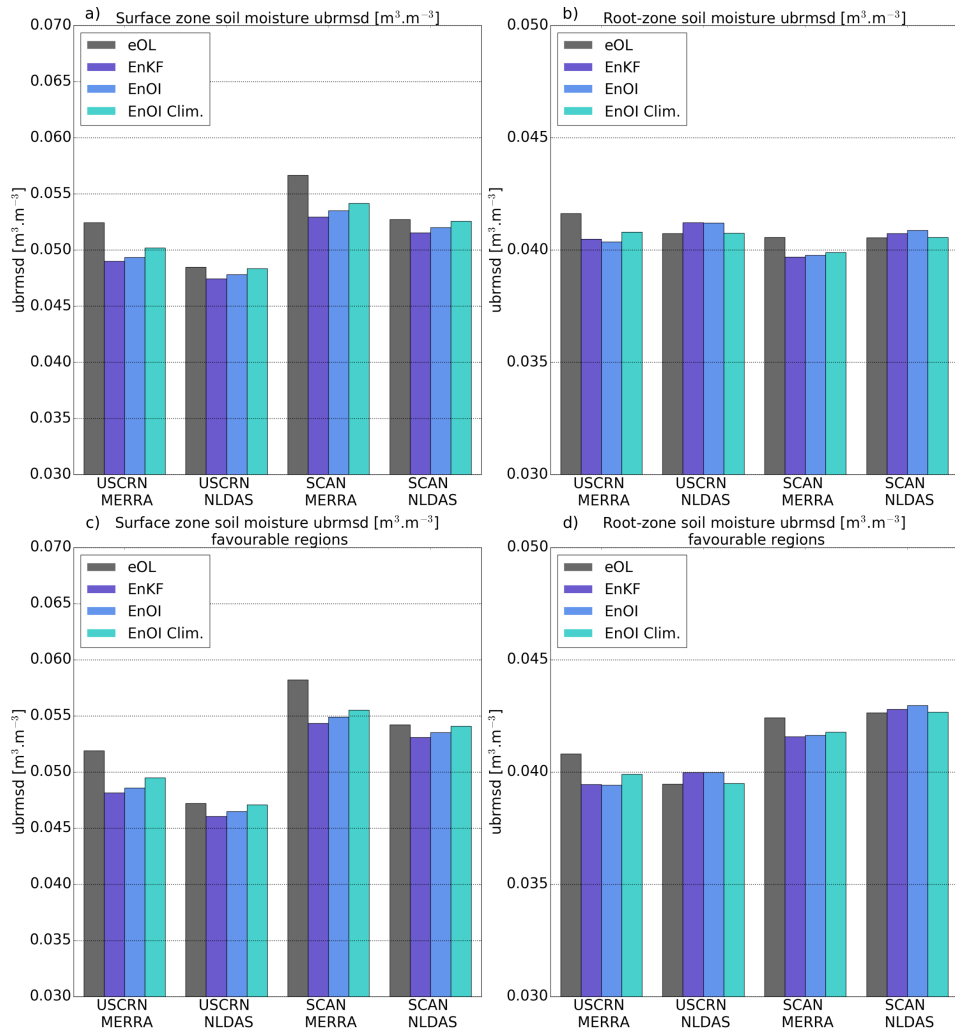


**Figure 9.** (a) unbiased root mean square difference (ubrmsd,  $\text{m}^3\text{m}^{-3}$ ) between the ensemble open-loop (eOL) root-zone soil moisture and the SCAN (triangles) and USCRN (circles) stations. Yellow/blue colours indicate low/high ubrmsd; (b)  $\Delta$ ubrmsd (EnKF minus eOL), blue/red colours indicate positive/negative impact of the analysis (148 of 223 stations improved); (c)  $\Delta$ ubrmsd (EnOI minus eOL) (127 of 223 stations improved); (d)  $\Delta$ ubrmsd (EnOI-Clim minus eOL) (130 of 223 stations improved).

Figure 10a–d present a summary of the ubrmsd statistics for all stations and stations located in regions favourable for soil moisture DA. For the surface zone, in Figure 10a, the largest improvements are seen for the EnKF\_MERRA2 ( $0.053/0.049 \text{ m}^3\text{m}^3$ ), while the EnOI\_MERRA2 ( $0.054/0.049 \text{ m}^3\text{m}^3$ ) and EnOI\_MERRA2-Clim ( $0.054/0.05 \text{ m}^3\text{m}^3$ ) also improve, relative to the eOL\_MERRA2 ( $0.057/0.052 \text{ m}^3\text{m}^3$ ), for SCAN and USCRN, respectively.

The EnKF\_MERRA2 has a spatial average for the surface  $\Delta$ ubrmsd of  $-0.0036 \text{ m}^3\text{m}^3$ ; this is larger in magnitude than the EnOI\_MERRA2 which has a value of  $-0.0031 \text{ m}^3\text{m}^3$  and the EnOI\_MERRA2-Clim with  $-0.0024 \text{ m}^3\text{m}^3$ . These values are close to what De Lannoy and Reichle [12] report in their study for the EnKF assimilating SMOS SM retrievals. Little change is seen in surface ubrmsd for the EnKF\_NLDAS, EnOI\_NLDAS and EnOI\_NLDAS-Clim when compared to the eOL\_NLDAS. However, as for the correlation, we note that the ubrmsd skill of the EnKF\_MERRA2 is close to that of the eOL\_NLDAS.

A root-zone ubrmsd target value of  $0.04\text{ m}^3\text{ m}^{-3}$  is set in the literature for SMAP-L4 products over favourable areas [38,49]. It is useful to compare our analysis results with that of the SMAP-L4 requirements. From the summary statistics in Figure 10d, we see that the SMAP-L4 requirements are met for all three filters over the USCRN stations using MERRA2 forcing. For the model runs using NLDAS forcing, the requirement is already met for the open-loop run over the USCRN stations. The domain average of the  $\Delta\text{ubrmsd}$  for all stations is smaller than for the surface zone,  $-0.001\text{ m}^3\text{ m}^{-3}$  (EnKF\_MERRA2),  $-0.001\text{ m}^3\text{ m}^{-3}$  (EnOI\_MERRA2) and  $-0.0007\text{ m}^3\text{ m}^{-3}$  (EnOI\_MERRA-Clim).



**Figure 10.** Summary statistics for the unbiased root mean square difference (ubrmsd,  $\text{m}^3\text{ m}^{-3}$ ) between the model and in situ data, computed for all stations (a,b) and favourable regions (c,d). (a) domain average surface zone ubrmsd, eOL (black), EnKF (dark blue), EnOI (light blue) and EnOI-Clim (turquoise); (b) domain average root-zone ubrmsd.

Interpreting the improvements in ubrmsd is difficult because, for dry regions, the ubrmsd is naturally low, since the variability of SM for dry regions is low [17,49]. Looking at the spatial distribution of the surface zone  $\Delta$ ubrmsd, in Figure 8b,c, most of the stations show negative values (positive analysis impact). This indicates that the improvements are due to the EnKF\_MERRA2 and the EnOI\_MERRA2 analysis, and not just a statistical artifact. We cannot say the same for the root-zone, as the spatial patterns here are more a mix between an increase and a decrease in skill. In addition, as for dry regions, root-zone SM has a naturally low variability making it difficult to interpret the results.

### 5.3. Satellite Skill: Comparison between ESA CCI, SMAP and SMOS Using the EnOI

In this section, we evaluate the results of the assimilation of the ESA CCI, SMAP and SMOS products. As in the intercomparison of the EnKF, EnOI and EnOI-Clim filters, we evaluate the results in this section against in situ stations. We apply the EnOI with the eOL ensemble as the background error-covariance, since it was found to have comparable skill to the EnKF, while being computationally cheaper—see Sections 5.1 and 5.2. In the assimilation of the three datasets, we use the MERRA-2 forcing, and the ESA CCI, SMAP and SMOS time-series that are all normalized to the eOL run using CDF-matching prior to the assimilation. The assimilation of the three different satellite products and the corresponding evaluation is done for the 31.03.2015 to 31.12.2016 time-period.

Figure 11a–b and Table 4 show the summary of the domain average surface and root-zone correlation between the eOL and in situ, EnOI\_ESA and in situ, EnOI\_SMAP and in situ and EnOI\_SMOS and in situ. The average metrics are only computed for stations covered by the ESA CCI product and in regions favourable for soil moisture DA, hence the domain mean differs from the one shown in Figure 7. The surface zone correlation in Figure 11a shows that the EnOI\_MERRA2 assimilation of SMAP data improves the correlation from 0.63 and 0.64 (eOL) to 0.68 and 0.7 at the 95% significance level for the SCAN and USCRN networks, respectively. The assimilation of SMAP is also seen to have the largest positive impact on the surface ubrmsd skill when comparing the three L2 datasets (Table 4). For the EnOI\_NLDAS analysis, only assimilation of SMAP data slightly improves the correlation when compared to the eOL run. The SMOS assimilation stays neutral, while the assimilation of the ESA CCI product has a decrease in skill compared to the eOL (Table 4).

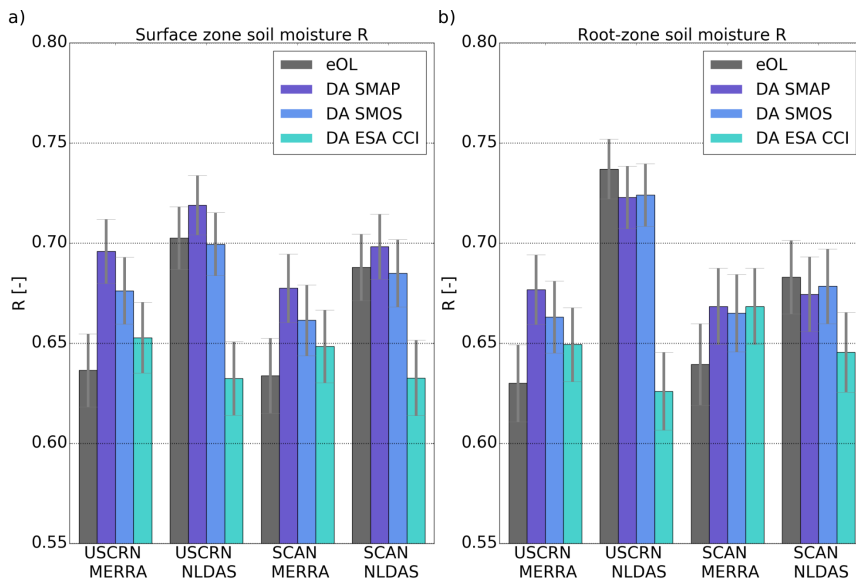
**Table 4.** Satellite skill metrics at SCAN and USCRN networks using MERRA-2 atmospheric forcing. The Pearson R correlation coefficient with half values of the 95% confidence interval (CI) for the individual experiments. The experiments with highest skill metrics are highlighted in bold.

Exp.	SCAN <sup>a,b</sup>			USCRN <sup>c,d</sup>		
	R	CI	Ubrmsd	R	CI	Ubrmsd
sfzsm (0 – 5) cm, (m <sup>3</sup> m <sup>-3</sup> )						
eOL_MERRA	0.63	±0.019	0.061	0.64	±0.018	0.054
EnOI_SMAP	<b>0.68</b>	<b>±0.017</b>	<b>0.057</b>	<b>0.70</b>	<b>±0.016</b>	<b>0.050</b>
EnOI_SMOS	0.66	±0.018	0.058	0.68	±0.017	0.051
EnOI_ESA	0.65	±0.018	0.059	0.65	±0.018	0.053
rzsm (0 – 100) cm, (m <sup>3</sup> m <sup>-3</sup> )						
eOL_MERRA	0.64	±0.020	0.043	0.63	±0.019	0.042
EnOI_SMAP	<b>0.67</b>	<b>±0.019</b>	<b>0.042</b>	<b>0.68</b>	<b>±0.017</b>	<b>0.041</b>
EnOI_SMOS	0.66	±0.019	0.043	0.66	±0.018	0.041
EnOI_ESA	0.67	±0.019	0.042	0.64	±0.018	0.042

Number of stations: <sup>a</sup>  $N_{sfzsm} = 86$ , <sup>b</sup>  $N_{rzsm} = 86$ , <sup>c</sup>  $N_{sfzsm} = 47$ , <sup>d</sup>  $N_{rzsm} = 41$ .

Summary statistics for the the root-zone correlations with in situ data are found in Table 4 and in Figure 11b. The EnOI\_MERRA2 assimilation of SMAP shows a statistically significant improvement for the USCRN stations (from 0.63 to 0.68 at a 95% level), and an improvement from 0.64 to 0.67 for the SCAN stations. Both the SMOS and the ESA CCI assimilation using EnOI\_MERRA2 are also found to increase the root-zone skill over the eOL run. A slight decrease in skill is seen for the EnOI\_NLDAS

assimilation of SMAP and SMOS; however, the confidence intervals are still overlapping. As seen in the surface zone, we also see a decrease in skill using the ESA CCI in the EnOI\_NLDAS assimilation. Spatial maps of  $\Delta R$  (not shown) indicate that the reason for the decrease in skill is not due to individual outliers, but an overall decrease in skill over several stations. We emphasize that the overall skill is computed for the same stations for the ESA CCI, SMAP and SMOS analysis, hence individual station characteristics should not have an impact on the difference in overall statistics between the three datasets. The most likely reason for the SMAP-L2 and SMOS-L2 analysis to yield better results than the ESA CCI analysis is that they are more sensitive to soil moisture (L-band), while most of the observations utilized in the ESA CCI product were C-band, for both the active and passive data. There is reason to expect that, when the ESA CCI product integrates more SMAP and SMOS data, the skill of the ESA CCI assimilation will increase. The assimilation of SMAP and SMOS showed slightly better performance for SMAP than for SMOS. These two satellites both operate in the L-band and use the same radiative transfer equation ( $\tau - \omega$  model). Al-Yaari et al. [74] compared SMAP and SMOS soil moisture data and found that SMAP had better skill. Our results agree with this skill difference between SMAP and SMOS as seen in [74], only here it is not seen by direct intercomparison of the retrieved soil moisture products, but via the SMAP land DA analysis performing better than the SMOS land DA analysis. They point at the difference in ancillary data (land cover dataset), model parameterizations, and RFI mitigation and detection to be some of the reasons for the difference in skill. In addition, the reported instrument error (for Level-1 data) is lower for SMAP ( $\sim 1$  K) compared to SMOS ( $\sim 3\text{--}3.5$  K) [75].



