An agenda for land data assimilation priorities: Realizing the promise of terrestrial water, energy, and vegetation observations from space

Sujay Kumar[1], Jana Kolassa[2,3], Rolf Reichle[2], Wade Crow[4], Gabrielle de Lannoy[5], Patricia de Rosnay[6], Natasha MacBean[7], Manuela Girotto[8], Andy Fox[9], Tristan Quaife[10], Clara Draper[11], Barton Forman[12], Gianpaolo Balsamao[6], Susan Steele-Dunne[13], Clement Albergel[14], Bertrand Bonan[15], Jean-Christophe Calvet[15], Jianzhi Dong[16], Hannah Liddy[17,18], Benjamin Ruston[9]

[1]Hydrological Sciences Lab, NASA Goddard Flight Center, Greenbelt, MD USA
[2]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD USA
[3]Science Systems and Applications, Inc., Lanham, MD, USA
[4]USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD USA
[5]Department of Earth and Environmental Sciences, KU Leuven, Heverlee, Belgium
[6]European Centre for Medium-Range Weather Forecasts, Reading, UK
[7]Department of Geography and Environment, Western University, London, Ontario, Canada
[8]Environmental Science and Policy Management, University of California, Berkeley, CA USA
[9]Joint Center for Satellite Data Assimilation, Boulder, CO USA.
[10]National Centre for Earth Observation, Department of Meteorology, University of Reading, Reading, UK
[11]NOAA Earth System Research Laboratories, Boulder, CO
[12]Department of Civil and Environmental Engineering, University of Maryland, College Park, MD USA
[13]Delft University of Technology, Department of Geoscience and Remote Sensing, Delft, The Netherlands
[14]European Space Agency Climate Office, ECSAT, Harwell Campus, Oxfordshire, Didcot, UK
[15]Centre National de Recherches Meteorologiques, University de Toulouse, Meteo-France, Toulouse, France
[16]Institute of Surface-Earth System Science, Tianjin University, Tianjin, China
[17]Center for Climate Systems Research, Columbia Climate School, Columbia University, NY USA.
[18]NASA Goddard Institute for Space Studies, NY, USA

Key Points:

- Land data assimilation has shown significant promise for short-term forecasting applications

Corresponding author: Sujay Kumar, `Sujay.V.Kumar@nasa.gov`

- Significant gaps remain in the current land data assimilation systems related to observation utilization and models
- Coordinated development with the modeling and observational community and adoption of technological enhancements are needed in the future

Abstract

The task of quantifying spatial and temporal variations in terrestrial water, energy, and vegetation conditions is challenging due to the significant complexity and heterogeneity of these conditions, all of which are impacted by climate change and anthropogenic activities. To address this challenge, Earth Observations (EOs) of the land and their utilization within data assimilation (DA) systems are vital. Satellite EOs are particularly relevant, as they offer quasi-global coverage, are non-intrusive, and provide uniformity, rapid measurements, and continuity. The past three decades have seen unprecedented growth in the number and variety of land remote sensing technologies launched by space agencies and commercial companies around the world. There have also been significant developments in land modeling and DA systems to provide tools that can exploit these measurements. Despite these advances, several important gaps remain in current land DA research and applications. This paper discusses these gaps, particularly in the context of using DA to improve model states for short-term numerical weather and sub-seasonal to seasonal predictions. We outline an agenda for land DA priorities so that the next generation of land DA systems will be better poised to take advantage of the significant current and anticipated shifts and advancements in remote sensing, modeling, computational technologies, and hardware resources.

Plain Language Summary

Satellite remote sensing measurements have enabled the monitoring of the Earth's land surface with unprecedented scale and frequency. These measurements allow us to monitor the changes on the land surface and understand the contribution of human activities toward them. The information from such observations is combined with the modeled estimates through data assimilation (DA) algorithms. This article discusses the progress made in the development of land DA systems and the major gaps that remain. The paper also outlines priorities that we need to consider in the development of next generation land DA systems so that the potential of land remote sensing measurements can be fully realized.

1 Introduction and premise

Warmer global temperatures associated with climate change have been shown to lead to more uncertainties in water availability both regionally and globally (Barnett et al., 2005; Milly et al., 2005; Berg et al., 2016; Konapala et al., 2020; Wu et al., 2020). There is increasing evidence for an uneven distribution of changes in key water flux rates such as precipitation and evaporation leading to more intense and frequent extreme events, with the lack of predictability in water availability exacerbating water scarcity (Oki & Kanae, 2006; Trenberth et al., 2014; Cook et al., 2018), water excess and natural disasters. The scientific literature shows that changes in atmospheric moisture transport will lead to consequences such as longer droughts in arid and semi-arid regions, more intense precipitation and snowfall, and earlier snow melt, among other changes (Trenberth et al., 2003; J. Huang et al., 2016; Ingram, 2016; Wainwright et al., 2021). Additionally, anthropogenic factors such as agricultural water use and industrial expansion, reservoir management, deforestation, and wetland drainage have not only quickened the pace of land transformation, but have also contributed to, often unpredictable, changes in the water, energy, and carbon cycles (Wood et al., 1997; McDonald et al., 2011; Vorosmarty & Sahagian, 2000; Haddeland et al., 2014; Mehran et al., 2017). Given the urgent need to understand the variability in water cycle fluxes, there have been several efforts to develop inventories of key water cycle fluxes (Trenberth et al., 2007; L'Ecuyer et al., 2015; Rodell et al., 2015; Vargas Godoy et al., 2021). The combined use of modeling and remote sensing estimates of water and energy states and fluxes is a key feature of these studies, which typically report closure of water and energy budgets with relatively small residuals and uncertainty ($\sim$4-8%) at coarse (monthly or larger) temporal scales. However, these studies also note that the imbalances and closure errors increase with increasing spatial and temporal resolution; consequently, concurrent high-quality Earth observations are needed to further constrain the modeled estimates and enable reliable assessment of water, energy, and carbon cycle budgets.

Equally important is the vegetation state, which is also a key factor in the coupling of the water, energy, and carbon cycles. Vegetation exerts key controls on surface water and energy partitioning and the associated response of terrestrial water, energy, and carbon fluxes to climatic drivers (Powell et al., 2013). Changes in vegetation impact surface albedo and surface roughness, interception of precipitation, and reduction in surface runoff (L. Zhang et al., 2015) and can explain up to 30% of the variance in precipitation and surface radiation (Green et al., 2017). Accurately representing vegetation processes in land surface models (LSMs) is thus critical for capturing the efficiency of carbon uptake and transpiration and – more importantly – how this efficiency is expected to change in a future climate.

Despite the importance of biospheric processes, their representation in many LSMs does not accurately capture observed vegetation behavior. The modeled biosphere-atmosphere coupling is much weaker than in observations, in part because the models generally underestimate the observed response of the biosphere to climate (Green et al., 2017; Papagiannopoulou et al., 2017). This can lead to substantial biases in near-surface states, such as air temperature and relative humidity, and can contribute to the misrepresentation of soil moisture dynamics (Koster et al., 2009) – problems that persist at the monthly and longer timescales associated with such vegetation anomalies.

In addition to their role in hydrological and climate prediction, coupled land surface-atmosphere models are integral to numerical weather prediction (NWP) and sub-seasonal to seasonal (S2S) forecasts, with their ability to represent heterogeneity in land use, soil type, vegetation types, soil moisture, the presence of snowpack, and the impact of these factors on the land-atmosphere exchange of energy and water. For example, through the soil-plant-atmosphere continuum, the land surface conditions can control the surface heat fluxes (both latent and sensible) with consequent impacts on air temperature, the height and stability of the atmospheric boundary layer, and occurrence of precipitation (Eltahir, 1998; Dirmeyer & Halder, 2016). This means land surface initialization can influence NWP and S2S forecast performance when the land-atmosphere coupling is strong and when there is sufficient variability, and also significant memory in land surface states (Koster et al., 2004, 2010, 2021)). Over recent years, regions of strong land-atmosphere coupling have been shown to have expanded, to now cover regions that never before exhibited strong land-atmosphere coupling (Dirmeyer et al., 2021). For example, soil moisture conditions have been shown to directly impact temperature and precipitation forecast skill, provided that the forecast model properly represents the processes linking land anomalies, surface fluxes, and atmospheric physics. Models with incorrect land initialization and/or poor land-atmosphere coupling behavior will have biases in temperature, humidity, and precipitation forecasts during periods of drought (Dirmeyer & Halder, 2016; Dirmeyer, 2018). Observational constraints are needed to improve such deficiencies (Dirmeyer et al., 2016).

Data assimilation (DA) is the typical approach used in Earth system modeling to combine information from observational data sources with dynamical process representations in numerical process models to develop estimates that are superior to those obtained by using just the data or model alone (Reichle, 2008a). The resulting estimate from DA is typically called the 'analysis' and is calculated according to the relative errors and uncertainties in the model and the observations. There is a wide variety of methods that are employed to generate such DA analyses, with the mathematical objective of minimizing the posterior error of the DA analysis, whilst being constrained by the available prior information about the system being modeled. The majority of such methods are founded in optimal control or Bayesian estimation theory (Nichols, 2010), which incorporates observations into models in a sequential or non-sequential manner (Bouttier & Courtier, 2002). While the atmosphere and ocean modeling communities (Ghil & Malanotte-Rizzoli, 1991; Malanotte-Rizzoli, 1997; Navon, 2009; Lahoz & Schneider, 2014) have led the charge in the methodological development and applications of DA, it is being increasingly adopted in land surface and hydrology modeling, particularly during the past two decades (Reichle et al., 2009; Houser et al., 2012; de Rosnay et al., 2014).

Given the importance of land surface initialization and land-atmosphere coupling to applications such as NWP and S2S forecasting, it is perhaps surprising that development of LSMs, and land-focused DA activities in particular, have been a relatively low priority for NWP and S2S development in the past. This prioritization is perhaps a reflection of how land surface modeling has developed in relation to atmospheric modeling. In an editorial piece in the American Geophysical Union's newsletter "EOS", Dirmeyer (2018) addressed this question directly, suggesting that a lack of observations of land surface states or land-atmosphere fluxes resulted in the development of land modeling schemes "*to ameliorate systematic atmospheric model errors... addressing one error by introducing additional ones*" (Dirmeyer, 2018). In such a circumstance, the realistic representation of land surface states, as obtained from land DA, may lead to only minimal or disadvantageos impacts on atmospheric forecasts (Crow et al., 2020). This inconsistency was more easily overlooked when high-quality observations of land states were not available. However, the latest generation of satellite and in-situ observations now allow for more rigorous verification of process-model improvements and initialization of forecasts.

In recent decades, there has been a tremendous expansion in the number of remote sensing platforms that provide observations of the land surface (Kimball, 2008; Lakshmi, 2014; McCabe et al., 2017; Balsamo et al., 2018). Over the years, monitoring capabilities from these platforms have encompassed a broad range of land surface/hydrology variables including land use, soil moisture, snow cover, snow water equivalent, land surface temperature, vegetation change/dynamics, terrestrial water storage, and surface water. The technological and scientific success of these spaceborne platforms have now yielded multidecadal observations for many variables (e.g., terrestrial water storage, soil moisture, leaf area index, snow, land surface temperature) enabling the production of long-term climate data records (Lettenmaier et al., 2015). Similarly, for some variables such as vegetation dynamics, soil moisture and precipitation, observations are now provided by multiple platforms, allowing the possibility of increased spatio-temporal coverage from remote sensing. More recent advances in land remote sensing have also enabled complementary remote sensing observations of key water and carbon cycle variables, e.g. precipitation from the Global Precipitation Measurement (GPM; Hou et al. (2014)) mission, surface soil moisture from Soil Moisture Active Passive (SMAP; Entekhabi et al. (2010a)) and Soil Moisture Ocean Salinity (SMOS; Kerr et al. (2010)) missions, terrestrial water storage from Gravity Recovery and Climate Experiment Follow-On (GRACE-FO; Flechtner et al. (2014); Kornfeld et al. (2019)) mission, Solar-induced chlorophyll fluorescence from the Orbiting Carbon Observatory-2 (OCO2; Sun et al. (2017)) mission, and greenhouse gases observing satellite (GOSAT; Doughty et al. (2022)).

Together, these observations facilitate the development of observation-only based water, energy, and carbon budget closure estimates (Trenberth et al., 2014). In addition to conventional space agency driven science missions, sensing platforms from commercial companies have also started to become more commonplace (Dash & Ogutu, 2016; Houborg & McCabe, 2016; Tollefson, 2017). There has been a growing recognition in the community that modeling systems that can effectively exploit such multi-sensor data are key to answering fundamental questions about global and regional water, energy, and carbon cycle changes (Durand et al., 2021). In particular, it is imperative that DA systems be designed to take advantage of this presumably "golden age" of land surface remote sensing (Stavros et al., 2017; Sellars et al., 2013; Schimel et al., 2019) and be designed and positioned to take advantage of upcoming missions as soon as data become available, such as the surface water storage from the upcoming Surface Water Ocean Topography (SWOT; Biancamaria et al. (2016)) mission, Surface Biology and Geology-designate observable (Cawse-Nicholson et al., 2021) or ESA's BIOMASS (Quegan et al., 2019).

The primary objective of this commentary article is to provide an agenda for the next generation land DA systems given the significant current and anticipated shifts and advancements in remote sensing, modeling, computational technologies, and hardware resources. Note that this article discusses these priorities primarily in the context of space-borne remote sensing observations of the land surface. Accurate land characterization has major impacts on atmospheric forecasts, water availability quantification, management of water and agricultural resources, and developmental issues facing the World's population growth (Collins et al., 2013). It is also critical for capturing the long-term carbon-cycle feedbacks of the biosphere to climate and meteorology and thus making more accurate projections of vegetation, water, and energy states with elevated CO_2 conditions (Canadell et al., 2021). Effective DA systems have a critical role to play in realizing the full information potential and promise of Earth science remote sensing measurements. The specific goals of this agenda paper are to:

- Define the most significant gaps in and priorities for the development of regional to global scale land DA systems for water, energy, and vegetation predictions at NWP and S2S lead times.
- Identify priorities for coordination between the modeling, remote sensing, and technology communities to improve the effectiveness of land DA.
- Provide a framework for future enhancements of DA systems.