**Figure 11.** Summary statistics for the assimilation of SMAP (dark blue), SMOS (light blue) and ESA CCI (turquoise), surface and root-zone Pearson R correlation coefficient. Statistics are only computed for stations covered by the ESA CCI product; this is to ensure a fair intercomparison. (a) domain average surface zone R, eOL (black), SMAP (dark blue), SMOS (light blue) and ESA CCI (turquoise); (b) domain average root-zone R.

## 6. Conclusions

In this study, we first compared three different data assimilation (DA) methods, the EnKF, the EnOI and the EnOI-Clim by assimilating the SMAP-L2 soil moisture (SM) product over the contiguous United States (CONUS) using MERRA-2 meteorological forcing. The correlation and ubrmsd skill



of the resulting DA analysis was evaluated against independent in situ data from the Soil Climate Analysis Network (SCAN) and the U.S. Climate reference network (USCRN). It was shown that the EnKF had the highest correlation skill (average of the SCAN and USCRN networks) when compared to in situ stations (0.685 surface and 0.681 root-zone), while the correlation skill of the EnOI was close to that of the EnKF (0.68 surface and 0.68 root-zone). The EnKF ubrmsd skill (average of the SCAN and USCRN networks) was  $0.051 \text{ m}^3 \text{ m}^3$  (surface) and  $0.04 \text{ m}^3 \text{ m}^3$  (root-zone), again, the EnOI skill was close with ubrmsd of  $0.0515 \text{ m}^3 \text{ m}^3$  (surface) and  $0.04 \text{ m}^3 \text{ m}^3$  (root-zone). Given that the EnOI is computationally cheaper than the EnKF and it still offers much of the same benefits, it could be a promising alternative to the EnKF. The downside is that it does not use a flow dependent ensemble spread as a measure of the analysis uncertainty. This might not be that important though, as the assimilation window is so long that the ensemble spread of the EnKF is close to that of the ensemble open-loop (eOL) spread, making the differences in background error between the two filters minor. This can be explained by the nature of the soil moisture dynamics, which is stable with respect to initial conditions [76]. Thus, the findings may be extrapolated to other stable dynamical systems. The EnOI-Clim performed poorest, especially for dry regions, where the climatological spread was found to be too low. However, there should be potential to improve the EnOI-Clim results by, for example, looking more into sampling methods of the background error.

The impact of the atmospheric forcing on the analysis results was evaluated by assimilating SMAP data using the EnKF and the EnOI with MERRA-2 (model derived) and NLDAS (observation corrected) atmospheric forcing as model input. The largest improvements from an eOL run in correlation and ubrmsd were seen when we used the MERRA-2 atmospheric forcing data. This dataset had no precipitation correction, yet the initial skill was high when comparing the eOL forced by MERRA-2 data to in situ stations. The EnKF and the EnOI assimilation of SMAP showed statistically significant improvements in the surface zone correlation at the 95% level, while root-zone improvements were only found for USCRN stations located in regions favourable for soil moisture DA. Our results indicate that land DA of SMAP-L2 data can account for random errors in the atmospheric forcing dataset. We base this statement on the fact that our EnKF and EnOI analysis using MERRA-2 forcing; both had comparable skill to the eOL run using NLDAS atmospheric forcing data. This could be important for regions of the world where high-quality meteorological forcing and observation data are missing, such as over the high northern latitudes and Africa.

Finally, we assimilated the ESA CCI, SMAP and SMOS SM products using the EnOI and MERRA-2 atmospheric forcing. Here, we wanted to evaluate the satellite impact on the land DA analysis. We computed surface and root-zone correlations between the three analysis products and the in situ SM data, over regions favourable for soil moisture DA. We found that assimilation of the SMAP-L2 product had the highest correlation skill, 0.69 for the surface and 0.68 for the root-zone SM. The SMOS-L2 product had the second highest overall correlation skill, 0.67 for the surface and 0.66 for the root-zone SM. While the ESA CCI product had the lowest correlation skill, 0.65 for the surface and 0.66 for the root-zone.

Land DA of SMAP data were found to perform better than assimilation of SMOS and ESA CCI data. The differences between the SMAP and SMOS analysis skill were minor, and our analysis results support other studies indicating that SMAP SM data perform slightly better than SMOS SM data when compared to in situ stations [74].

In general, our work supports other studies showing that assimilation of surface SM observations from satellites could improve the surface zone SM in land surface models [8,12,17,23,77]. It also agrees well with the same studies showing the difficulty of getting a statistically significant improvement in the root-zone SM. While the improvements in root-zone skill were minor, we did see that, using SMAP-L2 data, in combination with the MERRA-2 forcing improved the surface zone skill, even when using the computationally cheaper EnOI (in a statistically significant manner). This result highlights the potential of the EnOI and SMAP-L2 combination for improving estimates of surface soil moisture. However, more work is needed on how to propagate this information into the root-zone.

**Author Contributions:** J.B., P.D.H., W.A.L., L.B. and C.A. conceived and designed the data assimilation experiments. D.F., P.D.H., L.B. and C.A. contributed to setting up the land surface model and the data assimilation system. J.B. conducted the experiments, did the analysis and wrote the manuscript. All authors provided comments on the manuscript.

**Funding:** The main author is supported by the Research Council of Norway (NFR PhD-grant 239947, 2015–2019).

**Acknowledgments:** The main author would like to thank Gabriëlle De Lannoy, Eric Wood and Ming Pan for valuable discussions in the construction of this paper. We would also like to thank Sam-Erik Walker for his contributions to the development of the EnKF code. Finally, we would like to thank four anonymous reviewers and the editor for their help in improving this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Loew, A. Impact of surface heterogeneity on surface soil moisture retrievals from passive microwave data at the regional scale: The Upper Danube case. *Remote Sens. Environ.* **2008**, *112*, 231–248. [[CrossRef](#)]
- Seneviratne, S.I.; Corti, T.; Davin, E.L.; Hirschi, M.; Jaeger, E.B.; Lehner, I.; Orlowsky, B.; Teuling, A.J. Investigating Soil Moisture–Climate Interactions in a Changing Climate: A Review. *Earth Sci. Rev.* **2010**, *99*, 125–161. [[CrossRef](#)]
- Wood, A.W.; Lettenmaier, D.P. An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys. Res. Lett.* **2008**, *35*. [[CrossRef](#)]
- Li, H.; Luo, L.; Wood, E.F.; Schaake, J. The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *J. Geophys. Res. Atmos.* **2009**, *114*. [[CrossRef](#)]
- Thober, S.; Kumar, R.; Sheffield, J.; Mai, J.; Schäfer, D.; Samaniego, L. Seasonal Soil Moisture Drought Prediction over Europe Using the North American Multi-Model Ensemble (NMME). *J. Hydrometeorol.* **2015**, *16*, 2329–2344. [[CrossRef](#)]
- Shukla, S.; Sheffield, J.; Wood, E.F.; Lettenmaier, D.P. On the sources of global land surface hydrologic predictability. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 2781–2796. [[CrossRef](#)]
- Reichle, R.; McLaughlin, D.; Entekhabi, D. Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1708–1718. [[CrossRef](#)]
- Kumar, S.V.; Reichle, R.H.; Peters-Lidard, C.D.; Koster, R.D.; Zhan, X.; Crow, W.T.; Eylander, J.B.; Houser, P.R. A land surface data assimilation framework using the land information system: Description and applications. *Adv. Water Resour.* **2008**, *31*, 1419–1432. [[CrossRef](#)]
- Draper, C.S.; Mahfouf, J.F.; Walker, J.P. An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. *J. Geophys. Res. Atmos.* **2009**, *114*. [[CrossRef](#)]
- Lahoz, W.; Khattatov, B.; Ménard, R. *Data Assimilation: Making Sense of Observations*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 1–718.
- Lievens, H.; De Lannoy, G.J.M.; Al Bitar, A.; Drusch, M.; Dumedah, G.; Hendricks Franssen, H.J.; Kerr, Y.H.; Tomer, S.K.; Martens, B.; Merlin, O.; et al. Assimilation of SMOS Soil Moisture and Brightness Temperature Products Into a Land Surface Model. *Remote Sens. Environ.* **2015**, *180*, 292–304.
- De Lannoy, G.J.; Reichle, R.H. Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 4895–4911. [[CrossRef](#)]
- Albergel, C.; Munier, S.; Jennifer Leroux, D.; Dewaele, H.; Fairbairn, D.; Lavinia Barbu, A.; Gelati, E.; Dorigo, W.; Faroux, S.; Meurey, C.; et al. 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.* **2017**, *10*, 3889–3912. [[CrossRef](#)]
- Albergel, C.; Munier, S.; Bocher, A.; Bonan, B.; Zheng, Y.; Draper, C.; Leroux, D.J.; Calvet, J.C. 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.* **2018**, *10*, 1627. [[CrossRef](#)]
- Balsamo, G.; Agusti-Parareda, A.; Albergel, C.; Arduini, G.; Beljaars, A.; Bidlot, J.; Bousserez, N.; Boussetta, S.; Brown, A.; Buizza, R.; et al. Satellite and In Situ Observations for Advancing Global Earth Surface Modelling: A Review. *Remote Sens.* **2018**, *12*, 2038. [[CrossRef](#)]
- Schmugge, T.J.; Kustas, W.P.; Ritchie, J.C.; Jackson, T.J.; Rango, A. Remote sensing in hydrology. *Adv. Water Resour.* **2002**, *25*, 1367–1385. [[CrossRef](#)]

17. Koster, R.D.; Liu, Q.; Mahanama, S.P.P.; Reichle, R.H. Improved Hydrological Simulation Using SMAP Data: Relative Impacts of Model Calibration and Data Assimilation. *J. Hydrometeorol.* **2018**, doi:10.1175/JHM-D-17-0228.1. [[CrossRef](#)] [[PubMed](#)]
18. Reichle, R.H.; De Lannoy, G.J.M.; Liu, Q.; Koster, R.D.; Kimball, J.S.; Crow, W.T.; Ardizzone, J.V.; Chakraborty, P.; Collins, D.W.; Conaty, A.L.; et al. Global Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using Assimilation Diagnostics. *J. Hydrometeorol.* **2017**, *18*, 3217–3237. [[CrossRef](#)] [[PubMed](#)]
19. Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics (part 1). *J. Geophys. Res.* **1994**, *99*, 10143–10162. [[CrossRef](#)]
20. Bonan, B.; Albergel, C.; Zheng, Y.; Barbu, A.; Fairbairn, D.; Munier, S.; Calvet, J.C. An Ensemble Kalman Filter for the joint assimilation of surface soil moisture and leaf area index within LDAS-Monde: Application over the Euro-Mediterranean basin. *HESSD HYMEX 2019*, in press.
21. Oke, P.R. Assimilation of surface velocity data into a primitive equation coastal ocean model. *J. Geophys. Res.* **2002**, *107*, 3122. [[CrossRef](#)]
22. Evensen, G. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dyn.* **2003**, *53*, 343–367. [[CrossRef](#)]
23. Liu, Q.; Reichle, R.H.; Bindlish, R.; Cosh, M.H.; Crow, W.T.; de Jeu, R.; De Lannoy, G.J.M.; Huffman, G.J.; Jackson, T.J. The Contributions of Precipitation and Soil Moisture Observations to the Skill of Soil Moisture Estimates in a Land Data Assimilation System. *J. Hydrometeorol.* **2011**, *12*, 750–765. [[CrossRef](#)]
24. Maggioni, V.; Reichle, R.H.; Anagnostou, E.N. The Efficiency of Assimilating Satellite Soil Moisture Retrievals in a Land Data Assimilation System Using Different Rainfall Error Models. *J. Hydrometeorol.* **2013**, *14*, 368–374. [[CrossRef](#)]
25. Gelaro, R.; McCarty, W.; Suárez, M.J.; Todling, R.; Molod, A.; Takacs, L.; Randles, C.A.; Darmenov, A.; Bosilovich, M.G.; Reichle, R.; et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **2017**, *30*, 5419–5454. [[CrossRef](#)]
26. Mitchell, K.E. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.* **2004**, *109*, D07S90. [[CrossRef](#)]
27. Desroziers, G.; Berre, L.; Chapnik, B.; Poli, P. Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 3385–3396. [[CrossRef](#)]
28. Liu, Y.Y.; Dorigo, W.A.; Parinussa, R.M.; De Jeu, R.A.; Wagner, W.; McCabe, M.F.; Evans, J.P.; Van Dijk, A.I. Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sens. Environ.* **2012**, *123*, 280–297. [[CrossRef](#)]
29. Dorigo, W.A.; Gruber, A.; De Jeu, R.A.; Wagner, W.; Stacke, T.; Loew, A.; Albergel, C.; Brocca, L.; Chung, D.; Parinussa, R.M.; et al. Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sens. Environ.* **2015**, *162*, 380–395. [[CrossRef](#)]
30. Dorigo, W.; Wagner, W.; Albergel, C.; Albrecht, F.; Balsamo, G.; Brocca, L.; Chung, D.; Ertl, M.; Forkel, M.; Gruber, A.; et al. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sens. Environ.* **2017**, *203*, 185–215. [[CrossRef](#)]
31. Gruber, A.; Dorigo, W.A.; Crow, W.; Wagner, W. Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 6780–6792. [[CrossRef](#)]
32. Wagner, W.; Lemoine, G.; Rott, H. A method for estimating soil moisture from ERS Scatterometer and soil data. *Remote Sens. Environ.* **1999**, *70*, 191–207. [[CrossRef](#)]
33. Bartalis, Z.; Wagner, W.; Naeimi, V.; Hasenauer, S.; Scipal, K.; Bonekamp, H.; Figa, J.; Anderson, C. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). *Geophys. Res. Lett.* **2007**, *34*. [[CrossRef](#)]
34. Naeimi, V.; Bartalis, Z.; Wagner, W. ASCAT Soil Moisture: An Assessment of the Data Quality and Consistency with the ERS Scatterometer Heritage. *J. Hydrometeorol.* **2009**, *10*, 555–563. [[CrossRef](#)]
35. Owe, M.; de Jeu, R.; Holmes, T. Multisensor historical climatology of satellite-derived global land surface moisture. *J. Geophys. Res. Earth Surf.* **2008**, *113*. [[CrossRef](#)]
36. Loizu, J.; Massari, C.; Álvarez-Mozos, J.; Tarpanelli, A.; Brocca, L.; Casali, J. On the assimilation set-up of ASCAT soil moisture data for improving streamflow catchment simulation. *Adv. Water Resour.* **2018**, *111*, 86–104. [[CrossRef](#)]