Note that the present paper focuses on DA in the context of state estimation. Remote sensing observations can also be used to reduce errors associated with model parameters and structure. Challenges and priorities related to LSM calibration and parameter estimation, with a particular focus on coupled carbon-vegetation-water interactions related to longer-term biosphere-climate feedbacks, will be addressed in a separate companion paper (Quaife et al., in preparation). In parallel, a perspective is offered (De Lannoy et al., 2022) on how to address the increased complexity of land DA for either state or parameter estimation (or both) in the future.

2 Background

As described in Reichle (2008a), the primary objectives of DA include extending the spatio-temporal coverage of infrequent observational data, science translation of raw measurements from satellites (or other data sources), and optimal merging of information from multiple sources while preserving the basic constraints on the physical system. By incorporating observational inputs, DA also helps in reducing the uncertainty of estimated land states and fluxes within Earth systems models. This is particularly important given the proliferation of models and modeling systems, leading to large uncertainties purely from differences in conceptual formulations and parameterizations (Clark et al., 2017). The observational constraints from DA are a practical approach to reduce such uncertainties, so that reliable estimates of the water and carbon cycles, and their feedbacks, can be developed.

Taking advantage of the increased availability of environmental datasets, atmospheric DA has played fundamental roles in steadily improving regional and global environmental prediction capabilities (Benjamin et al., 2018; Hersbach et al., 2020). In particular, the prediction of precipitation, short-term extreme events, and applications of nowcasting (Kain et al., 2010; Mass, 2012) have hugely benefitted from the advances in DA techniques to the point that DA is considered an integral component of NWP systems (Kalnay, 2003). Similarly, DA has been key in global and regional ocean prediction and reanalysis systems (Moore et al., 2019; Zuo et al., 2019). These communities have also pioneered the development of the majority of variational and ensemble DA techniques used in operational and reanalysis environments. These methodological advances have been adapted by the land/hydrology/biogeochemistry communities to develop similar improvements in the land surface water and carbon cycle predictions (Reichle, 2008a; Rayner, 2010; de Rosnay et al., 2014; Lahoz & De Lannoy, 2014; Rayner et al., 2019; Fairbairn et al., 2019; Smith et al., 2020).

In the past several decades, there has been a significant increase in the development and application of DA to LSMs, with the majority using sequential filtering approaches. The use of variational DA approaches, which are popular in the atmospheric and ocean modeling communities, is less common in NWP and S2S land DA applications. Variational DA approaches are more widely used for optimizing vegetation and carbon cycle related parameters LSMs, with the goal of better estimating longer-term (annual to centennial) carbon stocks and fluxes (Kaminski et al., 2013; Pinnington et al., 2020; MacBean et al., 2022), although these approaches have also been used to improve LSM soil moisture predictions (Scholze et al., 2016; Pinnington et al., 2018; Raoult et al., 2021). Variational methods provide updates - by minimizing a specified cost function - that constrain the model forecasts with a set of observations. They require the use of linearized observation operators, and in some cases, a tangent linear or adjoint model that represents the sensitivity of model states backward in time (Courtier et al., 1993). Some variational applications exist for land DA that use tangent linear or adjoint models (e.g. Peylin et al. (2016); Raoult et al. (2016); Schurmann et al. (2016)) or adopting a linearity assumption (Hess, 2001; Balsamo et al., 2004) but, in general, the development of such models for the land surface has been more challenging due to the tendency for LSMs to contain non-differentiable threshold processes (Reichle et al., 2002a). In variational DA systems, the temporal evolution of the background error covariance is typically not modeled, whereas Kalman and particle filtering approaches allow dynamic updates of these estimates (Rabier & Liu, 2003).

Variants of Kalman filters are often used in sequential DA systems. The Extended Kalman Filter (EKF), the nonlinear version of the Kalman filter, is used in some land surface DA environments (Rudiger et al., 2010; Ghent et al., 2010; de Rosnay et al., 2013), by employing linearized observation operators and error variance propagation formulations (Reichle et al., 2002a). The ensemble Kalman filtering (EnKF) methods (Evensen, 2003), which are Monte-Carlo approximations of the Bayesian filtering process, provide a more convenient approach that is suitable for weakly non-linear models. In addition, due to its Monte Carlo formulation, the EnKF is easy to implement and provides a convenient means for representing the wide variety of errors impacting LSMs. The use of an ensemble of model realizations in the EnKF also allows for the implicit propagation of model errors, making it more computationally tractable and scalable for large systems. As a result, the EnKF has become widely used in the land surface community. Non-sequential extensions of the EnKF that allow information updates backwards over a time window have also been developed (Evensen & Leeuwen, 2000). Detailed methodological descriptions of various DA approaches are discussed in several prior review papers, including (Reichle, 2008a; van Leeuwen, 2009; Lahoz & De Lannoy, 2014; Lahoz & Schneider, 2014; Houtekamer & Zhang, 2016; van Leeuwen et al., 2019).

Land DA applications have been examined for a wide range of land surface water, energy, and carbon variables. Given the developmental legacy of LSMs as the boundary conditions to atmospheric models, a majority of the land DA efforts are focused on constraining variables such as soil moisture and snow that have relatively long-memory and a direct impact on the NWP initialization skill (de Rosnay et al., 2014; Santanello et al., 2019; Gomez et al., 2020; Benjamin et al., 2022). Many NWP centers currently employ some form of screen level assimilation of synoptic measurements of air temperature and relative humidity close to the surface (Mahfouf, 1991), updating soil moisture states within DA. Motivated by the need for improved land characterization for short-term forecasts and hydrology applications, a wide variety of studies have focused on the assimilation of more direct information of soil moisture from ground and spaceborne platforms, with a range of DA techniques (Houser et al., 1998; Margulis et al., 2002; Parajka et al., 2006; Drusch, 2007; Sabater et al., 2007; Kumar et al., 2012; Draper et al., 2012; Renzullo et al., 2014; Lievens et al., 2015; De Lannoy & Reichle, 2016a; Lievens et al., 2016; Kolassa et al., 2017a, 2017b; Lievens et al., 2017; Reichle et al., 2019; Rodriguez-Fernandez et al., 2019; De Lannoy et al., 2019). These studies encompass the assimilation of soil moisture measurements and retrievals from passive microwave satellite platforms ranging from Special Sensor Microwave Imager/Sounder (SSMI/S) in the early 1990s to more recent missions such as SMOS and SMAP that rely on L-band microwave sensors. While less numerous, other studies have focused on the assimilation of root-zone soil moisture proxy information derived from thermal-infrared remote sensing (e.g., Hain et al. (2012)).

In addition to soil moisture, assimilation of snow cover and snow water equivalent (SWE) relevant measurements have also been examined with approaches ranging from direct insertion to variants of Kalman filtering and smoothing. Snow cover observations from optical sensors, which do not directly provide quantitative information on the amount of snow in the snowpack are often assimilated using rule-based update methods (Rodell & Houser, 2004; Zaitchik & Rodell, 2009; Arsenault et al., 2013; De Lannoy et al., 2012). A few studies have also used the EnKF to assimilate snow cover observations (Clark et al., 2006; H. Su et al., 2008; Y.-F. Zhang et al., 2014; Toure et al., 2018). These studies typically demonstrate modest improvements for estimating ephemeral SWE and little benefit over deeper snowpacks, owing to the fact that the information about SWE in instantaneous snow cover observations tends to diminish for larger values of SWE, as snow cover saturates at 100%. To address the weak instantaneous correlation between snow cover and SWE, a number of studies (Durand et al., 2008; Girotto et al., 2014; Margulis et al., 2015; Oaida et al., 2019) employed smoothing approaches that assimilate all snow cover observations within a time window. Estimates of SWE from ground measurements, microwave, and lidar platforms have also been widely used for DA in snow hydrology studies (Andreadis & Lettenmaier, 2006; Slater & Clark, 2006; Liu et al., 2013; Magnusson et al., 2014; Margulis et al., 2015; Kumar et al., 2015a; C. Huang et al., 2017; Margulis et al., 2019; Smyth et al., 2019). Efforts to assimilate brightness temperature observations from microwave instruments and radar backscatter measurements using machine learning and radiative transfer model operators for improving snow estimation have also been reported over different regions of the world (Phan et al., 2014; Xue & Forman, 2017; Xue et al., 2018; Kwon et al., 2019; Park et al., 2021). Consequently, some operational centers around the world have enabled a form of land analysis to incorporate remote sensing information for NWP initialization (Barrett, 2003; Pullen et al., 2011; de Rosnay et al., 2013, 2014; Carrera et al., 2019; Wegiel et al., 2020; Gomez et al., 2020). These analyses also range from optimal interpolation to Kalman filter methodologies.

In addition to work within NWP and with offline LSMs, there exists a large set of applications that rely on accurate estimates of land surface conditions. Water resource planning efforts require reliable estimates of water fluxes and their variability to account for water availability issues and hydrological extremes. Consequently, these inputs are also becoming critically important for agricultural and food security assessments. Military agencies rely on accurate land characterization of surface states for mobility and trafficability assessments. Water storage in peatlands controls large carbon fluxes to the atmosphere. Estimates of river flow are important for reservoir management as well as transboundary water applications. Given these significant needs for land surface hydrology information, there have been land-focused systems (sometimes called ’offline’ environments) developed in the past several decades. The Data Assimilation Research Testbed (DART; Anderson et al. (2009)), developed at the National Center for Atmospheric Research has enabled ensemble DA for the land surface schemes of several Earth system models. The Global Land Data Assimilation System (GLDAS; Rodell et al. (2004)) and the North American Land Data Assimilation System (NLDAS; Mitchell et al. (2004)) were the two pioneering efforts that established software frameworks for driving LSMs with observationally constrained meteorology, with the eventual goal of enabling the assimilation of land surface states. Other examples of such land-only simulation products include MERRA-Land (Reichle et al., 2011), ERA-Interim/Land (Balsamo et al., 2015), and ERA5-Land (Munoz-Sabater et al., 2021), which supplement major atmospheric reanalysis products. Another land-only product is LDAS-Monde (Albergel et al., 2017, 2020), which jointly assimilates surface soil moisture and leaf area index. Moreover, the land-only SMAP Level-4 Soil Moisture product focuses on the assimilation of SMAP radiance observations (Reichle et al., 2019). A flexible environment called the Land Information System (LIS; Kumar et al. (2006)) was developed at NASA that allows for the interoperable use of multiple land and hydrology models and formal DA capabilities around them (Kumar et al., 2008). The capabilities of LIS are not only used in both GLDAS and NLDAS, but have also fostered many LDAS configurations and instances in different parts of the world (Goncalves et al., 2009; McNally et al., 2017; Kumar et al., 2018; Erlingis et al., 2021). The development of similar LDAS environments has also been reported in other organizations around the world (Lewis et al., 2012; Reichle et al., 2014; Sawada et al., 2015; Carrera et al., 2015; Albergel et al., 2017, 2020).

The development of land-only DA environments has enabled the assimilation of other land hydrology variables, including some that are not viable for operational NWP needs because of their long latency. For example, Terrestrial Water Storage anomalies from gravity missions such as GRACE have been used in DA configurations to develop inferences on subsurface changes (Zaitchik et al., 2008; Syed et al., 2008; B. Li et al., 2012; Kumar et al., 2016; Girotto et al., 2017; Zhao & Yang, 2018) The use of GRACE data has provided unprecedented information about the changes in Earth's groundwater storage and has uncovered unsustainable exploitation from groundwater pumping in many areas of the world (Rodell et al., 2009, 2018; Lo et al., 2016; Thomas & Famiglietti, 2019; Nie et al., 2019). The use of GRACE-based constraints also has enabled critical inputs for drought monitoring efforts, by providing an observation-informed estimate of groundwater changes, a critical water resource that varies at longer timescales than surface water (Houborg et al., 2012; B. Li et al., 2019; Getirana et al., 2020).

With the availability of prognostic representations of carbon processes within the third and fourth generation LSMs (Pitman, 2003; Fisher & Koven, 2020), there have been several efforts to utilize remotely sensed vegetation datasets through DA (e.g. in the aforementioned LDAS-Monde system). There is a long legacy of high-resolution measurements of canopy states from optical sensors such as leaf area index, normalized difference vegetation index, fraction of photosynthetically active radiation, and vegetation biomass. Assimilation of leaf area index to improve the estimation of crop yields, vegetation biomass, root zone soil moisture, and carbon fluxes has been demonstrated in several prior studies (Demarty et al., 2007; Quaife et al., 2008; Dente et al., 2008; Sabater et al., 2008; Jarlan et al., 2008; Barbu et al., 2011; Nearing et al., 2012; Fox et al., 2018; Kumar et al., 2019; Raczka et al., 2021). At coarser resolution, analogs of vegetation changes can also be obtained from passive microwave radiometry (Konings et al., 2016; Feldman et al., 2021). While optical sensor measurements are subject to coverage limitations from cloud obscuration, microwave measurements are nearly all-weather. The assimilation of vegetation optical depth estimates from microwave measurements has also shown promising results in improving land surface states, and it provides the opportunity to extend the spatio-temporal coverage of higher-resolution optical measurements (Kumar et al., 2020a; Mucia et al., 2021). Many of these studies have explicitly shown the significant promise of vegetation assimilation in capturing human management impacts such as agriculture, where anthropogenic activities lead to seasonal changes in vegetation growth (Mocko et al., 2021). Vegetation assimilation has also shown effectiveness in capturing vegetation disturbance features from fires, which often lead to changes in regional hydrological response (Kumar et al., 2021, 2022). DA offers a more practical approach to representing these changes on the land surface, which are difficult to characterize with process-based models alone (Pongratz et al., 2018).

Other reported land DA instances involve variables such as land surface temperature, albedo, and water surface elevation. Land surface temperature and albedo from thermal and optical sensors are variables at the heart of the surface energy balance and influence the latent and sensible heat fluxes at the land-atmosphere boundary. Therefore, assimilation of these variables should have a direct influence on weather and climate model forecast accuracy. Unlike variables such as soil moisture, snow, terrestrial water storage, and vegetation, the response timescale of land surface temperature is more rapid, with associated diurnal and seasonal cycles. Additional challenges related to bias correction at these finer timescales are described in studies of land surface temperature assimilation, which have employed both variational and filtering techniques (Castelli et al., 1999; Lakshmi, 2000; Bosilovich et al., 2007; Ghent et al., 2010; Reichle et al., 2010; Han et al., 2013; Draper et al., 2015; Lu et al., 2017). Comparatively, fewer studies examine the assimilation of albedo, despite the classification of albedo as an essential climate variable (Hollmann et al., 2013), particularly over large spatial scales. Most of the reported studies employ direct insertion approaches due to the lack of prognostic representation of albedo variables within LSMs and have reported some positive impacts on characterizing surface fluxes and snow states (Malik et al., 2012; Yin et al., 2016; Boussetta et al., 2015; Kumar et al., 2020b).

In addition to NWP and S2S, improving the initial conditions for hydrological forecasting applications through DA has also been the subject of many studies (e.g., Liu et al. (2012); Baugh et al. (2020)). Most of these focus on improving soil moisture (e.g., Loizu et al. (2018)) and snow states (e.g., D. Li et al. (2019)), though a few studies have also examined the assimilation of hydrological states and fluxes such as discharge (Clark et al., 2008; D.-J. Seo et al., 2009; Pauwels & De Lannoy, 2009; Thirel et al., 2010; DeChant & Moradkhani, 2011; Fairbairn et al., 2017). Estimates of water stage or elevation have been derived from radar altimeters since the early 1990s, providing nearly all-weather observations (Silva et al., 2010). The use of DA within hydrodynamic models has been examined to extend the spatio-temporal utility of these typically narrow swath altimeter measurements. These studies include efforts to assimilate river discharge, velocity, and water level (Andreadis et al., 2007; Neal et al., 2007; Ricci et al., 2011; Paiva et al., 2013; Ercolani & Castelli, 2017; Emery et al., 2020; El Gharamti et al., 2021). The improved surface water characterization over data-poor transboundary basins from the assimilation of remote sensing data is a highly relevant hydrological modeling challenge not routinely employed in hydrological prediction systems.

3 Major gaps in land data assimilation

As described in the previous section, there have been significant advances in the development of methodologies for land DA, the utilization of land surface measurements from various spaceborne platforms as well as in-situ observations, and the use of land DA for routine applications. Despite this progress, there are several significant remaining gaps and challenges, which are described in this section. Figure 1 shows a summary of the key topics described in this section.

3.1 Reliance on the assimilation of retrieval products

All DA methods require an "observation operator" or forward model to translate variables from model space to observation space, and vice versa. The simplest observation operator is the identity matrix where the assimilated value is either an observation (in-situ soil moisture, atmospheric temperature) or a retrieval from (typically remotely sensed) observations (e.g., soil moisture, snow mass, or leaf area index) of a variable that is also estimated by the LSM. To date, most of the reported land DA studies and applications have focused on the assimilation of retrieved products, whereas the atmospheric and ocean communities have made far more progress in the direct use of radiance measurements. The retrievals are obtained using radiative transfer or empirical models with their own underlying assumptions of parameter, land surface, and climate characteristics. These assumptions are often inconsistent with those simultaneously made in LSMs that are utilized in DA instances. In many cases, retrieval models are developed and calibrated with limited airborne and field campaign data. Subsequently, their application over different or large spatial extents leads to errors. The assimilation of satellite radiances as in the SMAP L4_SM algorithm (Reichle et al., 2019) is likely to reduce errors stemming from such inconsistencies and potentially make observation errors more easily quantifiable (Lewis et al., 2012; Lievens et al., 2017; de Rosnay et al., 2022), though forward radiative transfer models also suffer from uncertainties related to conceptualization and parameterization. As noted earlier, given the complexity of forward modeling for snow mass, machine learning models have also been used as observation operators for snow DA studies.

The historical reliance on retrieval products is a significant obstacle for the more widespread adoption of land DA in near-real-time or operational modeling environments. This is partly due to the fact that latency associated with retrieval products often does not meet the constraints of operational environments. Radiance assimilation also provides an opportunity to update states impacting the forward modeling of land surface emission (e.g., vegetation states controlling canopy opacity). To realize the more common use of radiance measurements within land DA environments, relevant forward modeling tools, whether they use physically-based radiative transfer models (e.g., Quaife et al. (2008); De Lannoy et al. (2013)) or employ statistical approaches (e.g., Aires et al. (2021); Kolassa et al. (2018)), are needed. Compared to the breadth of LSMs, only a handful of land relevant community forward modeling systems and tools are currently available (Parrens et al., 2014; Royer et al., 2017).

3.2 Limitations from developmental legacies

The legacy of LSMs has also significantly influenced the application of land DA instances. As mentioned earlier, the first generation LSMs were conceptualized as the boundary conditions to atmospheric models with the primary objective of improving atmospheric prediction. The focus of ensuring physical realism of land surface states has historically been a secondary priority. For example, while most of the major NWP centers around the world maintain detailed scorecards of prediction skill of atmospheric variables (Ebert et al., 2013; Haiden et al., 2018; Brown et al., 2021), such routine evaluations for land surface variables are uncommon. Consequently, process representations, and therefore the climatologies of variables such as soil moisture, have evolved specifically for each associated NWP system. It is well documented that soil moisture climatologies from different models are starkly different (Reichle et al., 2004), to the point that they are essentially model-specific representations of soil wetness (and should not even be called soil moisture; Koster et al. (2009)).

As noted earlier, the screen level assimilation approaches used in many NWP centers update soil moisture states instead of atmospheric states assuming a causative relationship between soil moisture boundary conditions and low-level atmospheric forecasts. Since these assumptions are only valid when weakly forced atmospheric conditions are present, the screen level DA scheme to adjust soil moisture often has become an approach that compensates for true underlying errors in other parts of the model (Draper et al., 2009). It is recognized that the resulting improvements in the low-level atmospheric forecasts are not necessarily due to improved representation of land surface states (Douville et al., 2000; Drusch & Viterbo, 2007), and may in fact degrade the surface states and hydrological consistency (e.g., Zsoter et al. (2019)). Consequently, the assimilation of observations or retrievals which have a direct relation to soil moisture becomes problematic in such DA environments. For example, studies have demonstrated that assimilation of satellite soil moisture data and screen level temperature and humidity observations often lead to diverging estimates of root zone soil moisture updates (Draper et al., 2011), though satellite soil moisture assimilation itself improves land surface states (Carrera et al., 2019; Munoz-Sabater et al., 2019). As we noted above, with the evolution of LSMs, they are also increasingly being used in environments where there is a direct need for accurate characterization of land surface, without being necessarily connected to atmospheric models and requirements. As the need to improve the physical realism of land surface states increases and direct measurements of land surface states become increasingly available, methodologies that treat the land surface states as an 'error sink' are fundamentally undercutting the potential of land DA. Localization strategies (Anderson, 2007) that selectively adjust portions of the model states most closely related to the observation may also be needed to circumvent such limitations.

3.3 Bias correction and efficiency issues

Another key issue related to the lack of physical realism of LSMs is the current practice of bias correction within land DA environments. In real data assimilation systems, systematic errors (including biases) between model estimates and observations are unavoidable, and they are typically caused by a combination of errors from instrument noise, retrieval issues, and model deficiencies, related to model structure and uncertainties in parameters and inputs. The model structural deficiencies are also tied to the legacy issues mentioned in Section 3.2. Additionally, the lack of physical realism in LSM estimates also stems from the difficulties of the complex heterogeneity of the land surface, which makes it difficult to develop spatially distributed assessments of biases from limited in-situ measurements (that are not spatially representative). Prior studies have explored both bias-aware and bias-blind approaches in land DA (Kumar et al., 2012). Bias-aware approaches typically attribute the source of the biases exclusively to the model or the observations and use the analysis increments to progressively estimate the bias (De Lannoy et al., 2007; Bosilovich et al., 2007; Reichle et al., 2010). While these approaches are useful for systems with transient changes in bias, the lack of realism in the model estimates or observations makes their application difficult. Bias-blind approaches are more common in land DA systems, where the relative biases between model estimates and observations are removed before assimilation, essentially focusing on the correction of errors in short-term or interannual variations. Approaches such as Cumulative Distribution Function (CDF)-matching and standard normal deviate scaling are commonly used in soil moisture assimilation systems to rescale the observations into the model climatology prior to assimilation (Reichle & Koster, 2004; Crow et al., 2005a; de Rosnay et al., 2020). The observation rescaling approaches require long data records of observation and modeled soil moisture estimates to develop scaling parameters, which are difficult with new satellite missions. Similarly, long model integrations with a consistent driving meteorology are required to establish the model climatology, which also often do not exist for operational environments because of frequent system updates. Moreover, regional modeling centers also often change their domains of focus, which makes the reliance on an established soil moisture climatology specific to each modeling domain impractical. The rescaling approaches are also documented to reduce the information transfer from retrievals by about 10% due to errors in empirical CDF computations (Nearing et al., 2018) and violation of theoretical assumptions underlying the matching of higher-order statistical moments in the CDF approach (Yilmaz & Crow, 2013).

Further, recent studies have shown that the rescaling approaches are problematic when stationarity assumptions about the model-observational biases do not hold true, dynamic changes in bias occur, or when human management factors or unmodeled processes are present in the observational signal and missing from the models (Kumar et al., 2015b; Girotto et al., 2017). In this regard, the land surface is unique compared to the atmosphere and ocean, as the direct influence of human activities (e.g., irrigation, vegetation disturbances, groundwater pumping, urbanization, reservoir management) is ubiquitous on the land. Because many of these processes are subjective in nature, their accurate representation within models is hard. For example, though irrigation representations have been developed within LSMs, the determination of irrigation onset, duration, and magnitude is difficult to specify accurately (Massari et al., 2021). The use of remote sensing observations through DA offers a more practical approach to incorporating such impacts, without the reliance on heuristic rules. However, as the unmodeled signals are often manifested as systematic errors or biases, new approaches are needed that can identify the root causes of bias, including bias from systematic instrument errors, model parameter and parameterization deficiencies, and unmodeled processes; otherwise, it will remain difficult if not impossible to preserve important observational signals (Kumar et al., 2015b).

3.4 Methodological limitations

As noted earlier, owing to the difficulties in deriving adjoint models for highly non-linear LSMs, Kalman filtering approaches have become the most commonly used methods for land DA, providing a computationally efficient alternative to variational DA systems with comparable performance (Reichle et al., 2002b; Caparrini et al., 2004; Kotsuki et al., 2022). One of the key assumptions in the serial applications of the EnKF is that model and observation errors are Gaussian and mutually and serially uncorrelated (Katzfuss et al., 2016), which are often not realistic in land DA systems. The particle filter (Moradkhani et al., 2005; Smyth et al., 2019) overcomes this limitation, by approximating the model posterior distribution with Monte Carlo sampling, but requires large ensemble sizes (Weerts & El Serafy, 2006; Dong et al., 2015). Another possibility are hybrid ensemble-variational techniques (Bannister, 2017), which solve the variational problem by approximating required properties of the adjoint and tangent linear using a model ensemble and are being increasingly used in NWP. Hybrid approaches also show promise for overcoming some of the limitations inherent in other ensemble DA techniques (such as limited sample size), and have been applied to land surface modeling problems (e.g., Pinnington et al. (2021)).

However, all of these ensemble DA systems can suffer from issues of filter collapse and degeneracy when adequate ensemble spread does not exist, which is sometimes addressed with techniques such as covariance inflation (Anderson & Anderson, 1999; Fox et al., 2018; Gharamti, 2018, 2021). These techniques also require significant tuning and knowledge of observation error covariances (Miyoshi & Yamane, 2007). The reliance on the ensemble to derive the model error covariances can be limiting in some instances when the state variable is bound by theoretical limits. For example, since fractional snow cover is constrained by zero as the lower limit and unity as the upper limit, the effective generation of spread in the ensemble becomes problematic when the model estimate is close to these bounds and has been shown to cause disadvantageous impacts within DA (Arsenault et al., 2013). Reliance on a single choice of surface meteorological forcing has also been shown to underestimate the land model error covariance and reduce the efficiency of land DA (Kumar et al., 2017). Studies have also shown that DA can lead to water balance errors (which are particularly important in coupled land-atmosphere systems) when the analysis increments are not explicitly constrained to conserve the water balance (Girotto et al., 2021). The use of weak constraints in land DA systems has been explored to reduce these inefficiencies (Pan & Wood, 2006; Yilmaz et al., 2011). Finally, most DA systems are designed to work with errors that are strictly random. Since systematic errors are more dominant in land hydrology processes, the primary focus on improving short-term errors through DA often limits the level of improvements possible through land DA (Section 3.3).