37. Draper, C.S.; Reichle, R.H. Assimilation of passive and active microwave soil moisture retrievals—Draper—2012—Geophysical Research Letters—Wiley Online Library. *Geophys. Res. Lett.* **2012**, *39*, 1–5.
38. Entekhabi, D.; Njoku, E.G.; O'Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; et al. The Soil Moisture Active Passive (SMAP) Mission. *Proc. IEEE* **2010**, *98*, 704–716. [[CrossRef](#)]
39. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.P.; Martinuzzi, J.M.; Font, J.; Berger, M. Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1729–1735. [[CrossRef](#)]
40. Kerr, Y.H.; Waldteufel, P.; Richaume, P.; Wigneron, J.P.; Ferrazzoli, P.; Mahmoodi, A.; Al Bitar, A.; Cabot, F.; Gruhier, C.; Juglea, S.E.; et al. The SMOS soil moisture retrieval algorithm. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1384–1403. [[CrossRef](#)]
41. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.P.; Delwart, S.; Cabot, F.; Boutin, J.; Escorihuela, M.J.; Font, J.; Reul, N.; Gruhier, C.; et al. The SMOS L: New tool for monitoring key elements of the global water cycle. *Proc. IEEE* **2010**, *98*, 666–687. [[CrossRef](#)]
42. Dorigo, W.A.; Wagner, W.; Hohensinn, R.; Hahn, S.; Paulik, C.; Xaver, A.; Gruber, A.; Drusch, M.; Mecklenburg, S.; Van Oevelen, P.; et al. The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1675–1698. [[CrossRef](#)]
43. Dorigo, W.; Xaver, A.; Vreugdenhil, M.; Gruber, A.; Hegyiová, A.; Sanchis-Dufau, A.; Zamojski, D.; Cordes, C.; Wagner, W.; Drusch, M. Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network. *Vadose Zone J.* **2013**, *12*. [[CrossRef](#)]
44. Schaefer, G.L.; Cosh, M.H.; Jackson, T.J. The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *J. Atmos. Ocean. Technol.* **2007**, *24*, 2073–2077. [[CrossRef](#)]
45. Diamond, H.J.; Karl, T.R.; Palecki, M.A.; Baker, C.B.; Bell, J.E.; Leeper, R.D.; Easterling, D.R.; Lawrimore, J.H.; Meyers, T.P.; Helfert, M.R.; et al. U.S. climate reference network after one decade of operations status and assessment. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 485–498. [[CrossRef](#)]
46. Palecki, M.A.; Bell, J.E. U.S. Climate Reference Network Soil Moisture Observations with Triple Redundancy: Measurement Variability. *Vadose Zone J.* **2013**. [[CrossRef](#)]
47. Zreda, M.; Desilets, D.; Ferré, T.P.; Scott, R.L. Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. *Geophys. Res. Lett.* **2008**, *35*. [[CrossRef](#)]
48. Zreda, M.; Shuttleworth, W.J.; Zeng, X.; Zweck, C.; Desilets, D.; Franz, T.; Rosolem, R. COSMOS: The cosmic-ray soil moisture observing system. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 4079–4099. [[CrossRef](#)]
49. Reichle, R.H.; De Lannoy, G.J.M.; Liu, Q.; Colliander, A.; Conaty, A.; Jackson, T.; Kimball, J.; Koster, R.D. *Soil Moisture Active Passive (SMAP) Project Assessment Report for the Beta-Release L4\_SM Data Product*; Technical Report NASA/TM-2015-104606; Goddard Space Flight Center: Greenbelt, MD, USA, 2015.
50. Reichle, R.H.; De Lannoy, G.J.M.; Liu, Q.; Ardzzone, J.V.; Colliander, A.; Conaty, A.; Crow, W.; Jackson, T.J.; Jones, L.A.; Kimball, J.S.; et al. Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using In Situ Measurements. *J. Hydrometeorol.* **2017**, *18*, 2621–2645. [[CrossRef](#)]
51. Reichle, R.H.; Koster, R.D. Bias reduction in short records of satellite soil moisture. *Geophys. Res. Lett.* **2004**, *31*. [[CrossRef](#)]
52. Drusch, M.; Wood, E.F.; Gao, H. Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. *Geophys. Res. Lett.* **2005**, *32*. [[CrossRef](#)]
53. Brocca, L.; Hasenauer, S.; Lacava, T.; Melone, F.; Moramarco, T.; Wagner, W.; Dorigo, W.; Matgen, P.; Martínez-Fernández, J.; Llorens, P.; et al. Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sens. Environ.* **2011**, *115*, 3390–3408. [[CrossRef](#)]
54. Kumar, S.V.; Reichle, R.H.; Harrison, K.W.; Peters-Lidard, C.D.; Yatheendradas, S.; Santanello, J.A. A comparison of methods for a priori bias correction in soil moisture data assimilation. *Water Resour. Res.* **2012**, *48*. [[CrossRef](#)]
55. Kolassa, J.; Reichle, R.; Liu, Q.; Cosh, M.; Bosch, D.; Caldwell, T.; Colliander, A.; Holifield Collins, C.; Jackson, T.; Livingston, S.; et al. Data Assimilation to Extract Soil Moisture Information from SMAP Observations. *Remote Sens.* **2017**, *9*, 1179. [[CrossRef](#)]
56. Masson, V.; Le Moigne, P.; Martin, E.; Faroux, S.; Alias, A.; Alkama, R.; Belamari, S.; Barbu, A.; Boone, A.; Bouysse, F.; et al. The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geosci. Model Dev.* **2013**, *6*, 929–960. [[CrossRef](#)]

57. Noilhan, J.; Mahfouf, J.F. The ISBA land surface parameterisation scheme. *Glob. Planet. Chang.* **1996**, *13*, 145–159. [[CrossRef](#)]
58. Calvet, J.C.; Noilhan, J.; Roujean, J.L.; Bessemoulin, P.; Cabelguenne, M.; Olioso, A.; Wigneron, J.P. An interactive vegetation SVAT model tested against data from six contrasting sites. *Agric. For. Meteorol.* **1998**, *92*, 73–95. [[CrossRef](#)]
59. Decharme, B.; Boone, A.; Delire, C.; Noilhan, J. Local evaluation of the Interaction between Soil Biosphere Atmosphere soil multilayer diffusion scheme using four pedotransfer functions. *J. Geophys. Res. Atmos.* **2011**, *116*. [[CrossRef](#)]
60. Boone, A.; Masson, V.; Meyers, T.; Noilhan, J. The Influence of the Inclusion of Soil Freezing on Simulations by a Soil–Vegetation–Atmosphere Transfer Scheme. *J. Appl. Meteorol.* **2000**, *39*, 1544–1569. [[CrossRef](#)]
61. Boone, A.; Etchevers, P. An Intercomparison of Three Snow Schemes of Varying Complexity Coupled to the Same Land Surface Model: Local-Scale Evaluation at an Alpine Site. *J. Hydrometeorol.* **2001**, *2*, 374–394. [[CrossRef](#)]
62. Faroux, S.; Kaptué Tchuenté, A.T.; Roujean, J.L.; Masson, V.; Martin, E.; Le Moigne, P. 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.* **2013**, *6*, 563–582. [[CrossRef](#)]
63. Wieder, W.R.; Boehnert, J.; Bonan, G.B.; Langseth, M. RegridDED Harmonized World Soil Database v1.2. 2014. Available online: <http://daac.ornl.gov> (accessed on 25 February 2019).
64. Noilhan, J.; Lacarrere, P. GCM grid-scale evaporation from mesoscale modeling. *J. Clim.* **1995**, *8*, 206–223. [[CrossRef](#)]
65. Sakov, P.; Oke, P.R. Implications of the Form of the Ensemble Transformation in the Ensemble Square Root Filters. *Mon. Weather Rev.* **2008**, *136*, 1042–1053. [[CrossRef](#)]
66. Fairbairn, D.; Barbu, A.L.; Mahfouf, J.F.; Calvet, J.C.; Gelati, E. Comparing the ensemble and extended Kalman filters for in situ soil moisture assimilation with contrasting conditions. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 4811–4830. [[CrossRef](#)]
67. Kumar, S.V.; Peters-Lidard, C.D.; Mocko, D.; Reichle, R.; Liu, Y.; Arsenaull, K.R.; Xia, Y.; Ek, M.; Riggs, G.; Livneh, B.; Cosh, M. Assimilation of Remotely Sensed Soil Moisture and Snow Depth Retrievals for Drought Estimation. *J. Hydrometeorol.* **2014**, *15*, 2446–2469. [[CrossRef](#)]
68. Yin, J.; Zhan, X.; Zheng, Y.; Hain, C.R.; Liu, J.; Fang, L. Optimal ensemble size of ensemble Kalman filter in sequential soil moisture data assimilation. *Geophys. Res. Lett.* **2015**, *42*, 6710–6715. [[CrossRef](#)]
69. Counillon, F.; Bertino, L. Ensemble Optimal Interpolation: Multivariate properties in the Gulf of Mexico. *Tellus Ser. A Dyn. Meteorol. Oceanogr.* **2009**, *61*, 296–308. [[CrossRef](#)]
70. De Lannoy, G.J.M.; Reichle, R.H. Global Assimilation of Multiangle and Multipolarization SMOS Brightness Temperature Observations into the GEOS-5 Catchment Land Surface Model for Soil Moisture Estimation. *J. Hydrometeorol.* **2016**, *17*, 669–691. [[CrossRef](#)]
71. Alberto, C.; Marc, B.; Laurent, B.; Geir, E. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdiscip. Rev. Clim. Chang.* **2018**, *9*, e535. [[CrossRef](#)]
72. Aalstad, K.; Westermann, S.; Schuler, T.V.; Boike, J.; Bertino, L. Ensemble-based assimilation of fractional snow-covered area satellite retrievals to estimate the snow distribution at Arctic sites. *Cryosphere* **2018**, *12*, 247–270. [[CrossRef](#)]
73. Mahfouf, J.F.; Bergaoui, K.; Draper, C.; Bouysse, F.; Taillefer, F.; Taseva, L. A comparison of two off-line soil analysis schemes for assimilation of screen level observations. *J. Geophys. Res. Atmos.* **2009**, *114*. [[CrossRef](#)]
74. Al-Yaari, A.; Wigneron, J.P.; Kerr, Y.; Rodriguez-Fernandez, N.; O’Neill, P.E.; Jackson, T.J.; De Lannoy, G.J.; Al Bitar, A.; Mialon, A.; Richaume, P.; et al. Evaluating soil moisture retrievals from ESA’s SMOS and NASA’s SMAP brightness temperature datasets. *Remote Sens. Environ.* **2017**, *193*, 257–273. [[CrossRef](#)] [[PubMed](#)]
75. De Lannoy, G.J.; Reichle, R.H.; Peng, J.; Kerr, Y.; Castro, R.; Kim, E.J.; Liu, Q. Converting between SMOS and SMAP Level-1 Brightness Temperature Observations over Nonfrozen Land. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1908–1912. [[CrossRef](#)]

76. Carrassi, A.; Hamdi, R.; Termonia, P.; Vannitsem, S. Short time augmented extended Kalman filter for soil analysis: A feasibility study. *Atmos. Sci. Lett.* **2012**, *13*, 268–274. [[CrossRef](#)]
77. Lievens, H.; Reichle, R.H.; Liu, Q.; De Lannoy, G.J.; Dunbar, R.S.; Kim, S.B.; Das, N.N.; Cosh, M.; Walker, J.P.; Wagner, W. Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. *Geophys. Res. Lett.* **2017**, *44*, 6145–6153. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



# Paper II

## **Monitoring Soil Moisture Drought over Northern High Latitudes from Space**

**Blyverket, J.;** Hamer, P.D.; Schneider, P.; Albergel, C.; Lahoz, W.A.



*Remote Sensing* 2019, 11(10). [doi.org/10.3390/rs11101200](https://doi.org/10.3390/rs11101200)





Article

# Monitoring Soil Moisture Drought over Northern High Latitudes from Space

Jostein Blyverket <sup>1,2,\*</sup> , Paul D. Hamer <sup>1</sup>, Philipp Schneider <sup>1</sup>, Clément Albergel <sup>3</sup>  and William A. Lahoz <sup>1,†</sup>

<sup>1</sup> Norwegian Institute for Air Research NILU, Instituttveien 18, 2007 Kjeller, Norway; pdh@nilu.no (P.D.H.); philipp.schneider@nilu.no (P.S.); wal@nilu.no (W.A.L.)