3.5 Lack of information transfer in land data assimilation efforts

A key appeal of DA is the potential to extend the information in observations to other model states that are connected to the observations. Land DA efforts have yielded mixed results in effectively realizing this goal, due to a variety of reasons. For example, the assimilation of surface soil moisture retrievals has been widely used to develop improvements in deeper soil moisture, with studies reporting varying degrees of success. The low skill of soil moisture retrievals from older sensors was a key factor in early soil moisture DA studies that reported small improvements in root zone estimation. Larger improvements in subsurface soil moisture estimation with DA have been reported with data from newer L-band sensors such as SMOS and SMAP (De Lannoy & Reichle, 2016a; Reichle et al., 2017). Similarly, larger improvements from DA are often reported with studies involving older or simpler soil moisture models (F. Li et al., 2010; Mladenova et al., 2019), as the information from remote sensing data compensates for the lack of skill in those models. When provided with high-quality boundary conditions, modern LSMs are fairly skillful in their soil moisture estimates. As a result, the added improvements from soil moisture DA featuring such models and high quality precipitation inputs are relatively small (Kumar et al., 2014; Kolassa et al., 2017b; E. Seo et al., 2021; Reichle et al., 2021a). The strength of the surface-subsurface soil moisture coupling within LSMs is a significant factor in the level of improvements obtained in the root zone through surface soil moisture assimilation (Kumar et al., 2009). Given that the true surface-subsurface coupling strength is largely unknown and significantly heterogeneous, these studies suggest that there is a strong model configuration dependence on the level of improvements in root zone soil moisture realized through surface soil moisture assimilation.

Attempts to extract information from the assimilation of remote sensing soil moisture observations about associated key water cycle variables such as evapotranspiration (ET), runoff, and river discharge have also produced mixed results. For example, most soil moisture and vegetation assimilation studies that examined the impact on ET and/or river discharge either report marginal improvements or degradations (Peters-Lidard et al., 2011; Fairbairn et al., 2017; Albergel et al., 2017; Kumar et al., 2018; Munoz-Sabater et al., 2019; Bonan et al., 2020). Similar to the earlier example on root zone soil moisture, the information transfer issues with other variables such as ET are also related to the inherent assumptions of coupling between the relevant variables (Crow et al., 2018, 2020). In reality, such coupling relationships are dependent on a number of factors including land surface characteristics, surface meteorology, and climate conditions (Dong et al., 2020a). For example, soil moisture has a strong influence on ET in a water limited domain but may not be the primary controlling factor in an energy limited domain (Kumar et al., 2020a). Similarly, LSMs often underestimate the coupling between simulated soil moisture and subsequent surface runoff and thus potentially squander valuable streamflow predictability associated with soil moisture assimilation (Crow et al., 2017, 2018, 2019). Consequently, the results from DA will be misleading if the LSM does not adequately represent such coupling relationships accurately, even when high quality observations are assimilated. The impact of such inherent model coupling features is also reflected in studies assimilating observations related to vegetation that have shown stronger impact on root zone soil moisture than from the assimilation of surface soil moisture retrievals (Albergel et al., 2017, 2020; Kumar et al., 2020a).

In other cases, the lack of success in information transfer is related to a lack of adequate model prognostics or limited physics. For example, the assimilation of land surface temperature within LSMs has been problematic because there are large discrepancies in the representation of surface temperature in models and the skin temperature derived from remote sensing instruments (Bosilovich et al., 2007). Partly due to the lack of an equivalent model prognostic, the land surface temperature assimilation results have yielded little improvement in surface fluxes (Reichle et al., 2010). Similarly, though terrestrial water storage retrievals from the GRACE mission have provided valuable insights on subsurface water storage changes around the world, their assimilation may produce erroneous trends because most LSMs only include a shallow groundwater representation (whereas GRACE may be observing deeper groundwater storage signals; Girotto et al. (2017)).

Another key factor in the lack of information transfer is the limited skill in observations themselves. Though there is often a direct connection between streamflow and melt from the snowpack, little success in this area is reported from snow DA studies that solely rely on remote sensing observations (Andreadis & Lettenmaier, 2006; Kumar et al., 2014; Zsoter et al., 2019). On the other hand, the use of more reliable data from airborne or ground measurements of snow has shown greater promise in improving streamflow (C. Huang et al., 2017; Lahmers et al., 2022), suggesting that significant improvements in remote sensing snow retrievals are needed to improve their utility. As noted above, despite the availability of multi-decadal observations of land surface variables such as land surface temperature, albedo, and leaf area index, these observations were used in relatively few DA studies, mainly because LSMs have large model biases or have not considered and developed appropriate prognostic representations of such variables.

3.6 Spatio-temporal characterization of error specifications

The specification of model and observation errors is a critical factor in the performance of land DA systems and remains a significant challenge (Crow & Loon, 2006). Observation errors must take into account instrument errors, errors in observation operators or retrieval algorithms (considered synonymous with representativeness errors (van Leeuwen, 2015)); consequently, observation error parameters are difficult to measure and specify in a spatially distributed manner. Though spatial correlations are present in the observation errors, they are often ignored in DA environments (Ying, 2020). Model error specification must consider limitations in model physics and uncertainties in boundary conditions and parameters, among other factors. It is generally recognized that it is more difficult to specify model errors than observation errors because the general lack of physical realism in most LSM outputs makes it difficult to perform direct comparisons against available ground/truth measurements (Reichle, 2008a). Additionally, the significant heterogeneity and complexity of land processes makes it difficult to develop spatially distributed estimates of model errors. Because of these limitations, the derivation of model error covariance specifications is often done through idealized experiments and largely remains a subjective process (Kumar et al., 2008, 2009; Carrera et al., 2015; Fox et al., 2018; Raczka et al., 2021). In fact, most DA studies to-date employ domain-wide uniform error specifications. DA diagnostics such as innovations and analysis increments are routinely used to evaluate error assumptions in DA configurations, by examining their deviations from the theoretical ideal values (e.g., Reichle et al. (2017)). These methods have also been employed within DA systems to adaptively diagnose and improve error variance specifications (Reichle et al., 2008a; Crow & Reichle, 2008). Similarly, the application of the triple collocation approach has also been shown to improve the estimation of soil moisture retrieval errors by removing the problematic assumption – required by innovation analysis – that errors in assimilated observations are serially uncorrelated (Crow & Berg, 2010). These theoretical approaches, however, become less effective when observational signals are dominated by unmodeled processes, which can be especially true for land DA systems (Kumar et al., 2015b). In land-ensemble filtering, the forecast ensemble spread is typically created by adding small perturbations sampled from randomly generated noise to the model states at regular time intervals. Recent studies have shown that this approach creates unrealistic estimates of forecast uncertainty, since the ensemble spread reflects the local model persistence of the applied perturbations, which is not necessarily an accurate representation of the model uncertainty (Draper, 2021). Explicit consideration of uncertainty from several contributing sources for more realistic model spread representation has been reported in some studies (Raeder et al., 2021; Dokoohaki et al., 2022). Additionally, the model background error can be underestimated due to the limitations of the specific model configurations (Kumar et al., 2017).

3.7 Limitations in the use of metrics and evaluation strategies

The typical approach in DA studies is to use the available (often point-scale) ground measurements as the truth for evaluating the impact of DA, with the inherent assumption being that the spatio-temporal scale of the model and remote sensing data is sufficiently fine to contain information about the in-situ measurements. This assumption may not always be valid, as shown in studies that have quantified information use efficiency within DA systems (Nearing et al., 2018). The assessments versus in-situ measurements also overlook the larger utility of remote sensing datasets to provide spatially distributed coverage, offering the possibility to capture spatial features and heterogeneity that cannot be obtained with in-situ measurements. Finally, the geographic distribution of in-situ stations is skewed toward data-rich regions, where model performance is generally good, whereas the impact of land DA is generally larger in data-sparse regions, which was demonstrated using instrumental variable approaches (e.g., Dong et al. (2019); Reichle et al. (2021a)), including the aforementioned triple collocation method (Gruber et al., 2020). Therefore, spatially distributed evaluation strategies and metrics must be emphasized in demonstrating the potential utility of remote sensing datasets. For example, the impacts of agricultural activity and man-made disturbances such as fires are hard to represent in models or capture through ground measurements. Assimilating vegetation datasets, on the other hand, has been effective in representing such features over large spatial scales (Kumar et al., 2020a, 2021; Mucia et al., 2021; Kumar et al., 2022).

Additionally, when representativeness differences are significant with variables such as soil moisture, DA evaluations mainly focus on the use of temporal anomaly metrics (Entekhabi et al., 2010b). The restricted focus on the evaluation of temporal variability has been shown to underestimate errors in soil moisture datasets (Dorigo et al., 2010) and miss out on capturing the impact of unmodeled processes (Kumar et al., 2015b). Evaluation strategies based on approaches such as information theory (Shannon, 1948), as used for instance in NWP context (Cardinali, 2009), that allow for the spatially distributed quantification of the information transfer and noise reduction can be considered to better demonstrate the utility of land remote sensing and DA (e.g. Balsamo et al. (2007)). Since the land is particularly impacted by anthropogenic impacts, land DA must also focus on demonstrating the characterization of such processes, beyond simply improving processes driven by natural variability. It must be stressed that some of these evaluation limitations are due to a profound lack of observational capability for many land surface variables. For example, the point-to-satellite footprint representativeness issues for soil moisture are more severe for the representation of spatial patterns than for temporal anomaly statistics (Chen et al., 2019). Datasets or observational techniques that allow for a direct evaluation of soil moisture patterns are generally lacking.

3.8 Multivariate and land reanalysis efforts

While there have been numerous efforts to develop consistent multi-decadal estimates of atmospheric and ocean states through reanalysis of all available remote sensing datasets, such efforts are largely non-existent in the land and hydrology community. This is partly because land DA development has historically lagged behind DA development in the atmosphere and ocean communities, despite the availability of long land surface and hydrology observational data records. In fact, multivariate and concurrent DA of land variables have only been reported in more recent years. An obvious beneficial impact from multi-sensor and multivariate DA setups has been in improving the temporal coverage of observational inputs within the DA environment. For example, since 2010, soil moisture retrievals from multiple passive and active microwave instruments have become available, which improves the temporal observational coverage from remote sensing. Similarly, vegetation conditions are available from both optical/thermal infrared and microwave sensors, with each contributing unique advantages (Section 2). Combining both types of vegetation measurements can help in providing a more continuous observational constraint with a DA environment (Kumar et al., 2020a; Mucia et al., 2021). This can be particularly important when temporally continuous coverage of land surface conditions is needed for monitoring of hazards such as fires and related vegetation disturbances (Kumar et al., 2021).

The concurrent handling of complementary information from multiple observation types is a significant challenge in multivariate DA environments (Montzka et al., 2012). Because soil moisture and snow are important in warm and cold seasons, respectively, the assimilation is, for the most part, mutually exclusive and thus relatively straightforward. The development of a reanalysis that incorporates soil and snow depth from multiple sensors across three decades has demonstrated improvements in land surface states, particularly with more modern sensors (Kumar et al., 2018). Resolving the disparities in temporal and spatial resolution differences across observational products is another challenge in multivariate DA systems (Tian et al., 2017). Generally, studies have reported added improvements from multivariate DA than those from corresponding univariate configurations (H. Su et al., 2010; Hain et al., 2012; Zhao & Yang, 2018; Abbaszadeh et al., 2020). Multivariate observational constraints may be necessary in instances where observations provide unique insights into related processes. For example, univariate assimilation of terrestrial water storage or soil moisture in places where groundwater pumping is prevalent may produce incorrect results, while their joint assimilation provides more consistent improvements across relevant water storage terms (Girotto et al., 2019; Tangdamrongsub et al., 2020). Given the simultaneous impact from anthropogenic processes and the interlinked nature of water cycle processes, emphasis on multivariate DA environments will be needed for highlighting the utility of land DA. There is also a significant need to develop land reanalysis products so that an observation-informed spatially and temporally continuous climate record of water, energy, and carbon cycle changes on the land surface can be established.

3.9 Coupled DA environments

Since the initial development of land DA was motivated by the need to improve NWP, most applications of land DA are limited to providing better initial conditions prior to a forecast. As noted earlier, such independent development of approaches can lead to divergent developmental priorities and inconsistent results with land and atmosphere systems. Coupling between the respective DA systems is needed to reduce the imbalance between them. Generally, there are two types of coupled DA systems (Penny & Hamill, 2017). In a weakly-coupled DA system, communication and data exchange between the land and the atmosphere only occurs through the physics within a coupled forecast model, which is the land DA approach commonly used in the context current NWP and reanalysis development (de Rosnay et al., 2013; Draper & Reichle, 2019; Carrera et al., 2019; Gomez et al., 2020; Reichle et al., 2021b). A strongly coupled DA environment, on the other hand, facilitates cross-component interactions during DA, to allow the use of cross-model error covariances during DA (Shahabadi et al., 2019). Though there is a recognition that strongly coupled DA systems enable increased exploitation of the information from observations and are better at addressing the imbalance issues, no such operational systems exist currently (Sluka et al., 2016). Coupled DA approaches are going to be pursued for their consistent surface-atmosphere initialization (for a "spatial-consistency" target), beneficial for NWP prediction or within a climate reanalysis. They, however, carry a substantially higher computational burden. Land-only DA systems are, therefore, likely to thrive and provide valuable and complementary solutions for frequent land reanalysis and permitting to initialize past reforecasts in support of anomaly-based predictions, such as those produced in the S2S framework (for a "time-consistency" target). While land DA and coupled DA approaches may have important methodological differences, they do face common challenges related to model and observation errors, characterization, handling of biases, quality control, compensating coupling errors (Vitart et al., 2019). Therefore several of the priorities identified hereafter are crosscutting.

4 Future priorities, new opportunities, and recommendations

Despite the challenges and gaps described above, DA methodologies are vital tools for realizing the promise of land surface observations. In this section, we outline the priorities for advancing land DA including the need for improvements in models and observations, co-development of models with the DA community, and embracing technological advancements, so that land DA environments can not only help to realize the full potential of land surface observations, but also help define the next generation observational needs. These enhancements will likely require an iterative process, and coordination across several interdisciplinary areas, as shown in Figure 2. A number of specific recommendations to realize this vision is outlined below.

4.1 Enable multivariate, multi-sensor DA

As noted earlier, the number and type of space-borne observations relevant to the land surface and hydrology continue to increase. This increased availability of observations presents unique challenges and opportunities for significantly advancing land and hydrologic prediction (McCabe et al., 2017). The simultaneous availability of multiple observation types of land surface processes provides concurrent water and carbon budget constraints, and has the potential to reduce errors in water and vegetation state estimates beyond what is achievable with numerical process models alone. This is particularly relevant in regions of the world where water cycle processes are dominated by anthropogenic impacts and univariate assimilation may provide misleading results.

The wealth of observational information available from multiple platforms is a unique opportunity to embrace a "systems approach" (Jenkins, 1969) to land DA that truly acknowledges the interconnected nature of water cycle processes and the human influence by enabling multivariate and multi-sensor environments. Focused site level analyses (e.g., Fox et al. (2018)) can be used to systematically diagnose the utility of specific observational inputs and to minimize representativeness errors. Such strategies can also be used to configure appropriate localization approaches to specify the extent of information transfer, and to reduce sampling errors from ensemble structures.

4.2 Develop coupled DA

In strongly coupled DA, multiple components from a coupled model are updated in a single DA step. Strongly coupled land-atmosphere DA expands the possibilities for multivariate and multisensor DA, and its potential advantages are being actively explored. There have been significant advances in the development of coupled Earth system models, through efforts such as the Earth System Modeling Framework (ESMF; www.earthsystemmodeling.org) that provide interoperable standards and tools for the community. Similar, concerted efforts to enable such paradigms (e.g. Joint Effort for Data Assimilation Infrastructure; JEDI; Tremolet and Auligne (2020)) and provide flexible and computationally scalable software infrastructure are needed to realize the vision of coupled Earth system DA environments (de Rosnay et al., 2022). Recent studies examining strongly coupled land-atmospheric DA systems have also stressed the challenges of addressing inconsistencies in spatial and temporal error correlation scales (across the land and atmosphere), the need for reliable forward modeling tools, and the recognition of the memory and timescale differences in land and atmospheric dynamics (Lin & Pu, 2020; Draper, 2021).

4.3 Improve the physical realism of LSMs

Many of the limitations of land DA environments are closely related to LSM deficiencies, as discussed in Section 3.2. Therefore, there needs to be a strong emphasis on improving the physical realism of LSM process descriptions and of retrieval products so that interpretation errors and efficiency loss issues can be reduced. Calibration of model parameters can be used to reduce systematic errors in models prior to DA - typically performed using optimization algorithms and treating the parameters as time-invariant (Zhou et al., 2020). Online DA tools have also been applied in this regard, simultaneously adjusting the model parameters and states. While computational cost is often prohibitive in employing these algorithms over large spatial extents and fine resolutions, more computationally efficient approaches founded in machine learning have been developed more recently (Tsai et al., 2021). Embracing these technological advances and improving the physical realism of the model and observations should reduce the reliance on rescaling approaches and enable the direct incorporation of observations within land DA systems. In addition to the parameters, LSM development must also include the refinement of the inherent coupling representations between relevant energy, water, and carbon cycle processes. Diagnostic studies that identified inconsistencies in coupling strength representations (Section 3.5) should serve as a guideline for such improvements. The same observational advances that enable the assimilation of higher-quality state and flux estimates into LSMs also provide opportunities for better assessing, and correcting, coupling strength biases in LSMs. If the objective of information transfer to related variables through DA is to be realized, such model process and parameter improvements are essential. Finally, LSM formulations must evolve to include relevant prognostics that can be employed in land DA environments. More broadly, there is a significant need to consider the requirements from assimilation needs in LSM development.

4.4 Enhance spaceborne observations

There is also a need for co-producing land observation datasets in coordination with the land DA community so that a consistent characterization of spatio-temporal errors of observations can be established. Given the prevalence and heterogeneity of bias in land modeling and assimilation, bias-aware DA algorithms that are built to diagnose and incorporate systematic errors during DA (as opposed to addressing them *a priori*) are needed (Dee, 2005; Pauwels et al., 2013; Ridler et al., 2017). However, such bias-aware systems are only practical if the climatology of the observations can be considered as the reference (Dong et al., 2020b). Efforts to reduce biases in land remote sensing retrievals or forward modeling approaches should be a priority to achieve these goals (Grant et al., 2008; Gao et al., 2020). Further, coordination across mission science teams and data producers to reduce errors and inconsistencies across products will be beneficial in improving the quality of observational products available for assimilation. Such data homogenization and harmonization efforts can reduce interpretation errors and foster the development of universally accepted observation benchmarks (McCabe et al., 2017). DA systems, in fact, offer natural environments for the harmonization of satellite data from multiple platforms, by appropriately handling the respective sources of errors and uncertainties. Land reanalysis and multivariate assimilation efforts incorporating multiple sources of data should be considered important goals for the land DA community. Such efforts are necessary to establish climate data records, which are important benchmarks for climate studies (C.-H. Su et al., 2016).

4.5 Exploit ground measurements

As noted in the introduction, the focus of this article is primarily on the assimilation of space-borne, spatially distributed measurements. Ground-based measurements of land surface variables such as soil moisture, snow, ET, and river flow are available from several measurement networks and focused field campaigns. These measurements have played and will continue to play an important role in enabling the systematic evaluation and optimization of DA analyses. Further coordination with the ground measurement community (e.g. the International Soil Moisture Network (ismn.geo.tuwien.ac.at), FLUXNET (fluxnet.org), and the U.S. National Ecological Observatory Network (NEON; www.neonscience.org)) to communicate the specific DA relevant needs is required to overcome some of the gaps discussed in Section 3. For example, reliable, co-located in-situ measurements of land surface states and fluxes (soil moisture, ET, runoff, LAI) can help in validating satellite data (e.g., SMOSREX (de Rosnay et al., 2006) and VALERI (Weiss et al., 2001)) and establishing true surface-subsurface coupling relationships, which are needed to improve the realism of such formulations in models and to ultimately improve the efficiency of DA methods. Similarly, focused ground-truthing efforts to characterize human management impacts (e.g., irrigation impact on soil moisture) would also be helpful in validating space-borne measurements over such areas. Much like the anticipated growth in the number of space-based platforms and measurements, the advent of new technologies is fostering an era of crowd-sourced observations from platforms such as mobile phones and inexpensive sensors that can be placed within a virtual network through them. For example, crowd-sourcing via smartphone apps and GPS devices has revolutionized how traffic information is collected and updated in real-time. While there are still challenges in ensuring the reliability and quality of such measurements for science applications, they offer new possibilities not just for evaluation, but also for DA, particularly if their availability becomes routine and spatially distributed.

4.6 Expand forward modeling capabilities

Advances in forward modeling capabilities are needed to reduce the reliance on retrieval products and to enhance the exploitation of observations in coupled DA systems. In addition to radiative transfer models, the land DA community could also leverage advances in machine learning to develop observation operators. Even if the relevant radiative transfer model is available for the translation of a certain type of observation, ensuring their reliability over different types of terrestrial, biosphere, and climatic conditions often requires significant calibration and validation efforts. The modern class of machine learning methods such as deep learning has shown great promise in its ability to perform hierarchical feature extraction and could effectively function as a forward operator within DA systems (e.g., Shan et al. (2022)). Generally, machine learning models are often criticized for their non-transparent nature, a hindrance to developing physical understanding and insights into the results generated by them. Recent advancements have also enabled 'explainable AI' methods (Montavon et al., 2018) that have specifically focused on increasing transparency of the decisions and inner workings of the machine learning algorithms. These methods are useful for providing improved understanding of the sensitivities of the inputs, which can lead to improved information exploitation to produce more skillful retrieval products or using such translations directly within DA systems. These methods can also be applied to understand when and where remote sensing measurements provide relevant information content, which can then be used in DA environments to screen observations and refine the quality control procedures to improve the observational information utilization. In addition to providing a statistically robust model, machine learning-based emulators can help in reducing the computational expense of retrieval and/or forward modeling (Rodriguez-Fernandez et al., 2015; Kolassa et al., 2018).

4.7 Incorporate human-driven land processes

A unique challenge and opportunity for land surface DA is to account for human activities for which we either do not have enough knowledge to implement process representations in models or which are inherently impossible to implement due to their subjective nature. As discussed earlier, the identification and inclusion of such signals and their attribution to the systematic differences stemming from them are needed in such instances. The use of machine learning methods can be potentially useful in this regard. Machine learning methods are also increasingly used for object detection and feature extraction applications (LeCun et al., 2015; Bengio, 2009), where the algorithm automatically learns and characterizes features from the data that are presented to it. As these methods are less 'feature engineering' dependent, they can be an effective tool for identifying unmodeled features and characterizing the sources of bias in land observations and models. Taking advantage of such capabilities and incorporating machine learning-based modernizations (Reichstein et al., 2019) are likely to result in better characterization of anthropogenic signals.

4.8 Exploit new space-borne observation technologies

With the availability of remote sensing opportunities afforded by smallsats and cubesats, commercial platforms for Earth observations are becoming more prevalent (Sandau, 2010; Llop et al., 2014). While space-agency missions provide long-term coverage for critical Earth processes, they tend to be costly and often require decades of development. The smallsats are less costly and can provide continuous coverage at fine resolutions, when launched in multiples. Additionally, other sensing opportunities from smart phones, unmanned aerial vehicles, and other close-to-Earth platforms are going to increase, fostering the era of "big data" (Sellars et al., 2013). While these measurements could be exploited within DA environments, they also likely present additional technology hurdles. As the size of the datasets become larger, there is an increasing shift to using cloud-based resources away from local computing resources. Effective and consistent quality control of these measurements from such distributed platforms can also be challenging. Modeling and DA systems must anticipate and adapt to such changing technologies if the significant potential of such datasets is to be realized.

4.9 Employ land DA in mission planning efforts

Land DA efforts have a critical role to play in the definition of the next generation of observational types and needs. Given the considerable resources required to implement Earth observing missions, their utility must be systematically assessed through observing system simulation experiments (OSSEs), prior to instrument development. DA is key to enabling OSSEs, which can help assess the relative utility of competing mission designs (e.g. Tan et al. (2007); Kaminski and Mathieu (2017)). Mature land DA environments can be used in this regard for the sub-selection of mission configurations and technologies and to define the level of accuracy required from the observations to meet science utility. Comprehensive multi-observation OSSEs can also help to identify the spatio-temporal gaps in the types of observations that are needed to improve the characterization of the water cycle. For example, snow and soil moisture processes may be seasonally dominant for water availability characterization in a region, but the accuracy, revisit, and spatial resolution requirements of those observations are likely to be different for water balance closure or water availability characterization needs. By enabling 'what-if' studies, OSSEs can formalize such requirements and influence the recommendations and definitions of future land hydrology observing systems. In a pioneering effort, the SMAP mission was informed by land DA-based OSSEs in its formulation phase (Crow et al., 2005b; Reichle et al., 2008a) and has been operationally generating a soil moisture assimilation product (Reichle et al., 2019). OSSEs have been used to quantify the anticipated utility of snow measurements from different sensors and technologies to guide the development of future snow missions (Garnaud et al., 2019; Kwon et al., 2021; Wrzesien et al., 2022). Similar land DA environments and efforts must be considered part of the lifecycle of future land surface satellite missions.

5 Summary and Conclusions

Land DA has been an active area of research for some decades now and has shown significant promise for improving land surface characterization and benefiting NWP, S2S and operational hydrology forecasting environments. However, to realize its full potential, the land DA community, in conjunction with the relevant observational and modeling communities should address the significant gaps and challenges laid out in this article. Separate configurations of land DA environments specifically tailored for different applications will likely be needed to overcome the constraints of developmental legacies, lack of model realism, and to foster the development and widespread acceptance of land DA as a necessary capability. We hope this agenda article will serve as the pathway for future land DA developmental priorities so that land DA systems will evolve into viable environments for addressing the critical science and water resource problems facing society through the exploitation of observational information.

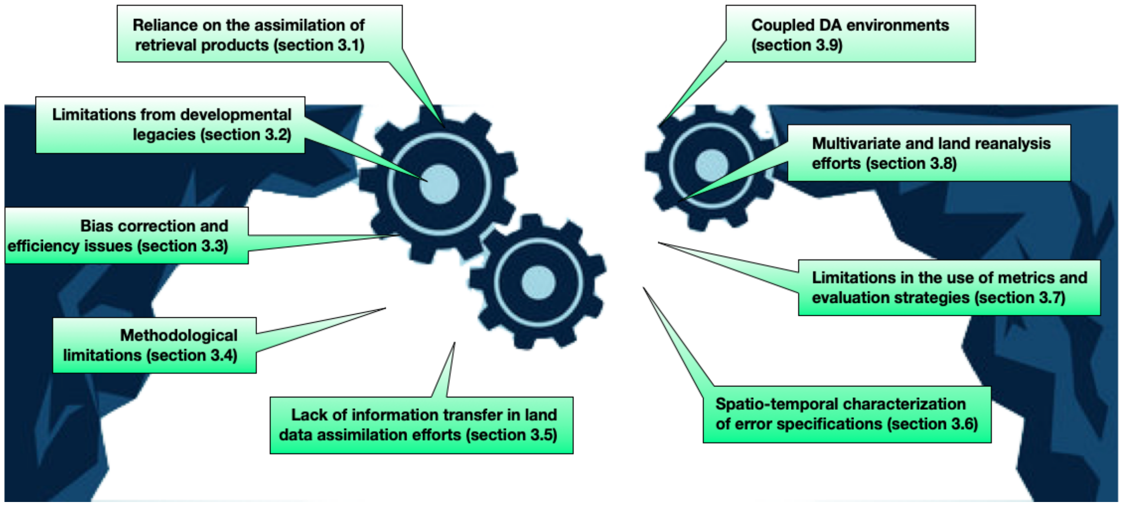

Figure 1. List of the major challenges and gaps in land data assimilation discussed in this article.