<sup>2</sup> Geofysisk Institutt, University of Bergen UiB, Allegaten 70, 5020 Bergen, Norway

<sup>3</sup> CNRM—Université de Toulouse, Météo-France, CNRS, 31057 Toulouse, France; clement.albergel@meteo.fr

\* Correspondence: jb@nilu.no

† Deceased.

Received: 30 March 2019; Accepted: 15 May 2019; Published: 20 May 2019



**Abstract:** Mapping drought from space using, e.g., surface soil moisture (SSM), has become viable in the last decade. However, state of the art SSM retrieval products suffer from very poor coverage over northern latitudes. In this study, we propose an innovative drought indicator with a wider spatial and temporal coverage than that obtained from satellite SSM retrievals. We evaluate passive microwave brightness temperature observations from the Soil Moisture and Ocean Salinity (SMOS) satellite as a surrogate drought metric, and introduce a Standardized Brightness Temperature Index (STBI). We compute the STBI by fitting a Gaussian distribution using monthly brightness temperature data from SMOS; the normal assumption is tested using the Shapiro-Wilk test. Our results indicate that the assumption of normally distributed brightness temperature data is valid at the 0.05 significance level. The STBI is validated against drought indices from a land surface data assimilation system (LDAS-Monde), two satellite derived SSM indices, one from SMOS and one from the ESA CCI soil moisture project and a standardized precipitation index based on in situ data from the European Climate Assessment & Dataset (ECA&D) project. When comparing the temporal dynamics of the STBI to the LDAS-Monde drought index we find that it has equal correlation skill to that of the ESA CCI soil moisture product (0.71). However, in addition the STBI provides improved spatial coverage because no masking has been applied over regions with dense boreal forest. Finally, we evaluate the STBI in a case study of the 2018 Nordic drought. The STBI is found to provide improved spatial and temporal coverage when compared to the drought index created from satellite derived SSM over the Nordic region. Our results indicate that when compared to drought indices from precipitation data and a land data assimilation system, the STBI is qualitatively able to capture the 2018 drought onset, severity and spatial extent. We did see that the STBI was unable to detect the 2018 drought recovery for some areas in the Nordic countries. This false drought detection is likely linked to the recovery of vegetation after the drought, which causes an increase in the passive microwave brightness temperature, hence the STBI shows a dry anomaly instead of normal conditions, as seen for the other drought indices. We argue that the STBI could provide additional information for drought monitoring in regions where the SSM retrieval problem is not well defined. However, it then needs to be accompanied by a vegetation index to account for the recovery of the vegetation which could cause false drought detection.

**Keywords:** SMOS; drought index; summer 2018 drought

## 1. Introduction

Droughts cost society billions of dollars every year, estimates from the World Meteorological Organization (WMO) show that in the European Union alone droughts cost around 6.2 billion USD per year [1]. This is also reflected in research activities on drought, which are rapidly increasing compared to research on other types of natural hazards globally [2]. It is important to implement tools that can monitor and warn about drought conditions, in order to mitigate and prevent losses from droughts [3,4]. Such tools will provide policy and decision makers with a quantitative measure of drought characteristics, allowing them to act upon scientifically based data. Drought indices from different sources, i.e., satellite platforms, models and in situ observations are crucial components of drought monitoring tools. By utilizing information (and creating drought indices) from multiple sources one avoids relying too much on just one source of information and the possible failure of this source to capture the drought.

In the spring and early summer of 2018 severe drought conditions developed in Nordic countries: Norway, Sweden, Finland and Denmark [5,6]. The dry surface conditions set preconditions for wildfires, decreased crop yield and increased crop failure, which resulted in large private and governmental economic losses [6–9]. In Norway alone the preliminary payout from the government to farmers (3 January 2019) reached 187 million USD, compared to 4.9 million USD per year on average for the 2008–2017 period [7]. Late winter and early spring precipitation deficit lead to a decrease in soil moisture (SM), which did not recover until late August and September [8]. For example, the rainfall from May to July in Lund, Sweden, was only about half of the previous low record, with observations dating back to 1748 [6]. Droughts are rare in the Nordic countries, and regional monitoring capabilities and preventive measures were lacking, likely increasing the negative impacts of the drought. Recent studies have found that climate change is likely to exacerbate droughts [10]; as a result, the drought will set in quicker and be more intense [11]. Although the Nordic region is projected to get wetter conditions on average under climate change [12], aridity is expected to increase during the boreal summer months [10], and thus a way of monitoring and mapping droughts over the northern regions is much needed. One way of doing this is by satellite remote sensing [13–15], as satellites could provide near-real-time observations covering large regions within a relative short time-period.

Satellite retrieval of surface soil moisture (SSM) over northern latitudes is difficult because of snow cover, high open water fraction, steep topography and dense boreal vegetation that affect the microwave emissions from the soil [16]. This eventually results in large areas where the retrievals are missing (masked), and hence the spatial and temporal coverage of satellite derived SM over this region is poor. Although the inversion from brightness temperature to SM might be ill posed over northern latitudes, the microwave signal will still carry information about water content in the vegetation (VWC) and soil system [16,17]. Thus, anomalies in the water content of the vegetation-soil system will be reflected in anomalies in the passive microwave brightness temperature. An advantage of using the raw microwave signal is that we do not introduce auxiliary data, as done when for example retrieving SM and vegetation optical depth (VOD). Computation of these variables usually depend on field observations extended to the Global scale, as a result there might be large uncertainties associated with these estimates, propagating into the SSM retrieval.

In this paper, we argue that when studying hydrological extremes, such as drought, we can omit the satellite SM retrieval problem over northern latitudes and look at the raw radiances (microwave brightness temperature,  $T_b$ ) instead. The rationale is that the  $T_b$  is a convolution of SM and VWC [14,18], hence it can be used to map drought (onset, extent and recovery) from space over northern latitudes, a region where SM retrieval products have large spatial and temporal gaps. These spatio-temporal gaps are present because of, for example dense boreal vegetation seen in Finland, Sweden and eastern parts of Norway, and because of frozen soil and snow cover in the Alpine regions in southern Norway and along the coast of northern Norway. Regions with frozen soil and snow cover during the boreal summer months (June, July and August) are limited (and small in extent),

thus we argue that it will not affect our conclusions regarding using the  $T_b$  as a drought index. In this work, we introduce the Standardized Brightness Temperature Index (STBI) for drought monitoring over northern high latitudes. Due to its close relation to SSM we compare and validate the STBI index against two SSM indices and a one month precipitation index. The one month precipitation and SSM indices have been found to closely describe agricultural drought [15,19,20].

This paper is divided into four parts, Section 1 introduces the paper, in Section 2 we present the remote sensing, precipitation and modelling data; we also introduce the methods for the computation of the standardized drought indices. In Section 3.1 we evaluate the temporal dynamics of the STBI index using the Standardized Precipitation Index (SPI) from the gridded E-OBS in situ rainfall dataset, and two Standardized Soil moisture Indices (SSI), one from the National Centre for Meteorological Research (CNRS) Météo-France Land Data Assimilation System Monde (LDAS-Monde), and one from the European Space Agency Climate Change Initiative (ESA CCI) satellite derived SM product. In Section 3.2 a case study of the summer 2018 Nordic drought is used to evaluate the STBI drought monitoring capabilities. Finally, in Section 4 we present our conclusions.

## 2. Data and Methods

### 2.1. Remote Sensing Data

Launched in November 2009 by the European Space Agency (ESA), the Soil Moisture and Ocean Salinity (SMOS) satellite is dedicated to measure passive microwave emissions in the L-band from the Earth surface [16]. We use the SMOS Level-2 SMUDP2 version 650 reprocessed (2010–2017) and the operational (April, May, June, July, August and September 2018) data for both the SM and horizontally polarized brightness temperature ( $T_{bH}$ ). Thus, the SMOS SM and  $T_b$  indices are based on the same baseline climatology. The data are obtained from the ESA SMOS dissemination service [21]. The SMOS retrieval algorithm simultaneously retrieves SM and vegetation optical depth by using information from multi-angle observations of  $T_b$  at horizontal and vertical polarization. The SMOS retrieval is done by minimizing the difference between the satellite observed and model simulated  $T_b$ , using the L-band Microwave Emission of the Biosphere model (L-MEB) [16,22]. The horizontal polarization is chosen because other studies show that it is more sensitive to SSM than the vertical polarization [23]. However, we saw very small to no difference when applying the vertical polarization instead of the horizontal polarization in the computation of the microwave drought index, we therefore only show results for the horizontal polarization.

At L-band the  $T_{bH}$  is sensitive to SM in the upper 0–5 cm of the soil [24], and for very dry and sandy regions even deeper than this [16]. Even though the L-band penetration depth is deeper during very dry conditions a limitation of the satellite derived drought index will be the sensing depth, we are still unable to quantify the amount of water in the root-zone. However, the STBI can provide valuable information on surface water availability when plants develop [15] and for monitoring preconditions of wildfire.

The microwave emissions are larger for a dry soil than for a wet soil [25], and the satellite observed  $T_{bH}$  also depends on the effective soil and canopy temperature [26]. In addition, the  $T_{bH}$  is linked to the VWC; an increase in VWC leads to an increase in the observed brightness temperature [18]. Effectively, this means that under dry vegetation conditions a larger fraction of the observed brightness temperature over vegetated areas will come from the soil, as the vegetation masking of the signal will be smaller than under wet conditions.

The SMOS Level-2 swath data are gridded to the Equal Area Scalable Earth (EASE) version 2.0 36 km grid using a nearest neighbour method; this is done to avoid smoothing from an interpolation scheme. The SMOS data are extracted for the period 1 July 2010 until 1 October 2018 (April, May and June 2010 are not utilized, following [27]). We only use the morning overpass to ensure that the land-atmosphere system is as close as possible to thermal equilibrium. This means that the effective temperature of the

soil-vegetation system is not dominated by one of the compartments (surface or vegetation canopy). Which is important since the  $T_b$  is sensitive to the soil-vegetation effective temperature [16,24].

The  $T_{bH}$  data are screened for values outside a range of 100–320 K [27]. Other than that we do not do any detailed quality control, because part of this work is to see if the SMOS  $T_{bH}$  data contains drought information regardless of grid-cell properties. Monthly  $T_{bH}$  climatology is computed by averaging the  $\sim 6$  a.m. overpasses; this is done for April, May, June, July, August and September from 2010 (except April, May and June 2010) until 2018. Only grid-cells with nine years of data are included in the climatology, except for April, May and June where we use eight years of data.

The monthly satellite derived SM from the ESA CCI SM project is extracted from the Copernicus Climate Change Service (C3S) [28,29]. This is a state-of-the-art satellite derived SSM product, and will therefore act as an important dataset for validation of the STBI index. In addition, by comparing the STBI with the SSI\_ESA\_CCI index we can identify regions where the STBI could provide supplementary information to an index based on satellite derived SSM. We utilize the COMBINED product, which is a combination of SM retrievals from passive and active satellite sensors, such as METOP-A, METOP-B, AMSR2 and SMOS [30]. The COMBINED product is posted on a  $0.25^\circ$  regular longitude/latitude grid. The dataset spans from 1979 until present; however, because of spatial and temporal gaps in the product, we only use data from April 2010 until October 2018 (i.e., the same time-period as the SMOS-L2 product). This also ensures that the climatologies for the standardized indices are computed over the same time-period.

## 2.2. Precipitation Data

In this study, we use the E-OBS version 17.0 precipitation dataset, which corresponds of in situ rain gauge data posted on a  $0.25^\circ$  grid [31]. This dataset has undergone extensive quality control by the data providers. Furthermore, the data have been designed to provide the best estimate of grid-cell averages rather than point values. The data providers computed spatially distributed precipitation values using a three-step interpolation method, first interpolating the monthly precipitation totals, then interpolating daily anomalies using universal kriging, then finally, combining the monthly and daily estimates [31]. By using this three-step method they were able to take into account that the in situ stations are located in different climate zones across Europe. Data for June, July, August and September 2018 are not included in v17.0 and were therefore downloaded separately. The E-OBS dataset spans from 1st January 1950 until 1st October 2018. The one month Standardized Precipitation Index (SPI-1) is applied to create a measure of drought, which is independent from the STBI ( $T_{bH}$ ) data. Accumulated total precipitation for individual months is computed by summarizing daily precipitation (mm/day) for each month separately from 1950 until October 2018.

## 2.3. LDAS-Monde Soil Moisture Data

Soil moisture analysis data are from the Land Data Assimilation System Monde (LDAS-Monde) [32], which has recently been applied to monitor and forecast the impact of the 2018 summer drought on vegetation over central Europe [33]. We run the LDAS-Monde system over the Nordic region using ERA-5 reanalysis atmospheric forcing data and the ISBA (Interaction between Soil Biosphere and Atmosphere) land surface model [34,35] within the SURFEX v.8.1 (SURface EXternalisée) modelling framework [36]. Surface SM derived from the METOP satellite platforms and Leaf Area Index (LAI) observation data from the Copernicus Global Land (CGL) service are assimilated into the LDAS-Monde system using a simplified extended Kalman Filter (SEKF) [37–40]. The LDAS-Monde system is setup at a  $0.25^\circ$  regular longitude/latitude grid. Monthly means for the 2010 to 2018 period are created from the 6 a.m. SSM model data; this is done to correspond as closely as possible with the SMOS overpass time and the  $T_{bH}$  observations.