1053 ## Acronyms

1054 **DA** Data Assimilation

1055 **EO** Earth Observation

1056 **LSM** Land Surface Model

1057 **NWP** Numerical Weather Prediction

1058 **S2S** Sub-seasonal to Seasonal

1059 **GPS** Global Positioning System

1060 **GPM** Global Precipitation Measurement

1061 **SMAP** Soil Moisture Active Passive

1062 **SMOS** Soil Moisture Ocean Salinity

1063 **GRACE-FO** Gravity Recovery And Climate Experiment - Follow On

1064 **GRACE** Gravity Recovery And Climate Experiment

1065 **OCO-2** Orbiting Carbon Observatory-2

1066 **SWOT** Surface Water and Ocean Topography

1067 **EKF** Extended Kalman Filter

1068 **EnKF** Ensemble Kalman Filter

1069 **SSMI/S** Special Sensor Microwave Imager/Sounder

1070 **SWE** Snow Water Equivalent

1071 **GLDAS** Global Land Data Assimilation System

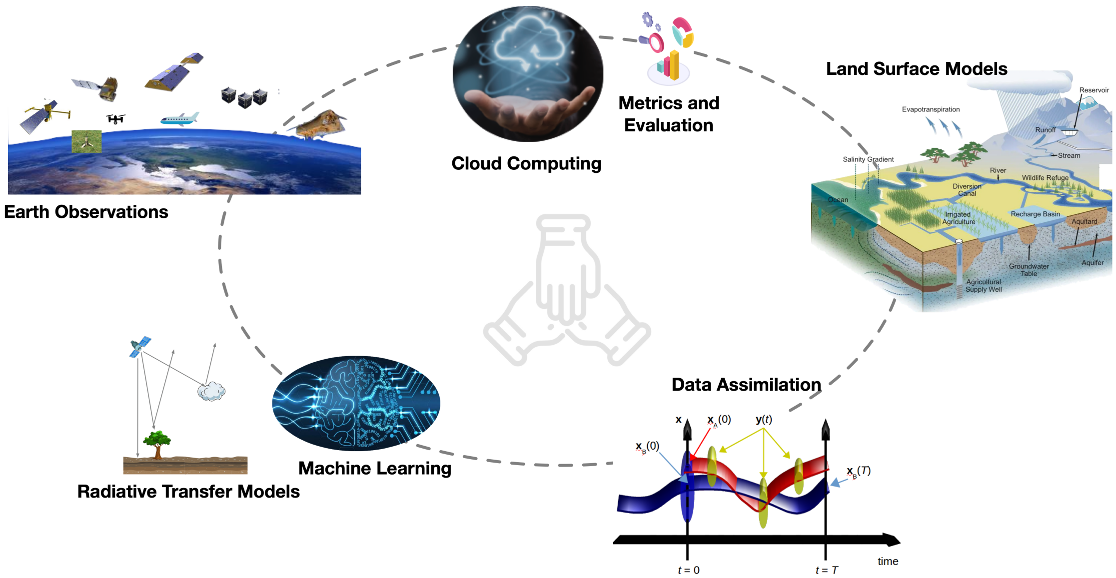

Figure 2. A vision for the next generation land DA systems, with closer developmental coordination with the LSM and observation communities, adoption of technological enhancements in machine learning and cloud computing, and advancement in the use of metrics and evaluation strategies to realize the potential of land surface observations.

NLDAS North American Land Data Assimilation System

LIS Land Information System

CDF Cumulative Distribution Function

ET Evapotranspiration

ESMF Earth System Modeling Framework

JEDI Joint Effort for Data Assimilation Infrastructure

OSSE Observing System Simulation Experiment

NASA National Aeronautics and Space Administration

USDA-ARS United States Department of Agriculture - Agricultural Research Service

NOAA National Oceanic and Atmospheric Administration

6 Open Research

This article is focused on discussing the challenges and priorities in the field of land data assimilation and does not include the specific use of any particular software or results involving specific data products.

Acknowledgments

This manuscript is in part an outcome of an Analysis, Integration, and Modeling of the Earth System (AIMES) workshop "Tackling Technical Challenges in Land Data Assimilation" held virtually in June 2021. In addition, the workshops organized by the International Earth Surface Working Group in past years (2017-2022) have also contributed to the formulation of this article.

References

Abbaszadeh, P., Gavahi, K., & Moradkhani, H. (2020). Multivariate remotely sensed and in-situ data assimilation for enhancing community WRF-Hydro model forecasting. *Advances in Water Resources*, *145*, 103721. doi: 10.1016/j.advwatres.2020.103721

Aires, F., Weston, P., de Rosnay, P., & Fairbairn, D. (2021). Statistical approaches to assimilate ASCAT soil moisture information -I. Methodologies and first assessment. *Quarterly Journal of the Royal Meteorological Society*, *147*(736), 1823–1852.

Albergel, C., Munier, S., Leroux, D., Dewaele, H., Fairbairn, D., & Barbu, A. (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.

Albergel, C., Zheng, Y., Bonan, B., Dutra, E., Rodriguez-Fernandez, N., & Munier, S. (2020). Data assimilation for continuous global assessment of severe conditions over terrestrial surfaces. *Hydrology and Earth System Sciences*, *24*(9), 4291–4316.

Anderson, J. (2007). Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D: Nonlinear Phenomena*, *230*(1), 99-111. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0167278906002168` (Data Assimilation) doi: https://doi.org/10.1016/j.physd.2006.02.011

Anderson, J., & Anderson, S. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, *127*(12), 2741–2758.

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Avellano, A. (2009). The Data Assimilation Research Testbed: A Community Facility. *Bulletin of the American Meteorological Society*, *90*(9), 1283 - 1296. Retrieved from `https://journals.ametsoc.org/view/journals/bams/90/9/2009bams2618_1.xml` doi: 10.1175/2009BAMS2618.1

Andreadis, K., Clark, E., Lettenmaier, D., & Alsdorf, D. (2007). Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, *34*(10).

Andreadis, K., & Lettenmaier, D. (2006). Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, *29*(6), 872–886. doi: 10.1016/j.advwatres.2005.08.004

Arsenault, K., Houser, P., De Lannoy, G., & Dirmeyer, P. (2013). Impacts of snow cover fraction data assimilation on modeled energy and moisture budgets. *Journal of Geophysical Research: Atmospheres*, *118*(14), 7489–7504. doi: 10.1002/jgrd.50542

Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., ... Zeng, X. (2018). Satellite and In Situ Observations for Advancing Global Earth Surface Modelling: A Review. *Remote Sensing*, *10*(12). Retrieved from https://www.mdpi.com/2072-4292/10/12/2038 doi: 10.3390/rs10122038

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., ... Vitart, F. (2015). ERA-Interim/Land: a global land surface reanalysis data set. *Hydrology and Earth System Sciences*, *19*(1), 389–407.

Balsamo, G., Bouyssel, F., & Noilhan, J. (2004). A simplified bi-dimensional variational analysis of soil moisture from screen-level observations in a mesoscale numerical weather-prediction model. *Quarterly Journal of the Royal Meteorological Society*, *130*(598), 895–915. doi: 10.1256/qj.02.215

Balsamo, G., Mahfouf, J.-F., Belair, S., & Deblonde, G. (2007). A Land Data Assimilation System for Soil Moisture and Temperature: An Information Content Study. *Journal of Hydrometeorology*, *8*(6), 1225–1242. doi: 10.1175/2007JHM819.1

Bannister, R. (2017). A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 607–633. doi: 10.1002/qj.2982

Barbu, A., Calvet, J.-C., Mahfouf, J.-F., Albergel, C., & Lafont, S. (2011). Assimilation of Soil Wetness Index and Leaf Area Index into the ISBA-A-gs land surface model: grassland case study. *Biogeosciences*, *8*(7), 1971–1986.

Barnett, T., Adam, J., & Lettenmaier, D. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303–309. doi: 10.1038/nature04141

Barrett, A. (2003). *National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC.* National Snow and Ice Data Center. (Cooperative Institute for Research in Environmental Sciences)

Baugh, C., de Rosnay, P., Lawrence, H., Jurlina, T., Drusch, M., Zsoter, E., & Prudhomme, C. (2020). The Impact of SMOS Soil Moisture Data Assimilation within the Operational Global Flood Awareness System (GloFAS). *Remote Sensing*. doi: 10.3390/rs12091490

Bengio, Y. (2009). Learning Deep Architectures. *Foundations and Trends in Machine Learning*, *2*(1), 1–127.

Benjamin, S., Brown, J., Brunet, G., Lynch, P., Saito, K., & Schlatter, T. (2018). 100 Years of Progress in Forecasting and NWP Applications. *Meteorological Monographs*, *59*, 13.1–13.67. doi: 10.1175/AMSMONOGRAPHS-D-18-0020.1

Benjamin, S., Smirnova, T., James, E., Lin, L.-F., Hu, M., Turner, D., & He, S. (2022). Land-snow data assimilation including a moderately coupled initialization method applied to NWP. *Journal of Hydrometeorology*. doi: 10.1175/JHM-D-21-0198.1

Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S., & Hurk, B. (2016). Land-atmosphere feedbacks amplify aridity increase over land under global warming. *Nature Climate Change*, *6*(9), 869–874. doi: 10.1038/nclimate3029

Biancamaria, S., Lettenmaier, D., & Pavelsky, T. (2016). The SWOT mission and its capabilities for land hydrology. In *Remote sensing and water resources* (pp. 117–147). Springer.

Bonan, B., Albergel, C., Zheng, Y., Barbu, A., Fairbairn, D., Munier, S., & Calvet, J.-C. (2020). An ensemble square root filter for the joint assimilation of surface soil moisture and leaf area index within the Land Data Assimilation System LDAS-Monde: application over the Euro-Mediterranean region. *Hydrology and Earth System Sciences*, *24*(1), 325–347.

Bosilovich, M., Radakovich, J., Silva, A., Todling, R., & Verter, F. (2007). Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system. *Journal of the Meteorological Society of Japan. Ser. II*, *85A*, 205–228. Retrieved from `https://doi.org/10.2151/jmsj.85A.205` doi: 10.2151/jmsj.85A.205

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

Bouttier, F., & Courtier, P. (2002). *Data assimilation concepts and methods March 1999* (Vol. 718).

Brown, B., Jensen, T., Gotway, J., Bullock, R., Gilleland, E., & Fowler, T. (2021). The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bulletin of the American Meteorological Society*, *102*(4), E782-E807.

Canadell, J., Monteiro, P., Costa, M., Cunha, L., Cox, P., & Alexey, V. (2021). Global carbon and other biogeochemical cycles and feedbacks. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

Caparrini, F., Castelli, F., & Entekhabi, D. (2004). Variational estimation of soil and vegetation turbulent transfer and heat flux parameters from sequences of multisensor imagery. *Water Resources Research*, *40*(12).

Cardinali, C. (2009). Monitoring the observation impact on the short-range forecast. *Quarterly Journal of the Royal Meteorological Society*, *135*(638), 239–250. doi: 10.1002/qj.366

Carrera, M., Belair, S., & Bilodeau, B. (2015). The Canadian Land Data Assimilation System (CaLDAS): Description and Synthetic Evaluation Study. *Journal of Hydrometeorology*, *16*(3), 1293–1314. doi: 10.1175/JHM-D-14-0089.1

Carrera, M., Bilodeau, B., Belair, S., Abrahamowicz, M., Russell, A., & Wang, X. (2019). Assimilation of Passive L-band Microwave Brightness Temperatures in the Canadian Land Data Assimilation System: Impacts on Short-Range Warm Season Numerical Weather Prediction. *Journal of Hydrometeorology*, *20*(6), 1053–1079. doi: 10.1175/JHM-D-18-0133.1

Castelli, F., Entekhabi, D., & Caporali, E. (1999). Estimation of surface heat flux and an index of soil moisture using adjoint-state surface energy balance. *Water Resources Research*, *35*(10), 3115–3125. doi: 10.1029/1999WR900140

Cawse-Nicholson, K., Townsend, P., Schimel, D., Assiri, A., Blake, P., & Buongiorno, M. (2021). NASA's surface biology and geology designated observable: A perspective on surface imaging algorithms. *Remote Sensing of Environment*, *257*, 112349. doi: 10.1016/j.rse.2021.112349

Chen, F., Crow, W., Cosh, M., Colliander, A., Asanuma, J., & Berg, A. (2019).

Uncertainty of Reference Pixel Soil Moisture Averages Sampled at SMAP Core Validation Sites. *Journal of Hydrometeorology*, *20*(8), 1553–1569. doi: 10.1175/JHM-D-19-0049.1

Clark, M., Bierkens, M., Samaniego, L., Woods, R., Uijlenhoet, R., Bennett, K., ... Peters-Lidard, C. (2017). The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. *Hydrol. Earth Syst. Sci*, *21*(7), 3427–3440. doi: 10.5194/hess-21-3427-2017

Clark, M., Rupp, D., Woods, R., Zheng, X., Ibbitt, R., & Slater, A. (2008). Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to upyear states in a distributed hydrological model. *Advances in Water Resources*, *31*(10), 1309–1324.

Clark, M., Slater, A., Barrett, A., Hay, L., McCabe, G., Rajagopalan, B., & Leavesley, G. (2006). Assimilation of snow covered area information into hydrologic and land-surface models. *Advances in Water Resources*, *29*(8), 1209–1221. doi: 10.1016/j.advwatres.2005.10.001

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., ... Booth, B. (2013). Long-term climate change: projections, commitments and irreversibility. In *Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 1029–1136). Cambridge University Press.