#### 2.4. Computation of the Standardized microwave Brightness Temperature Index (STBI)

In this section we introduce the new Standardized microwave Brightness Temperature Index (STBI). In this work the STBI is based on SMOS data. However, it can also be estimated based on data from other L-band satellites, for example, the Soil Moisture Active Passive (SMAP) NASA mission [24]. The STBI\_SMOS is computed assuming that the  $Tb_H$  in each grid-cell follows a Gaussian probability distribution. This assumption is tested using the Shapiro-Wilk test, where the null hypothesis is that our sample comes from a normally distributed population. The Shapiro-Wilk test is based on computation of the correlation between the data and the Gaussian quantile function based on their ranks, and it is a powerful test for normality of our data [41]. This test only checks if the data were drawn from a normal distribution, it does not check what the parameters of that distribution might be. Our null-hypothesis is that the data are normally distributed. If the  $p$ -value is smaller than a chosen  $\alpha$  value then the null-hypothesis is rejected and there is evidence that the data are not normally distributed. If the  $p$ -value is larger than the chosen  $\alpha$  value we cannot reject the null-hypothesis that the data are normally distributed, hence the data are likely normally distributed. Here we follow [15,19] and use a significance level of  $\alpha = 0.05$ .

To fit the Gaussian distribution to the  $Tb_H$  data we use the maximum likelihood method (MLM); this is done separately for each grid-cell. By fitting the data grid-cell by grid-cell we are able to take into account the climatology of the individual grid-cells, and therefore it should be more accurate than fitting the data over the whole domain, i.e., using one set of parameters for the whole domain. As an alternative to the MLM we could, for example, have used the method of moments, however, we argue that for this first attempt to compute the STBI it is not crucial for the conclusions of our paper which method we use to fit the data. We do not compute the goodness-of-fit between our data and the parametric distributions, since the parameters have been fitted using the available data [41]. A goodness-of-fit test is then likely to not reject the null hypothesis, i.e., that the observed data were drawn from the distribution being fitted, when in fact it should have been rejected. We compute the PDF of monthly  $Tb_H$  data, for each summer month. By integrating over  $(0, Tb_H^i)$  we find the probability of a given  $Tb_H^i$  value. This value is then converted to a standardized index using:

$$STBI = \Phi^{-1}(p(Tb_H^i)), \quad (1)$$

where  $\Phi^{-1}$  is the inverse standard normal distribution with zero mean and a standard deviation of one. The standardization is based on an approximation detailed in [42].

#### 2.5. Computation of the Standardized Soil moisture Index (SSI)

For comparison to the STBI\_SMOS index we compute three standardized SM indices (SSI\_ESA\_CCI, SSI\_LDAS and SSI\_SMOS), they are computed by assuming a Beta distribution for the underlying SM data [15,19]. For the LDAS data, we average daily SM values at 6 a.m. (which correspond to the SMOS overpass time) into monthly values. We use the monthly ESA CCI data provided by the C3S, to compute the SSI\_ESA\_CCI index. For the SMOS SM product we average available overpasses into a monthly climatology and compute the SSI\_SMOS based on this. The Beta probability distribution is given as:

$$f(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)}, \quad (2)$$

where  $B = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ ,  $\theta$  is the volumetric SM content,  $\Gamma$  is the gamma function,  $\alpha$  and  $\beta$  are shape parameters. First we find the upper and lower limit on SM for each individual grid-cell and month. We assume that the first/last 10% of the sorted SM values are linearly related to their empirical distribution function. After the computation of the upper and lower SM values, we find the Beta distribution shape parameters ( $\alpha$  and  $\beta$ ) using the MLM. We then use Equation (2) to compute the PDF of monthly SM, for each summer month. By integrating over  $(0, \theta)$  we find the probability of a  $\theta$  value.

This value is then used in Equation (1), to find the standardized index (SSI). Negative/positive SSI values are below /above the average climatology of SM and indicate a dry/wet period.

## 2.6. Computation of the Standardized Precipitation Index (SPI)

For the sake of comparison with the land surface drought indices (STBI and SSI) we also compute a Standardized Precipitation Index (SPI). The SPI is frequently used in studies and monitoring of meteorological drought, see for example [3,20]. The SPI is used to characterize droughts at time-scales of 1 to 36 months. On shorter time-scales the SPI is found to be closely related to SM drought, while at longer time-scales it is more closely related to groundwater drought. We therefore follow the approaches in [15,20] where they use the one month SPI for comparison to other SM drought indices. The general interpretation of the SPI is that it expresses the number of standard deviations the anomaly deviates from the long-term mean. In the computation of the SPI-1 we use a non-parametric standardization approach. We apply an empirical relationship because the length of the dataset allows this (69 years), and we avoid assuming one constant parametric distribution function for each grid-cell [42]. It has been found that the SPI can be sensitive to the choice of parametric distribution, and in addition, different recommendations on what parametric distribution to use for modelling precipitation are reported, see [43] and references therein. We also choose to use the whole length of the precipitation data, because we want to study the 2018 drought in a historical perspective. This is one of the limitations of the STBI index, since it only goes back to 2010. The same approach has been applied by [42], where they compute a standardized relative humidity index based on 11 years of satellite data. On such short time-scales the drought index can provide valuable information about current conditions, however, it cannot put the extreme event in a historical perspective. The empirical probabilities of the E-OBS precipitation data are computed for each individual grid-cell, using the empirical Gringorten plotting position [44].

$$p(\text{rainf}) = \frac{i - 0.44}{n + 0.12}, \quad (3)$$

where  $i$  is the rank of the precipitation data from the smallest value, and  $n$  is the sample size. The constants 0.44 and 0.12 are unique for this plotting position. The empirical probabilities are converted to a standardized index using Equation (1).

Negative or positive SPI-1/SSI values indicate a below (dry) or above (wet) average climatology for the precipitation or SM, respectively. For the STBI\_SMOS a high and therefore warm  $Tb_H$  reflects drier conditions, while a low and cold  $Tb_H$  reflects wetter conditions. We therefore multiply the STBI\_SMOS with  $-1$ , for the sake of comparison with the SPI-1 and SSI. An overview of the computed drought indices are given in Table 1.

**Table 1.** Overview of drought indices. Baseline climatology, availability, PDF used for fitting and the resulting index name.

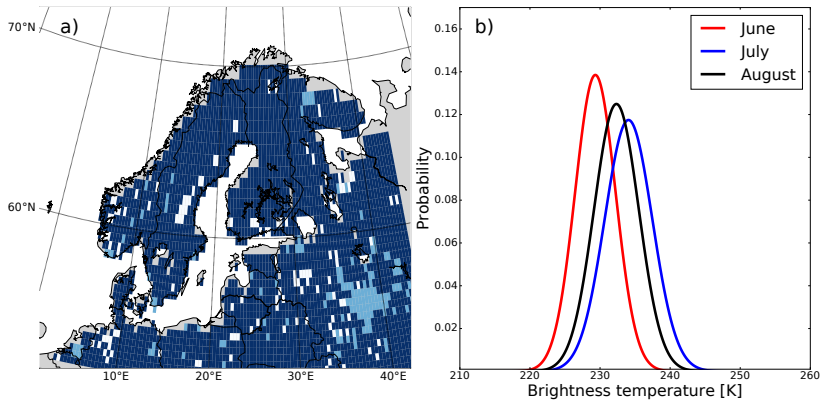
Dataset	Time-Period	Latency	Fitting Distribution	Index
SMOS-L2 Tb	2010–2018	8–12 h	Gaussian	STBI_SMOS
SMOS-L2 SM	2010–2018	8–12 h	Beta	SSL_SMOS
LDAS-Monde SM	2010–2018	n/a	Beta	SSL_LDAS
ESA CCI COMBINED SM	2010–2018	10 days	Beta	SSL_ESA_CCI
E-OBS Precip.	1950–2018	1 month	Empirical	SPI-1

### 3. Results and Discussion

#### 3.1. Evaluation of the Proposed Standardized Microwave Brightness Temperature Index

##### 3.1.1. $T_{bH}$ Probability Distribution

We show the result of the Shapiro-Wilk test on the normality assumption for  $T_{bH}$  during the boreal summer (June, July and August) months. In Figure 1a grid-cells in dark blue show where the null-hypothesis was not rejected, light blue grid-cells show where the null-hypothesis was rejected. White regions over land show where we had less than eight years of data for the Shapiro-Wilk test, these regions are excluded in the calculation.



**Figure 1.** (a) Shapiro-Wilk test for normality of the  $T_{bH}$  distribution shown for July. Dark blue regions null-hypothesis is not rejected, i.e., the data appears to be normally distributed. Light blue regions null-hypothesis is rejected, and white regions (over land) had too few years of data for testing. (b) Example PDFs of the fitted brightness temperature for boreal summer months, June (red), July (blue) and August (black) for the grid-cell covering  $63.7^\circ$  N and  $12.0^\circ$  E.

The Gaussian fit to the  $T_{bH}$  for June (red), July (blue) and August (black) are shown for one grid-cell ( $63.7^\circ$  N and  $12.0^\circ$  E) in Figure 1 b). The distributions show that June has a lower mean  $T_{bH}$  than July and August, with July being on average the warmest. The drought index value is computed by integrating the PDFs over  $(0, T_{bH})$ . The integral is approximated by a summation up to the  $T_{bH}$  value of interest.

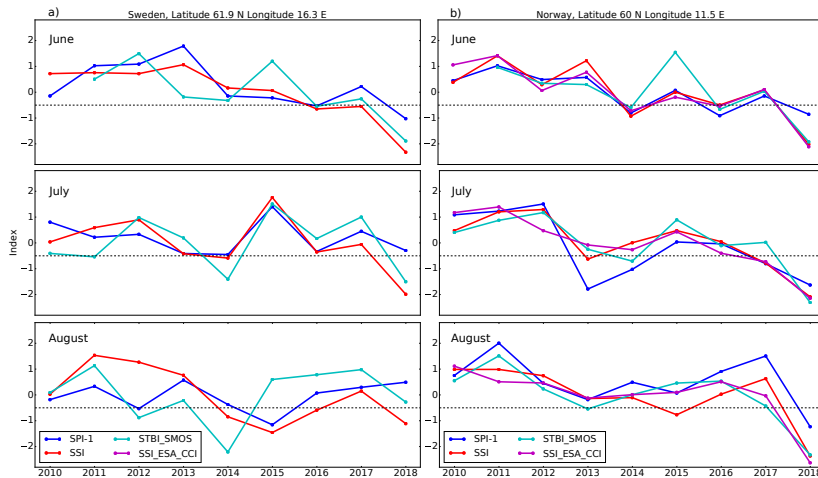
##### 3.1.2. Temporal and Spatial Patterns of the Drought Indices

Figure 2 shows time-series of the SPI-1 (blue), STBI\_SMOS (cyan), SSI\_LDAS (red) and SSI\_ESA\_CCI (magenta) over a grid-cell in Sweden ( $61.9^\circ$  N and  $16.3^\circ$  E) (a) and Norway ( $60.0^\circ$  N and  $11.5^\circ$  E) (b) from 2010 until 2018 for the boreal summer months, June, July and August. The two regions were selected to represent a region with and without the SSI\_ESA\_CCI data, and therefore show how the STBI\_SMOS can represent regions where SM retrievals are masked. Furthermore, these two regions were affected by the 2018 summer drought, as seen from the negative anomalies in the indices for 2018.

Depending on the severity, a drought can be classified into a drought scale or D-scale, detailed explanation of this scale is provided in [3], see their Table 1. In this classification an SSI below 0.5 is defined as being abnormally dry, and corresponding to a 30% probability of occurrence. The D0 value as described in [3] is the lowest on the drought scale ranging from D0 to D4, where D1 is a moderate drought and D4 is described as an exceptional drought. Following [42], we use the D0 as a drought onset threshold, we see that for the regions in Figure 2a,b, severe drought conditions (see 2018 summer) is captured by the STBI\_SMOS. The STBI\_SMOS does not only capture dry events, it also captures



years where a month is wetter than normal (index larger than zero). However, for positive index anomalies there seems to be more false events (e.g., June 2015 in Sweden, and June 2015 in Norway) than for the dry events.

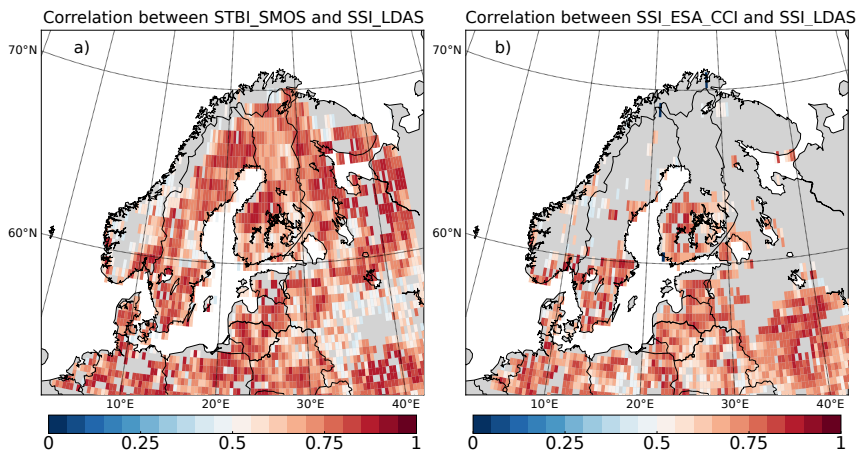


**Figure 2.** (a) Time-series of the SPI-1 (blue), STBI\_SMOS (cyan) and SSI\_LDAS (red) over a grid-cell in Sweden ( $61.9^{\circ}$  N and  $16.3^{\circ}$  E) from 2010 until 2018 for the boreal summer months June, July and August. The horizontal black dotted line indicate D0 drought conditions, see text for further explanation. Please note that for this region there were no observations for the computation of the SSI\_ESA\_CCI. (b) Same as a) but for a grid-cell in Norway ( $60.0^{\circ}$  N and  $11.5^{\circ}$  E), here observations from the ESA CCI were available for the computation of the SSI\_ESA\_CCI drought index. For the different indices the grid-cell sizes are:  $0.36^{\circ}$  for the STBI\_SMOS and  $0.25^{\circ}$  for the SPI-1, SSI\_LDAS and SSI\_ESA\_CCI

We evaluate how well the STBI\_SMOS, SSI\_ESA\_CCI and SPI-1 could capture the temporal dynamics of the SM drought by computing the correlation coefficients between the SSI\_LDAS and the other metrics (STBI\_SMOS, SSI\_ESA\_CCI and SPI-1). The LDAS-Monde index is then used as the reference index. This is justified by the fact that it incorporates both model and observation data in a data assimilation system. Other studies have shown that land data assimilation systems are able to correct for errors in precipitation datasets, and as a result, improve the representation of SSM (see for example [45]). Another example is provided by [46], where the authors show that the LDAS-Monde improves the representation of the 2012 US corn belt drought. We do not claim that the LDAS-Monde based data are superior to the other sources of data over this region. However, because of its use of state-of-the-art atmospheric forcing from ERA-5 and by combining SSM and LAI from satellites with model derived SSM and LAI it should be better than comparing with a model alone or to the satellite observations alone. It is also the best we can do for providing an independent reference index with a spatial scale close to that of the satellites.