Cook, B., Mankin, J., & Anchukaitis, K. (2018). Climate Change and Drought: From Past to Future. *Current Climate Change Reports*, *4*(2), 164–179. doi: 10.1007/s40641-018-0093-2

Courtier, P., Derber, J., Errico, R., Louis, J.-F., & VukiCEviC, T. (1993). Important literature on the use of adjoint, variational methods and the Kalman filter in meteorology. *Tellus A: Dynamic Meteorology and Oceanography*, *45*(5), 342–357. doi: 10.3402/tellusa.v45i5.14898

Crow, W., & Berg, M. (2010). An improved approach for estimating observation and model error parameters in soil moisture data assimilation. *Water Resources Research*, *46*(12). doi: 10.1029/2010WR009402

Crow, W., Chan, S., Entekhabi, D., Houser, P., Hsu, A., & Jackson, T. (2005b). An observing system simulation experiment for hydros radiometer-only soil moisture products. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(6), 1289–1303. doi: 10.1109/TGRS.2005.845645

Crow, W., Chen, F., Reichle, R., & Liu, Q. (2017). L band microwave remote sensing and land data assimilation improve the representation of prestorm soil moisture conditions for hydrologic forecasting. *Geophysical Research Letters*, *44*(11), 5495–5503. doi: 10.1002/2017GL073642

Crow, W., Chen, F., Reichle, R., & Xia, Y. (2019). Diagnosing Bias in Modeled Soil Moisture/Runoff Coefficient Correlation Using the SMAP Level 4 Soil Moisture Product. *Water Resources Research*, *55*(8), 7010–7026. doi: 10.1029/2019WR025245

Crow, W., Chen, F., Reichle, R., Xia, Y., & Liu, Q. (2018). Exploiting Soil Moisture, Precipitation, and Streamflow Observations to Evaluate Soil Moisture/Runoff Coupling in Land Surface Models. *Geophysical Research Letters*, *45*(10), 4869–4878. doi: 10.1029/2018GL077193

Crow, W., Gomez, C., Sabater, J., Holmes, T., Hain, C., & Lei, F. (2020). Soil Moisture-Evapotranspiration Overcoupling and L-Band Brightness Temperature Assimilation: Sources and Forecast Implications. *Journal of Hydrometeorology*, *21*(10), 2359–2374. doi: 10.1175/JHM-D-20-0088.1

Crow, W., Koster, R., Reichle, R., & Sharif, H. (2005a). Relevance of time-varying and time-invariant retrieval error sources on the utility of spaceborne soil moisture products. *Geophysical Research Letters*, *32*(24).

Crow, W., & Loon, E. (2006). Impact of Incorrect Model Error Assumptions on the Sequential Assimilation of Remotely Sensed Surface Soil Moisture. *Journal of Hydrometeorology*, *7*(3), 421–432. doi: 10.1175/JHM499.1

Crow, W., & Reichle, R. (2008). Comparison of adaptive filtering techniques for land surface data assimilation. *Water Resources Research*, *44*(8). doi: 10.1029/2008WR006883

Dash, J., & Ogutu, B. (2016). Recent advances in space-borne optical remote sensing systems for monitoring global terrestrial ecosystems. *Progress in Physical Geography*, *40*(2), 322–351.

de Rosnay, P., Calvet, J.-C., Kerr, Y., Wigneron, J.-P., Lemaitre, F., Escorihuela, M. J., . . . Waldteufel, P. (2006). SMOSREX: A long term field campaign experiment for soil moisture and land surface processes remote sensing. *Remote Sensing of Environment*, *102*(3), 377-389. Retrieved from https://

www.sciencedirect.com/science/article/pii/S0034425706000976 doi: https://doi.org/10.1016/j.rse.2006.02.021

DeChant, C., & Moradkhani, H. (2011). Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrology and Earth System Sciences*, *15*(11), 3399–3410.

Dee, D. (2005). Bias and data assimilation. *Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography*, *131*(613), 3323–3343.

De Lannoy, G., Bechtold, M., Albergel, C., Brocca, L., Calvet, J.-C., Carrassi, A., . . . Steele-Dunne, S. (2022). Perspective on satellite-based land data assimilation to estimate water cycle components in an era of advanced data availability and model sophistication. *Frontiers in Water*, *4*. Retrieved from https://www.frontiersin.org/articles/10.3389/frwa.2022.981745 doi: 10.3389/frwa.2022.981745

De Lannoy, G., de Rosnay, P., & Reichle, R. (2019). *Soil moisture data assimilation*. Springer Nature.

De Lannoy, G., & Reichle, R. (2016a). Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model. *Hydrology and Earth System Sciences*, *20*(12), 4895–4911.

De Lannoy, G., & Reichle, R. (2016b). 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.

De Lannoy, G., Reichle, R., Arsenault, K., Houser, P., Kumar, S., Verhoest, N., & Pauwels, V. (2012). Multiscale assimilation of Advanced Microwave Scanning Radiometer-EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado. *Water Resources Research*, *48*(1).

De Lannoy, G., Reichle, R., Houser, P., Pauwels, V., & Verhoest, N. (2007). Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. *Water Resources Research*, *43*(9). doi: 10.1029/2006WR005449

De Lannoy, G., Reichle, R., & Pauwels, V. (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.

Demarty, J., Chevallier, F., Friend, A., Viovy, N., Piao, S., & Ciais, P. (2007). Assimilation of global MODIS leaf area index retrievals within a terrestrial biosphere model. *Geophysical Research Letters*, *34*(15).

Dente, L., Satalino, G., Mattia, F., & Rinaldi, M. (2008). Assimilation of leaf area index derived from ASAR and MERIS data into CERES-Wheat model to map wheat yield. *Remote Sensing of Environment*, *112*(4), 1395–1407. doi: 10.1016/j.rse.2007.05.023

de Rosnay, P., Balsamo, G., Albergel, C., Munoz-Sabater, J., & Isaksen, L. (2014). Initialisation of Land Surface Variables for Numerical Weather Prediction. *Surveys in Geophysics*, *35*(3), 607–621. Retrieved from `https://doi.org/10.1007/s10712-012-9207-x` doi: 10.1007/s10712-012-9207-x

de Rosnay, P., Browne, P., Boisseson, D., Fairbairn, D., Hirahara, Y., & Ochi, K. (2022). Coupled assimilation at ECMWF: current status, challenges and future developments. *Quarterly Journal of the Royal Meteorological Society*. (in press.) doi: 10.1002/qj.4330

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., & Isaksen, L. (2013). A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, *139*(674), 1199–1213. doi: 10.1002/qj.2023

de Rosnay, P., Munoz-Sabater, J., Albergel, C., Isaksen, L., English, S., Drusch, M., & Wigneron, J.-P. (2020). SMOS brightness temperature forward modelling and long term monitoring at ECMWF. *Remote Sensing of Environment*, *237*, 111424. doi: 10.1016/j.rse.2019.111424

Dirmeyer, P. (2018). Coupled from the start. *Eos, Transactions American Geophysical Union*, *99*. doi: 10.1029/2018EO095367

Dirmeyer, P., Balsamo, G., Blyth, E., Morrison, R., & Cooper, H. (2021). Land-Atmosphere Interactions Exacerbated the Drought and Heatwave Over Northern Europe During Summer 2018. *AGU Advances*, *2*(2), e2020AV000283. doi: 10.1029/2020AV000283

Dirmeyer, P., & Halder, S. (2016). Sensitivity of Numerical Weather Forecasts to Initial Soil Moisture Variations in CFSv2. *Weather and Forecasting*, *31*(6), 1973–1983. doi: 10.1175/WAF-D-16-0049.1

Dirmeyer, P., Wu, J., Norton, H., Dorigo, W., Quiring, S., & Ford, T. (2016). Confronting Weather and Climate Models with Observational Data from Soil Moisture Networks over the United States. *Journal of Hydrometeorology*, *17*(4), 1049–1067. doi: 10.1175/JHM-D-15-0196.1

Dokoohaki, H., Morrison, B. D., Raiho, A., Serbin, S. P., Zarada, K., Dramko, L., & Dietze, M. (2022). Development of an open-source regional data assimilation system in PEcAn v. 1.7.2: application to carbon cycle reanalysis across the contiguous US using SIPNET. *Geoscientific Model Development*, *15*(8), 3233–3252. Retrieved from `https://gmd.copernicus.org/articles/15/3233/2022/` doi: 10.5194/gmd-15-3233-2022

Dong, J., Crow, W., Reichle, R., Liu, Q., Lei, F., & Cosh, M. (2019). A global assessment of added value in the SMAP Level 4 soil moisture product relative to its baseline land surface model. *Geophysical Research Letters*, *46*(12), 6604–6613.

Dong, J., Crow, W., Tobin, K., Cosh, M., Bosch, D., & Starks, P. (2020b). Comparison of microwave remote sensing and land surface modeling for surface soil moisture climatology estimation. *Remote Sensing of Environment*, *242*, 111756. doi: 10.1016/j.rse.2020.111756

Dong, J., Dirmeyer, P., Lei, F., Anderson, M., Holmes, T., Hain, C., & Crow, W. (2020a). Soil Evaporation Stress Determines Soil Moisture-Evapotranspiration Coupling Strength in Land Surface Modeling. *Geophysical Research Letters*, *47*(21), e2020GL090391. doi: 10.1029/2020GL090391

Dong, J., Steele-Dunne, S., Judge, J., & Giesen, N. (2015). A particle batch smoother for soil moisture estimation using soil temperature observations. *Advances in Water Resources*, *83*, 111–122. doi: 10.1016/j.advwatres.2015.05.017

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., & Naeimi, V. (2010). Error characterisation of global active and passive microwave soil moisture datasets. *Hydrology and Earth System Sciences*, *14*(12), 2605–2616. Retrieved from `https://hess.copernicus.org/articles/14/2605/2010/` doi: 10.5194/hess-14-2605-2010

Doughty, R., Kurosu, T., Parazoo, N., Kohler, P., Wang, Y., Sun, Y., & Frankenberg, C. (2022). Global GOSAT, OCO-2, and OCO-3 solar-induced chlorophyll fluorescence datasets. *Earth System Science Data*, *14*(4), 1513–1529.

Douville, H., Viterbo, P., Mahfouf, J.-F., & Beljaars, A. (2000). Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data. *Monthly Weather Review*, *128*(6), 1733–1756.

Draper, C. (2021). Accounting for Land Model Uncertainty in Numerical Weather Prediction Ensemble Systems: Toward Ensemble-Based Coupled Land-Atmosphere Data Assimilation. *Journal of Hydrometeorology*, *22*(8), 2089–2104. doi: 10.1175/JHM-D-21-0016.1

Draper, C., Mahfouf, J., & Walker, J. (2009). An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. *Journal of Geophysical Research: Atmospheres*, *114*(D20).

Draper, C., Mahfouf, J.-F., & Walker, J. (2011). Root zone soil moisture from the assimilation of screen-level variables and remotely sensed soil moisture. *Journal of Geophysical Research: Atmospheres*, *116*(D2). doi: 10.1029/2010JD013829

Draper, C., & Reichle, R. (2019). Assimilation of satellite soil moisture for improved atmospheric reanalyses. *Monthly Weather Review*, *147*(6), 2163–2188.

Draper, C., Reichle, R., De Lannoy, G., & Liu, Q. (2012). Assimilation of passive and active microwave soil moisture retrievals. *Geophysical Research Letters*, *39*(4). doi: 10.1029/2011GL050655

Draper, C., Reichle, R., De Lannoy, G., & Scarino, B. (2015). A Dynamic Approach to Addressing Observation-Minus-Forecast Bias in a Land Surface Skin Temperature Data Assimilation System. *Journal of Hydrometeorology*, *16*(1), 449–464. doi: 10.1175/JHM-D-14-0087.1

Drusch, M. (2007). Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set. *Journal of Geophysical Research: Atmospheres*, *112*(D3). doi: 10.1029/2006JD007478

Drusch, M., & Viterbo, P. (2007). Assimilation of screen-level variables in ECMWF's Integrated Forecast System: A study on the impact on the forecast quality and analyzed soil moisture. *Monthly Weather Review*, *135*(2), 300–314.

Durand, M., Barros, A., Dozier, J., Adler, R., Cooley, S., Entekhabi, D., … Steele-Dunne, S. (2021). Achieving Breakthroughs in Global Hydrologic Science by Unlocking the Power of Multisensor, Multidisciplinary Earth Observations. *AGU Advances*, *2*(4), e2021AV000455. doi: 10.1029/2021AV000455

Durand, M., Molotch, N., & Margulis, S. (2008). A Bayesian approach to snow water equivalent reconstruction. *Journal of Geophysical Research: Atmospheres*, *113*(D20). doi: 10.1029/2008JD009894

Ebert, E., Wilson, L., Weigel, A., Mittermaier, M., Nurmi, P., & Gill, P. (2013). Progress and challenges in forecast verification. *Meteorological Applications*, *20*(2), 130–139.

El Gharamti, M., McCreight, J. L., Noh, S. J., Hoar, T. J., RafieeiNasab, A., & Johnson, B. K. (2021). Ensemble streamflow data assimilation using WRF-Hydro and DART: novel localization and inflation techniques applied to Hurricane Florence flooding. *Hydrology and Earth System Sciences*, *25*(9), 5315–5336. Retrieved from https://hess.copernicus.org/articles/25/5315/2021/ doi: 10.5194/hess-25-5315-2021

Eltahir, E. (1998). A Soil Moisture-Rainfall Feedback Mechanism: 1. Theory and observations. *Water Resources Research*, *34*(4), 765–776. doi: 10.1029/97WR03499

Emery, C., Biancamaria, S., Boone, A., Ricci, S., Rochoux, M., Pedinotti, V., & David, C. (2020). Assimilation of wide-swath altimetry water elevation anomalies to correct large-scale river routing model parameters. *Hydrology and Earth System Sciences*, *24*(5), 2207–2233.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., … Van Zyl, J. (2010a). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, *98*(5), 704–716.

Entekhabi, D., Reichle, R., Koster, R., & Crow, W. (2010b). Performance metrics for soil moisture retrievals and application requirements. *Journal of Hydrometeorology*, *11*(3), 832–840.

Ercolani, G., & Castelli, F. (2017). Variational assimilation of streamflow data in distributed flood forecasting. *Water Resources Research*, *53*(1), 158–183.

Erlingis, J., Rodell, M., Peters-Lidard, C., Li, B., Kumar, S., & Famiglietti, J. (2021). A High-Resolution Land Data Assimilation System Optimized for the Western United States. *JAWRA Journal of the American Water Resources Association*, *57*(5), 692–710. doi: 10.1111/1752-1688.12910

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), 343–367.