In the computation of the correlation coefficient, we used June, July and August (boreal summer) data together from 2010 until 2018 to increase the number of data-points. The domain average was only computed for grid-cells where the Pearson correlation value was statistically significant at the 0.05 level. Figure 3a shows the Pearson correlation coefficient between the STBI\_SMOS and SSI\_LDAS, regions with no values (grey regions over land) had a correlation not different from zero at the 0.05 significance level. Most of the domain has a high correlation, except regions in south central Norway and from mid-Norway to northern Norway. In Figure 3b the Pearson correlation coefficient between the SSI\_ESA\_CCI and SSI\_LDAS is shown. Here large regions in the Nordic countries (Norway, Sweden and Finland) have non significant correlation. This is likely because the SSI\_ESA\_CCI index has large regions with missing data for the individual months, thus resulting in

a non-significant correlation and discarded values (grey regions in Figure 3). Summary statistics for the spatial correlations are shown in Table 2.



**Figure 3.** (a) Pearson correlation coefficient between the STBI\_SMOS and SSI\_LDAS, red regions indicate a high positive correlation, masked (grey) regions have a correlation not significantly different from zero. (b) Same as a) but for the Pearson correlation between the SSI\_ESA\_CCI and SSI\_LDAS

The STBI\_SMOS has a correlation with the SSI\_LDAS of 0.71, it also has the highest number of grid-cells with a statistically significant correlation value ( $N = 2437$  out of 2997 grid-cells (81 %)). The SSI\_ESA\_CCI index had a correlation of 0.70 with the SSI\_LDAS index, and significant values in  $N = 1523$  out of 2997 (~ 51 %) grid-cells. Finally, the SPI-1 correlation with SSI\_LDAS was 0.56 for  $N = 1537$  out of 2997 (~ 51 %) grid-cells. The high correlation between the STBI\_SMOS and the SSI\_LDAS indicate that the STBI\_SMOS is able to capture the variability in the SM over the Nordic region as good as the SSI\_ESA\_CCI index. The number of grid-cells with statistically significant correlation values are higher for the STBI\_SMOS than for the SSI\_ESA\_CCI, hence it provides better spatial coverage than the satellite derived SM index. To check that the high-correlation is not only found in regions where the SSI\_ESA\_CCI data were missing, we also compute the correlation for grid-cells covered by both products, see Table 2. Here the mean correlation is only taken for grid-cells where we have data for all the four indices (resulting in 800 of 2997 land grid-cells being covered, i.e., 27 %).

**Table 2.** Pearson R correlation coefficient between the SSI\_LDAS and the, STBI\_SMOS, SSI\_ESA\_CCI and the SPI-1. Computed for individual grid-cells (27 datapoints) and summed over the whole domain (all columns) and over grid-cells with overlap between all datasets (overlap columns).  $N$  indicates grid-cells with statistically significant correlation at the 0.05 level, total number of land grid-cells were 2997.

Index	All		Overlap	
	R	N	R	N
STBI_SMOS	0.71	2437	0.70	800
SSI_ESA_CCI	0.70	1523	0.70	800
SPI-1	0.56	1537	0.56	800

### 3.2. Case Study of the Summer 2018 Drought

To further evaluate the performance of the STBI for drought mapping we utilized the 2018 summer drought over the Nordic countries as a case study.

### 3.2.1. Comparison between the STBI\_SMOS, SSI\_LDAS, SSI\_ESA\_CCI, SPI-1 and SSI\_SMOS

The limited number of reliable satellite derived SM observations in the SMOS-L2 (Figure 4q–t) and ESA CCI COMBINED product (Figure 4i–l) motivated our attempt to describe the 2018 Nordic drought using the observed brightness temperature ( $T_{bH}$ ). In addition to the poor coverage, the Standardized Soil moisture Index for SMOS (SSI\_SMOS) exhibits noisy patterns, and little resemblance to the SSI\_LDAS in Figure 4. Comparing the STBI in Figure 4a–d) to the SSI\_ESA\_CCI in Figure 4i–l) we see that the STBI\_SMOS has a better spatial coverage than the SSI\_ESA\_CCI. Large regions over Sweden and northern Finland are not covered by the SSI\_ESA\_CCI. This problem is addressed by using the  $T_{bH}$  data.

Figure 4a,e,i,m,q show the STBI\_SMOS, SSI\_LDAS, SSI\_ESA\_CCI, SPI-1 and SSI\_SMOS for May 2018 over the Nordic region. The SPI-1 indicates a precipitation deficit for most of the domain. The STBI\_SMOS, SSI\_LDAS and SSI\_ESA\_CCI show that northern parts of Norway and the mountain regions in the south of Norway are wetter (colder for the STBI) than usual. This signal might come from late snowmelt wetting the soil in the northern latitudes and the mountainous regions in southern Norway. We also note that southern parts of Sweden and Finland are drier (warmer) than usual for the STBI\_SMOS, SSI\_LDAS and SSI\_ESA\_CCI. In general, the spatial patterns for May are very similar for the STBI and SSI\_LDAS, although the STBI overestimates the wet regions in northern Norway and in Finland.

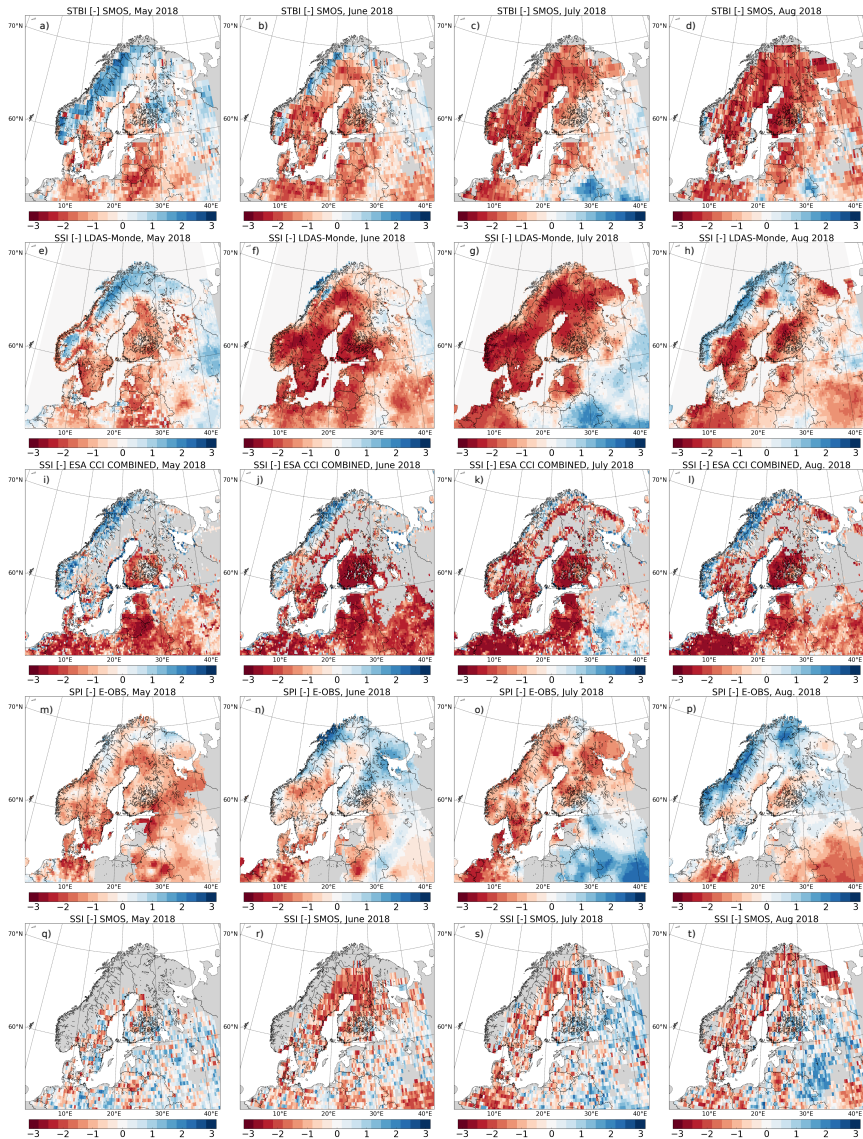
Next we examine the indices during June 2018. Figure 4b,f,j,n,r show the STBI\_SMOS, SSI\_LDAS, SSI\_ESA\_CCI, SPI-1 and SSI\_SMOS, respectively. The dry conditions seen in the SPI-1 continue in eastern Norway, southern Sweden and Denmark. Northern parts of Norway and most of Finland experience rainy conditions seen from the SPI-1. Eastern Norway, Sweden, Finland, Denmark and the Baltic countries have a dry anomaly in the STBI\_SMOS, SSI\_LDAS and SSI\_ESA\_CCI. Northern parts of Sweden and Finland have missing values for the SSI\_ESA\_CCI index; however, the STBI\_SMOS shows similar patterns as the SSI\_LDAS, except for the wet regions in southwestern and northern Norway. When comparing the STBI\_SMOS to the SSI\_LDAS we see that the STBI\_SMOS captures the wet (cold) regions in the east of the domain. The reason why the STBI is able to capture wet/dry anomalies is because of its close relationship to SSM. This is why L-band radiometers are used for SM retrievals in the first place. The large difference in dielectric constant between water ( $\sim 80$ ) and dry soil ( $\sim 3$ ), results in a large range in the emissivity of the soil, and therefore also the soil brightness temperature [47]. Low/high brightness temperature indicates that the surface is wet/dry. The STBI can therefore also detect wet anomalies (such as water ponding on the surface), which can be useful in monitoring flood and/or heavy precipitation events. Typically, for heavy precipitation events where you get water ponding on the surface, the SM retrieval does not make any sense, and the retrieval is omitted [16].

In July in Figure 4c,g,k,o,s the SPI-1 shows a dry anomaly for Norway, Sweden, Finland and Denmark. In July, drought conditions were dominant over most of the domain, except for regions in the south central and east, which is reflected in all of the indices. Again, there are gaps in the SSI\_ESA\_CCI over large regions of Sweden and Finland. These gaps are not present in the STBI\_SMOS, which is consistent with the SSI\_LDAS, showing dry anomalies for this region. Close to normal conditions in northern parts of Poland are found for both the STBI\_SMOS and the SSI\_LDAS for July; this is not seen for the SSI\_ESA\_CCI.

In August most of Norway experienced wetter than usual conditions (seen from the SPI-1), this is reflected in the SSI\_ESA\_CCI and the SSI\_LDAS, but not in the STBI\_SMOS, see Figure 4d,h,l,p. One reason for this could be precipitation intercepted by the vegetation, increasing the VWC, again increasing the emissivity from the vegetation [18]. Higher emissivity from a recovering vegetation could therefore mask out precipitation events and cause a false drought signal, as seen in August 2018. We see this as an argument for simultaneously retrieving SSM and VOD, as applying a constant (in time) VOD value would likely cause the SSM retrieval to be drier than expected because of the increased  $T_b$  from the vegetation.

The SPI-1, SSI\_LDAS and SSI\_ESA\_CCI are all on the same spatial scale ( $\sim 25$  km), however the STBI\_SMOS is posted on a coarser ( $\sim 36$  km) grid. In Figure 2 we plotted the time-series of the grid-cell closest to the given latitude/longitude, and despite the spatial mismatch between the STBI and the other indices, we see that there is a good temporal agreement which is also reflected in the Pearson correlation

coefficient in Table 2. Droughts are by nature large scale events (as seen in Figure 4), thus a small spatial mismatch between our indices should not alter our conclusions on the spatial patterns of the drought.



**Figure 4.** Drought indices, blue/red is above/below average precipitation,  $Tb_H$  or soil moisture. Grey colour indicates regions without data. Columns from left to right are for May, June, July and August. (a–d) Standardized microwave Brightness Temperature Index (STBI\_SMOS). (e–h) Standardized Soil moisture Index (SSI\_LDAS). (i–l) Standardized Soil moisture Index ESA CCI (SSI\_ESA\_CCI). (m–p) Standardized Precipitation Index (SPI-1). (q–t) Standardized Soil moisture Index SMOS (SSI\_SMOS).

### 3.2.2. Drought Severity

Following the D-scale [3], an SSI below  $-1.3$  is defined as a severe drought (D2 conditions). In Figure 5 we have plotted the STBI\_SMOS (a–d), SSI\_LDAS (e–h), SSI\_ESA\_CCI (i–l) and SPI-1 (m–p) for D2 conditions for May, June, July and August in 2018. We have not included the SSI\_SMOS in Figure 5, because from Figure 4q–t we saw that the SMOS SM data were noisy and had large spatial gaps. The difference between the one month land surface indices (STBI\_SMOS, SSI\_LDAS and SSI\_ESA\_CCI) and the SPI-1 is most likely due to the lag time between the meteorological drought (one month SPI) and the agricultural drought (one month SSI). The SPI-1 has a shorter memory than the SSI, hence a dry SPI-1 in month  $i$  is often followed by a dry SSI in month  $i + 1$ , even though the precipitation is back to normal conditions in month  $i + 1$ . Using the SSI\_LDAS as a reference we see that the STBI\_SMOS is able to capture regions in severe drought where the SSI\_ESA\_CCI has masked values from the retrieval. This can be seen in southern Norway (July) and northern and central Sweden (July). Comparing Figure 5c,g,k we see that northern parts of Poland do not have severe drought conditions for the SSI\_LDAS, and this is captured by the STBI\_SMOS but not by the SSI\_ESA\_CCI. On the other hand, the spatial pattern of the SSI\_ESA\_CCI drought severity in June has better agreement with the SSI\_LDAS than the comparison of the STBI\_SMOS versus the SSI\_LDAS. In August, the STBI\_SMOS (Figure 2d) is overestimating the regions experiencing severe drought conditions in northern Norway and Sweden, when compared to the SSI\_LDAS (Figure 5h)).

### 3.2.3. Drought Onset and Recovery

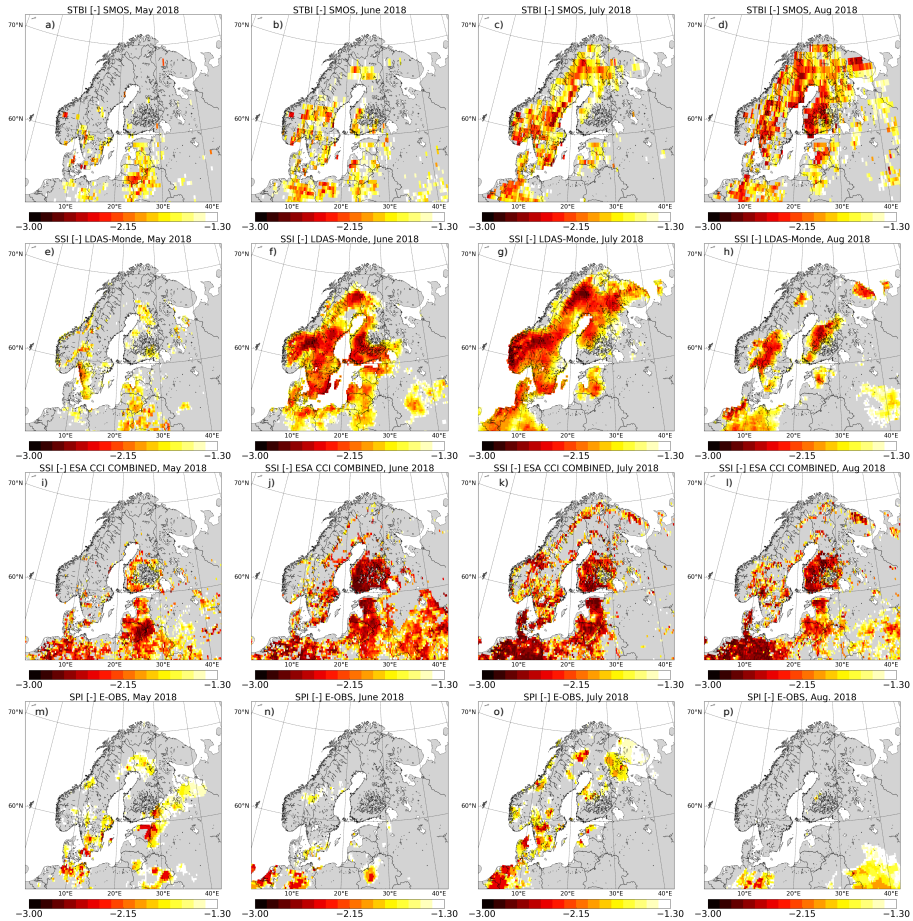
Accurate monitoring of drought onset and recovery could help farmers and decision makers minimize the negative impacts of a drought. Here we evaluate the temporal evolution of the STBI\_SMOS index against the temporal evolution of the SSI\_LDAS SSI\_ESA\_CCI and SPI-1 during the 2018 summer drought. As a consequence of the drought several regions in the Nordic countries experienced wildfires and agricultural losses [8,48], here we have chosen three sites to represent such conditions. In Figure 6 we have selected grid-cells in the vicinity of (a) Jokkmokk municipality, Sweden, (b) Tovaasen, Sweden and (c) Nes in Akershus municipality, Norway. These regions experienced large wildfires and agricultural droughts during the summer 2018 heatwave [6–8,48]. The horizontal dotted black line shows the D0 condition (moderate drought). The first thing to note is that in Figure 6a,b there are no data for the SSI\_ESA because these grid locations are flagged in the retrieval algorithm. This limits the use of the SSI\_ESA\_CCI over regions in the Nordic countries for drought monitoring and mapping. Hence a reason for choosing grid-cells where we have no SSI\_ESA\_CCI data is to show that the STBI\_SMOS can be used to monitor the drought in these regions.

The Jokkmokk municipality lies above the Arctic circle in northern Sweden and it experienced large wildfires during the 2018 summer. In Figure 6a we see that the precipitation deficit (low SPI-1) starting in May causes the STBI\_SMOS (cyan) and SSI\_LDAS (red) to fall below D0 conditions in June. The close to normal SPI-1 conditions in June has little impact on the land surface indices (STBI\_SMOS and SSI\_LDAS). Precipitation deficit in July and only close to normal SPI-1 conditions in August and September, results in a slow recovery of the land surface indices for the Jokkmokk site.

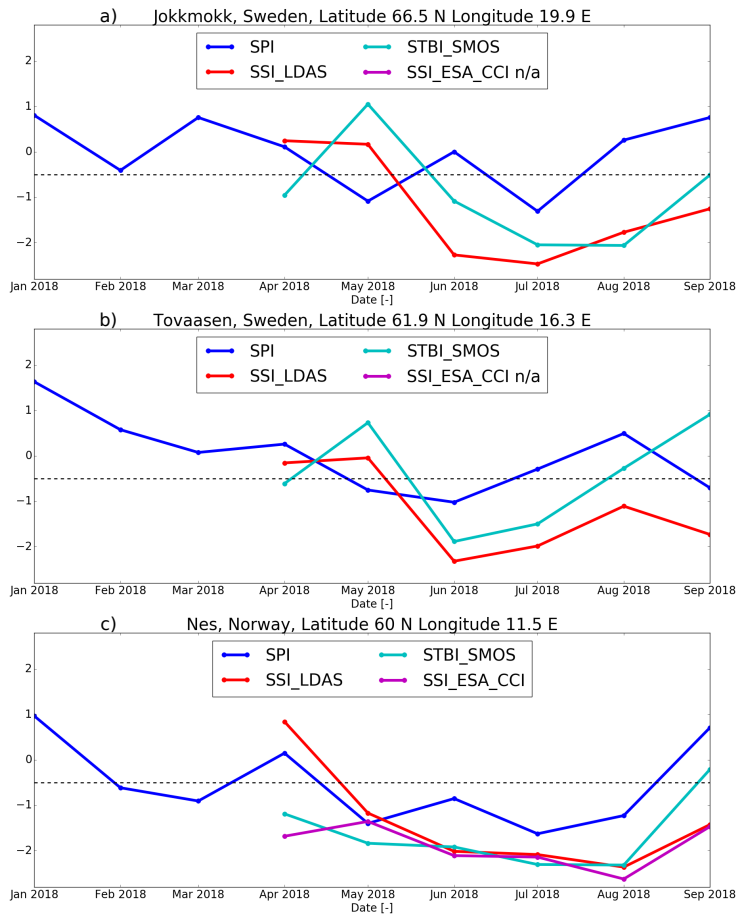
Tovaasen lies in the Ljusdalen municipality, a region in Sweden which experienced large wildfires in mid-July 2018. In Figure 6b we see that the SPI-1 (blue) is close to normal for February, March and April. In May and June the precipitation deficit leads to a decrease in the STBI\_SMOS (cyan) and SSI\_LDAS (red). In August the SPI-1 is close to normal conditions, but this is not enough for the STBI\_SMOS and SSI\_LDAS to recover. In September the STBI\_SMOS and the SSI\_LDAS diverges, with the STBI\_SMOS showing drought recovery while the SSI\_LDAS more closely follows the SPI-1 and shows drought conditions.

Much of the agriculture in Norway lies in the south-eastern parts of the country and here we choose a grid-cell which covers Nes in Akershus municipality. In Figure 6c we see that low SPI-1 conditions in the February and March likely caused abnormally dry (warm) conditions in April for the STBI\_SMOS (cyan) and SSI\_ESA\_CCI (magenta). The continued precipitation deficit in May,

June and July was propagated into the land seen by the low STBI\_SMOS, SSI\_LDAS and SSI\_ESA\_CCI. Here the three land surface drought indices follow each other closely during the dry spell in May, June, July and August.



**Figure 5.** Regions with severe drought conditions, drought index  $< -1.3$  for May, June, July and August. Red regions are severe drought conditions, while yellow and white are regions with less severe drought conditions. Grey colour indicates regions where the drought indices are larger than  $-1.3$ . (a–d) STBI  $< -1.3$ . (e–h) SSI\_LDAS  $< -1.3$ . (i–l) SSI\_ESA\_CCI  $< -1.3$ . (m–p) SPI-1  $< -1.3$ .



**Figure 6.** SPI-1 (blue), SSI\_LDAS (red), STBI\_SMOS (cyan) and SSI\_ESA\_CCI (magenta) time series. Black dotted horizontal line indicate D0 drought conditions. (a) Jokkmokk municipality, Sweden, (b) Tovaasen, Sweden, (c) Nes in Akershus municipality, Norway. Latitudes are given in degrees north (N) and longitudes are given in degrees east (E).

The case study for the 2018 summer show that the STBI\_SMOS has potential to supplement information to drought monitoring over the Nordic region. In particular, we see that it was able to monitor the drought in regions where data from the SM retrievals were missing. The STBI\_SMOS did however miss the transition to a wet anomaly for large regions in Norway in August 2018 (Figure 4d).

#### 4. Conclusions

In this study, we outlined a new approach for directly applying passive microwave brightness temperature to monitor and map drought over the Nordic countries. We propose a standardized index (STBI) based on passive microwave brightness temperature data. The rationale behind this choice is that the  $Tb_H$  convolves information about soil moisture, soil temperature and vegetation water content, which are all important factors in drought and flood monitoring. The brightness temperature also provides a better spatial and temporal coverage than the retrieved soil moisture, because we avoid the retrieval problem, which is problematic over northern latitudes owing to dense vegetation, strong topography, high water fraction and snow cover. The brightness temperature is also available

earlier than than the retrieved soil moisture, which will benefit the drought monitoring capabilities of the index.

We found that the STBI\_SMOS metric was able to capture the spatial patterns of the 2018 Nordic drought, especially for the very dry conditions seen in July 2018, when comparing it to the SSI from LDAS-Monde. As seen for two test sites in Sweden and one in Norway, the STBI\_SMOS drought onset and end were in line with the SSI\_LDAS and SPI-1. The STBI\_SMOS was also characterized by a one month lag compared to the SPI-1 (as often seen in land surface drought metrics [19]), indicating that it contained information about soil/vegetation moisture, and not only about land surface temperature. One criticism of the index is that it did not capture the drought end/recovery seen in August 2018 for our case study. This is likely linked to the fact that when the vegetation recovers from drought the VWC increases resulting in an increase of Tb [18]. We therefore stress that the index should be used as a supplement to other monitoring tools, and not as a replacement.

The results from this work show that passive microwave observations (in the L-band) could be implemented in a Nordic drought monitoring system. We expect that the STBI could be a supplement to modelling tools, and that downscaling of the index would enhance its applicability for drought monitoring at decision making scales. In the future it would be possible to calculate the STBI for observations from more recently launched L-band satellites, such as the SMAP NASA mission [24]. The performance of passive microwave observations in the C-band should also be investigated for drought monitoring over northern latitudes because the temporal span of these missions are longer than the L-band missions, and hence a more reliable estimate of the ( $T_{bH}$ ) climatology can be computed. The method could also be expanded to other regions of the world, where retrieval of soil moisture is difficult. In addition, to fully investigate the advantages of the STBI future work should consider computing the STBI over regions where we know that the SM retrieval is better understood. This study was also the first attempt to monitor agricultural drought over the Nordic region from space and compare the skill of a space based drought index with that of a state-of-the-art land surface data assimilation system (LDAS-Monde). We expect that future development of the STBI\_SMOS metric could benefit farmers, decision makers and others depending on information concerning agricultural drought over the Nordic countries.

**Author Contributions:** J.B., P.D.H., W.A.L., C.A. and P.S. conceived and designed the study. C.A. designed and ran the land surface data assimilation system. J.B. did the analysis and wrote the manuscript. All authors provided comments on the manuscript.

**Funding:** The main author is supported by the Research Council of Norway (NFR PhD-grant 239947, 2015–2019).

**Acknowledgments:** The corresponding author would like to dedicate this paper to the memory of William A. Lahoz. The corresponding author would also like to thank Laurent Bertino for helpful comments in the preparation of this paper. The authors would like to thank the Copernicus Global Land Service for providing the satellite-derived LAI products and soil moisture for the LDAS-Monde assimilation. We would also like to thank the Copernicus Climate Change Service (C3S) for providing the ESA CCI COMBINED soil moisture data, and the ERA-5 data used for atmospheric forcing in the LDAS-Monde system. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (<http://ensembles-eu.metoffice.com>) and the data providers in the ECAD project (<http://www.ecad.eu>). The SMOS data are available online from ESA. The Copernicus Climate Change soil moisture data are available at <https://cds.climate.copernicus.eu>. The LDAS-Monde derived data and analysis scripts can be accessed at <ftp://ftp.nilu.no/Pub/nilu/jostein/Data/>. Finally, we would like to thank the reviewers for their comments on the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gerber, N.; Mirzabaev, A. *Benefits of Action and Costs of Inaction: Drought Mitigation and Preparedness—A Literature Review*; Technical report; World Meteorological Organization (WMO) and Global Water Partnership (GWP): Geneva, Switzerland; Stockholm, Sweden, 2017.
2. Emmer, A. Geographies and Scientometrics of Research on Natural Hazards. *Geosciences* **2018**, *8*, 382. [[CrossRef](#)]
3. Svoboda, M. Drought Monitor. *Bull. Am. Meteorol. Soc.* **2002**, *83*, 1181–1190. [[CrossRef](#)]



4. Luo, L.; Wood, E.F. Monitoring and predicting the 2007 U.S. drought. *Geophys. Res. Lett.* **2007**, *34*. [[CrossRef](#)]
5. Lentze, G. Newsletter No. 157—Autumn 2018. ECMWF Newsletter. 2018. Available online: <https://www.ecmwf.int/node/18705> (accessed on 19 May 2019).
6. WMO. *The State of the Global Climate in 2018*; Technical Report; World Meteorological Organization: Geneva, Switzerland, 2018. Available online: <https://public.wmo.int/en/our-mandate/climate/wmo-statement-state-of-global-climate> (accessed on 19 May 2019).
7. Landbruksdirektoratet. Available online: <https://www.landbruksdirektoratet.no/no/statistikk/landbrukserstatning/klimarelaterte-skader-og-tap/avlingssvikt> (accessed on 28 March 2019).
8. Skaland, R.G.; Colleuille, H.; Andersen, A.S.H.; Mamen, J.; Grinde, L.; Tajet, H.T.T.; Lundstad, E.; Sidselrud, L.F.; Tunheim, K.; Hanssen-Bauer, I.; et al. *Tørkesommeren 2018*; Technical Report; The Norwegian Meteorological Institute: Oslo, Norway, 2019. Available online: [https://fido.nrk.no/ccfcfb66f38035154dd25ba51c2573ae231d397583bee2a4e545ae0b6e3fc2dd/Tørkesommeren%2018\\_\\_.pdf](https://fido.nrk.no/ccfcfb66f38035154dd25ba51c2573ae231d397583bee2a4e545ae0b6e3fc2dd/Tørkesommeren%2018__.pdf) (accessed on 19 May 2019).
9. Liberto, T.D. A Hot, Dry Summer Has Led to Drought in Europe in 2018. Available online: <https://www.climate.gov/news-features/event-tracker/hot-dry-summer-has-led-drought-europe-2018> (accessed on 19 May 2019).
10. Samaniego, L.; Thober, S.; Kumar, R.; Wanders, N.; Rakovec, O.; Pan, M.; Zink, J.; Sheffield, E.; Wood, E.F.; Marx, A. Anthropogenic warming exacerbates European soil moisture droughts. *Nat. Clim. Chang.* **2018**, *8*, 421–426. [[CrossRef](#)]
11. Trenberth, K.E.; Dai, A.; Van Der Schrier, G.; Jones, P.D.; Barichivich, J.; Briffa, K.R.; Sheffield, J. Global warming and changes in drought. *Nat. Clim. Chang.* **2014**, *4*, 17–22. [[CrossRef](#)]
12. Greve, P.; Gudmundsson, L.; Seneviratne, S.I. Regional scaling of annual mean precipitation and water availability with global temperature change. *Earth Syst. Dyn.* **2018**, *9*, 227–240. [[CrossRef](#)]
13. Mu, Q.; Zhao, M.; Kimball, J.S.; McDowell, N.G.; Running, S.W. A remotely sensed global terrestrial drought severity index. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 83–98. [[CrossRef](#)]
14. AghaKouchak, A.; Farahmand, A.; Melton, F.S.; Teixeira, J.; Anderson, M.C.; Wardlow, B.D.; Hain, C.R. Remote sensing of drought: Progress, challenges and opportunities. *Rev. Geophys.* **2015**, *53*. [[CrossRef](#)]
15. Sadri, S.; Wood, E.F.; Pan, M. Developing a drought-monitoring index for the contiguous US using SMAP. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 6611–6626. [[CrossRef](#)]
16. Kerr, Y.H.; Waldteufel, P.; Richaume, P.; Wigneron, J.P.; Ferrazzoli, P.; Mahmoodi, A.; Al Bitar, A.; Cabot, F.; Gruhier, C.; Juglea, S.E.; et al. The SMOS soil moisture retrieval algorithm. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1384–1403. [[CrossRef](#)]
17. Fernandez-Moran, R.; Al-Yaari, A.; Mialon, A.; Mahmoodi, A.; Al Bitar, A.; De Lannoy, G.; Rodriguez-Fernandez, N.; Lopez-Baeza, E.; Kerr, Y.; Wigneron, J.-P. SMOS-IC: An alternative SMOS soil moisture and vegetation optical depth product. *Remote Sens.* **2017**, *9*, 457. [[CrossRef](#)]
18. Jones, A.S.; Vukićević, T.; Vonder Haar, T.H. A Microwave Satellite Observational Operator for Variational Data Assimilation of Soil Moisture. *J. Hydrometeorol.* **2004**, *5*, 213–229. [[CrossRef](#)]
19. Sheffield, J.; Goteti, G.; Wen, F.; Wood, E.F. A simulated soil moisture based drought analysis for the United States. *J. Geophys. Res. D Atmos.* **2004**, *109*, 1–19. [[CrossRef](#)]
20. Yuan, X.; Ma, Z.; Pan, M.; Shi, C. Microwave remote sensing of short-term droughts during crop growing seasons. *Geophys. Res. Lett.* **2015**, *42*, 4394–4401. [[CrossRef](#)]
21. ESA. ESA SMOS Online Dissemination Service. Available online: <https://smos-diss.eo.esa.int/oads/access/> (accessed on 19 May 2019).
22. Wigneron, J.P.; Kerr, Y.; Waldteufel, P.; Saleh, K.; Escorihuela, M.J.; Richaume, P.; Ferrazzoli, P.; de Rosnay, P.; Gurney, R.; Calvet, J.C.; et al. L-band Microwave Emission of the Biosphere (L-MEB) Model: Description and calibration against experimental data sets over crop fields. *Remote Sens. Environ.* **2007**, *107*, 639–655. [[CrossRef](#)]
23. Njoku, E.G.; Wilson, W.J.; Yueh, S.H.; Dinardo, S.J.; Li, F.K.; Jackson, T.J.; Lakshmi, V.; Bolten, J. Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2659–2673. [[CrossRef](#)]
24. Entekhabi, D.; Njoku, E.G.; O'Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; et al. The soil moisture active passive (SMAP) mission. *Proc. IEEE* **2010**, *98*, 704–716. [[CrossRef](#)]

25. De Rosnay, P.; Drusch, M.; Boone, A.; Balsamo, G.; Decharme, B.; Harris, P.; Kerr, Y.; Pellarin, T.; Polcher, J.; Wigneron, J.-P. AMMA land surface model intercomparison experiment coupled to the community microwave emission model: ALMIP-MEM. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]
26. Drusch, M.; Holmes, T.; de Rosnay, P.; Balsamo, G. Comparing ERA-40-Based L-Band Brightness Temperatures with Skylab Observations: A Calibration/Validation Study Using the Community Microwave Emission Model. *J. Hydrometeorol.* **2009**, *10*, 213–226. [CrossRef]
27. De Lannoy, G.J.; Reichle, R.H. Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 4895–4911. [CrossRef]
28. Dorigo, W.; Gruber, A.; De Jeu, R.A.; Wagner, W.; Stacke, T.; Loew, A.; Albergel, C.; Brocca, L.; Chung, D.; Parinussa, R.M.; et al. Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sens. Environ.* **2015**, *162*, 380–395. [CrossRef]
29. Dorigo, W.; Wagner, W.; Albergel, C.; Albrecht, F.; Balsamo, G.; Brocca, L.; Chung, D.; Ertl, M.; Forkel, M.; Gruber, A.; et al. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sens. Environ.* **2017**, *203*, 185–215. [CrossRef]
30. Scanlon, T.; Chung, D.; Paulik, C.; Kidd, R. *Product User Guide and Specification*; Technical Report; ECMWF: Reading, UK, 2018. Available online: <http://datastore.copernicus-climate.eu/c3s/published-forms/c3sprod/satellite-soil-moisture/product-user-guide-v2.3.pdf> (accessed on 19 May 2019).
31. Haylock, M.R.; Hofstra, N.; Klein Tank, A.M.G.; Klok, E.J.; Jones, P.D.; New, M. A European daily high-resolution gridded dataset of surface temperature and precipitation. *J. Geophys. Res. Atmos.* **2008**, *113*, D20119. [CrossRef]
32. Albergel, C.; Munier, S.; Leroux, J.D.; Dewaele, H.; Fairbairn, D.; Lavinia Barbu, A.; Gelati, E.; Dorigo, W.; Faroux, S.; Meurey, C.; et al. 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.* **2017**, *10*, 3889–3912. [CrossRef]
33. Albergel, C.; Dutra, E.; Bonan, B.; Zheng, Y.; Munier, S.; Balsamo, G.; de Rosnay, P.; Muñoz-Sabater, J.; Calvet, J.C. Monitoring and Forecasting the Impact of the 2018 Summer Heatwave on Vegetation. *Remote Sens.* **2019**, *11*, 520. [CrossRef]
34. Noilhan, J.; Mahfouf, J.F. The ISBA land surface parameterisation scheme. *Glob. Planet. Chang.* **1996**, *13*, 145–159. [CrossRef]
35. Calvet, J.C.; Noilhan, J.; Roujean, J.L.; Bessemoulin, P.; Cabelguenne, M.; Olioso, A.; Wigneron, J.P. An interactive vegetation SVAT model tested against data from six contrasting sites. *Agric. For. Meteorol.* **1998**, *92*, 73–95. [CrossRef]
36. Masson, V.; Le Moigne, P.; Martin, E.; Faroux, S.; Alias, A.; Alkama, R.; Belamari, S.; Barbu, A.; Boone, A.; Bouyssel, F.; et al. The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geosci. Model Dev.* **2013**, *6*, 929–960. [CrossRef]
37. Mahfouf, J.F.; Bergaoui, K.; Draper, C.; Bouyssel, F.; Taillefer, F.; Taseva, L. A comparison of two off-line soil analysis schemes for assimilation of screen level observations. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]
38. Albergel, C.; Calvet, J.C.; Mahfouf, J.F.; Rüdiger, C.; Barbu, A.L.; Lafont, S.; Roujean, J.-L.; Walker, J.P.; Crapeau, M.; Wigneron, J.-P. Monitoring of water and carbon fluxes using a land data assimilation system: A case study for southwestern France. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 1109–1124. [CrossRef]
39. Barbu, A.L.; Calvet, J.C.; Mahfouf, J.F.; Albergel, C.; Lafont, S. Assimilation of Soil Wetness Index and Leaf Area Index into the ISBA-A-gs land surface model: Grassland case study. *Biogeosciences* **2011**, *8*, 1971–1986. [CrossRef]
40. De Rosnay, P.; Drusch, M.; Vasiljevic, D.; Balsamo, G.; Albergel, C.; Isaksen, L. A simplified extended kalman filter for the global operational soil moisture analysis at ECMWF. *Q. J. R. Meteorol. Soc.* **2013**, *139*, 1199–1213. [CrossRef]
41. Wilks, D.S. *Statistical Methods in Atmospheric Sciences*; Academic Press: Cambridge, MA, USA, 2005; Volume 100, p. 676. [CrossRef]
42. Farahmand, A.; AghaKouchak, A.; Teixeira, J. A vantage from space can detect earlier drought onset: An approach using relative humidity. *Sci. Rep.* **2015**, *5*. [CrossRef] [PubMed]
43. Farahmand, A.; AghaKouchak, A. A generalized framework for deriving nonparametric standardized drought indicators. *Adv. Water Resour.* **2015**, *76*, 140–145. [CrossRef]
44. Gringorten, I.I. A plotting rule for extreme probability paper. *J. Geophys. Res.* **1963**, *68*, 813–814. [CrossRef]

45. Blyverket, J.; Hamer, P.D.; Bertino, L.; Albergel, C.; Fairbairn, D.; Lahoz, W.A. An Evaluation of the EnKF vs. EnOI and the Assimilation of SMAP, SMOS and ESA CCI Soil Moisture Data over the Contiguous US. *Remote Sens.* **2019**, *11*, 478. [[CrossRef](#)]
46. Albergel, C.; Munier, S.; Bocher, A.; Bonan, B.; Zheng, Y.; Draper, C.; Leroux, D.J.; Calvet, J.C. 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.* **2018**, *10*, 1627. [[CrossRef](#)]
47. Njoku, E.G.; Entekhabi, D. Passive microwave remote sensing of soil moisture. *J. Hydrol.* **1996**, *184*, 101–129. [[CrossRef](#)]
48. Watts, J. The Swedish Town on the Frontline of the Arctic Wildfires. 2018. Available online: <https://www.theguardian.com/world/2018/jul/30/the-swedish-town-on-the-frontline-of-the-arctic-wildfires> (accessed on 7 February 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Graphic design: Communication Division, UIB / Print: Skjipes Kommunikasjon AS



[uib.no](http://uib.no)

ISBN: 9788230858387 (print)  
9788230847572 (PDF)