Evensen, G., & Leeuwen, P. (2000). An Ensemble Kalman Smoother for Nonlinear Dynamics. *Monthly Weather Review*, *128*(6), 1852–1867. doi: 10.1175/1520

-0493(2000)128

Fairbairn, D., Barbu, A., Napoly, A., Albergel, C., Mahfouf, J.-F., & Calvet, J.-C. (2017). The effect of satellite-derived surface soil moisture and leaf area index land data assimilation on streamflow simulations over France. *Hydrology and Earth System Sciences*, *21*(4), 2015–2033.

Fairbairn, D., de Rosnay, P., & Browne, P. (2019). The New Stand-Alone Surface Analysis at ECMWF: Implications for Land-Atmosphere DA Coupling. *Journal of Hydrometeorology*, *20*(10), 2023–2042. doi: 10.1175/JHM-D-19-0074.1

Feldman, A., Short Gianotti, D., Konings, A., Gentine, P., & Entekhabi, D. (2021). Patterns of plant rehydration and growth following pulses of soil moisture availability. *Biogeosciences*, *18*(3), 831–847.

Fisher, R., & Koven, C. (2020). Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), 2018 001453. doi: 10.1029/2018MS001453

Flechtner, F., Morton, P., Watkins, M., & Webb, F. (2014). Status of the GRACE follow-on mission. In *Gravity, geoid and height systems* (pp. 117–121). Springer.

Fox, A., Hoar, T., Anderson, J., Arellano, A., Smith, W., & Litvak, M. (2018). Evaluation of a Data Assimilation System for Land Surface Models using CLM4.5. *Journal of Advances in Modeling Earth Systems*. doi: 10.1029/2018MS001362

Gao, L., Sadeghi, M., Ebtehaj, A., & Wigneron, J.-P. (2020). A temporal polarization ratio algorithm for calibration-free retrieval of soil moisture at L-band. *Remote Sensing of Environment*, *249*, 112019. doi: 10.1016/j.rse.2020.112019

Garnaud, C., Belair, S., Carrera, M., Derksen, C., Bilodeau, B., & Abrahamowicz, M. (2019). Quantifying Snow Mass Mission Concept Trade-Offs Using an Observing System Simulation Experiment. *Journal of Hydrometeorology*, *20*(1), 155–173. doi: 10.1175/JHM-D-17-0241.1

Getirana, A., Rodell, M., Kumar, S., Beaudoing, H., Arsenault, K., & Zaitchik, B. (2020). GRACE Improves Seasonal Groundwater Forecast Initialization over the United States. *Journal of Hydrometeorology*, *21*(1), 59–71. doi: 10.1175/JHM-D-19-0096.1

Gharamti, M. E. (2018). Enhanced Adaptive Inflation Algorithm for Ensemble Filters. *Monthly Weather Review*, *146*(2), 623 - 640. Retrieved from `https://`

journals.ametsoc.org/view/journals/mwre/146/2/mwr-d-17-0187.1.xml doi: 10.1175/MWR-D-17-0187.1

Gharamti, M. E. (2021). Hybrid Ensemble-Variational Filter: A Spatially and Temporally Varying Adaptive Algorithm to Estimate Relative Weighting. *Monthly Weather Review*, *149*(1), 65 - 76. Retrieved from https://journals.ametsoc.org/view/journals/mwre/149/1/mwr-d-20-0101.1.xml doi: 10.1175/MWR-D-20-0101.1

Ghent, D., Kaduk, J., Remedios, J., Ardo, J., & Balzter, H. (2010). Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. *Journal of Geophysical Research: Atmospheres*, *115*(D19). doi: 10.1029/2010JD014392

Ghil, M., & Malanotte-Rizzoli, P. (1991). *Data Assimilation in Meteorology and Oceanography* (Vol. 33; R. Dmowska & B.-A. G. Saltzman, Eds.). Elsevier. doi: 10.1016/S0065-2687(08)60442-2

Girotto, M., De Lannoy, G., Reichle, R., Rodell, M., Draper, C., Bhanja, S., & Mukherjee, A. (2017). Benefits and pitfalls of GRACE data assimilation: A case study of terrestrial water storage depletion in India. *Geophysical Research Letters*, *44*(9), 4107–4115. doi: 10.1002/2017GL072994

Girotto, M., Margulis, S., & Durand, M. (2014). Probabilistic SWE reanalysis as a generalization of deterministic SWE reconstruction techniques. *Hydrological Processes*, *28*(12), 3875–3895. doi: 10.1002/hyp.9887

Girotto, M., Reichle, R., Rodell, M., Liu, Q., Mahanama, S., & De Lannoy, G. (2019). Multi-sensor assimilation of SMOS brightness temperature and GRACE terrestrial water storage observations for soil moisture and shallow groundwater estimation. *Remote Sensing of Environment*, *227*, 12–27. doi: 10.1016/j.rse.2019.04.001

Girotto, M., Reichle, R., Rodell, M., & Maggioni, V. (2021). Data Assimilation of Terrestrial Water Storage Observations to Estimate Precipitation Fluxes: A Synthetic Experiment. *Remote Sensing*. doi: 10.3390/rs13061223

Gomez, B., Charlton-Perez, C., Lewis, H., & Candy, B. (2020). The Met Office Operational Soil Moisture Analysis System. *Remote Sensing*. doi: 10.3390/rs12223691

Goncalves, L., Shuttleworth, W., Vila, D., Larroza, E., Bottino, M., & Herdies, D.

(2009). The South American Land Data Assimilation System (SALDAS) 5-Yr Retrospective Atmospheric Forcing Datasets. *Journal of Hydrometeorology*, *10*(4), 999–1010. doi: 10.1175/2009JHM1049.1

Grant, J., Saleh-Contell, K., Wigneron, J., Guglielmetti, M., Kerr, Y., & Schwank, M. (2008). Calibration of the L-MEB Model Over a Coniferous and a Deciduous Forest. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(3), 808–818. doi: 10.1109/TGRS.2007.914801

Green, J., Konings, A., Alemohammad, S., Berry, J., Entekhabi, D., & Kolassa, J. (2017). Regionally strong feedbacks between the atmosphere and terrestrial biosphere. *Nature Geoscience*, *10*(6), 410–414.

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., & Calvet, J.-C. (2020). Validation practices for satellite soil moisture retrievals: What are (the) errors? *Remote Sensing of Environment*, *244*, 111806. doi: 10.1016/j.rse.2020.111806

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Florke, M., & Hanasaki, N. (2014). Global water resources affected by human interventions and climate change. *Proceedings of the National Academy of Sciences*, *111*(9), 3251 – 3256. doi: 10.1073/pnas.1222475110

Haiden, T., Janousek, M., Bidlot, J., Buizza, R., Ferranti, L., Prates, F., & Vitart, F. (2018). Evaluation of ECMWF forecasts, including the 2018 upgrade. In *European Centre for Medium Range Weather Forecasts.* Reading, UK.

Hain, C., Crow, W., Anderson, M., & Mecikalski, J. (2012). An ensemble Kalman filter dual assimilation of thermal infrared and microwave satellite observations of soil moisture into the Noah land surface model. *Water Resources Research*, *48*(11). doi: 10.1029/2011WR011268

Han, X., Hendricks Franssen, H.-J., Li, X., Zhang, Y., Montzka, C., & Vereecken, H. (2013). Joint Assimilation of Surface Temperature and L-Band Microwave Brightness Temperature in Land Data Assimilation. *Vadose Zone Journal*, *12*(3). doi: 10.2136/vzj2012.0072

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Munoz-Sabater, J., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Hess, R. (2001). Assimilation of screen-level observations by variational soil moisture analysis. *Meteorology and Atmospheric Physics*, *77*(1), 145–154. doi: 10.1007/s007030170023

Hollmann, R., Merchant, C., Saunders, R., Downy, C., Buchwitz, M., & Cazenave, A. (2013). The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables. *Bulletin of the American Meteorological Society*, *94*(10), 1541–1552. doi: 10.1175/BAMS-D-11-00254.1

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., … Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701–722.

Houborg, R., & McCabe, M. (2016). High-Resolution NDVI from planet's constellation of earth observing nano-satellites: A new data source for precision agriculture. *Remote Sensing*, *8*(9), 768.

Houborg, R., Rodell, M., Li, B., Reichle, R., & Zaitchik, B. (2012). Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, *48*(7). doi: 10.1029/2011WR011291

Houser, P., De Lannoy, G., & Walker, J. (2012). Hydrologic data assimilation. In T. by & J. (Eds.), *Approaches to Managing Disaster-Assessing Hazards, Emergencies and Disaster Impacts* (pp. 41–64).

Houser, P., Shuttleworth, W., Famiglietti, J., Gupta, H., Syed, K., & Goodrich, D. (1998). Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*, *34*(12), 3405–3420.

Houtekamer, P. L., & Zhang, F. (2016). Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, *144*(12), 4489 - 4532. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/144/12/mwr-d-15-0440.1.xml` doi: 10.1175/MWR-D-15-0440.1

Huang, C., Newman, A., Clark, M., Wood, A., & Zheng, X. (2017). Evaluation of snow data assimilation using the ensemble Kalman filter for seasonal streamflow prediction in the western United States. *Hydrology and Earth System Sciences*, *21*(1), 635–650.

Huang, J., Ji, M., Xie, Y., Wang, S., He, Y., & Ran, J. (2016). Global semi-arid climate change over last 60 years. *Climate Dynamics*, *46*(3), 1131–1150. doi: 10.1007/s00382-015-2636-8

Ingram, W. (2016). Increases all round. *Nature Climate Change*, *6*(5), 443–444. doi: 10.1038/nclimate2966

Jarlan, L., Balsamo, G., Lafont, S., Beljaars, A., Calvet, J.-C., & Mougin, E. (2008). Analysis of leaf area index in the ECMWF land surface model and impact on latent heat and carbon fluxes: Application to West Africa. *Journal of Geophysical Research: Atmospheres*, *113*(D24).

Jenkins, G. (1969). *The systems approach*. Citeseer.

Kain, J., Xue, M., Coniglio, M., Weiss, S., Kong, F., & Jensen, T. (2010). Assessing Advances in the Assimilation of Radar Data and Other Mesoscale Observations within a Collaborative Forecasting-Research Environment. *Weather and Forecasting*, *25*(5), 1510–1521. doi: 10.1175/2010WAF2222405.1

Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability*. Cambridge university press.

Kaminski, T., Knorr, W., Schurmann, G., Scholze, M., Rayner, P., & Zaehle, S. (2013). The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges. *Journal of Geophysical Research: Biogeosciences*, *118*(4), 1414–1426. doi: 10.1002/jgrg.20118

Kaminski, T., & Mathieu, P.-P. (2017). Reviews and syntheses: Flying the satellite into your model: on the role of observation operators in constraining models of the Earth system and the carbon cycle. *Biogeosciences*, *14*(9), 2343–2357.

Katzfuss, M., Stroud, J., & Wikle, C. (2016). Understanding the ensemble Kalman filter. *The American Statistician*, *70*(4), 350–357.

Kerr, Y., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., … Mecklenburg, S. (2010). The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle. *Proceedings of the IEEE*, *98*(5), 666–687. doi: 10.1109/JPROC.2010.2043032

Kimball, J. (2008). Earth Observation of Global Change: The Role of Satellite Remote Sensing in Monitoring the Global Environment. *Eos, Transactions American Geophysical Union*, *89*(32), 294. doi: 10.1029/2008EO320009

Kolassa, J., Reichle, R., & Draper, C. (2017a). Merging active and passive microwave observations in soil moisture data assimilation. *Remote Sensing of Environment*, *191*, 117–130. doi: 10.1016/j.rse.2017.01.015

Kolassa, J., Reichle, R., Liu, Q., Alemohammad, S., Gentine, P., & Aida, K. (2018).

Estimating surface soil moisture from SMAP observations using a Neural Network technique. *Remote Sensing of Environment*, *204*, 43–59. doi: 10.1016/j.rse.2017.10.045

Kolassa, J., Reichle, R., Liu, Q., Cosh, M., Bosch, D., & Caldwell, T. (2017b). Data assimilation to extract soil moisture information from SMAP observations. *Remote Sensing*, *9*(11), 1179.

Konapala, G., Mishra, A., Wada, Y., & Mann, M. (2020). Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. *Nature Communications*, *11*(1), 3044. Retrieved from `https://doi.org/10.1038/s41467-020-16757-w` doi: 10.1038/s41467-020-16757-w

Konings, A., Piles, M., Rotzer, K., McColl, K., Chan, S., & 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. doi: 10.1016/j.rse.2015.11.009

Kornfeld, R., Arnold, B., Gross, M., Dahya, N., Klipstein, W., Gath, P., & Bettadpur, S. (2019). GRACE-FO: The Gravity Recovery and Climate Experiment Follow-On Mission. *Journal of Spacecraft and Rockets*, *56*(3), 931–951. doi: 10.2514/1.A34326

Koster, R., DeAngelis, A., Schubert, S., & Molod, A. (2021). Asymmetry in Subseasonal Surface Air Temperature Forecast Error with Respect to Soil Moisture Initialization. *Journal of Hydrometeorology*, *22*(10), 2505–2519. doi: 10.1175/JHM-D-21-0022.1

Koster, R., Guo, Z., Yang, R., Dirmeyer, P., Mitchell, K., & Puma, M. (2009). On the Nature of Soil Moisture in Land Surface Models. *Journal of Climate*, *22*(16), 4322–4335. doi: 10.1175/2009JCLI2832.1

Koster, R., Mahanama, S., Yamada, T., Balsamo, G., Berg, A., & Boisserie, M. (2010). Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters*, *37*(2). doi: 10.1029/2009GL041677

Koster, R., Suarez, M., Liu, P., Jambor, U., Berg, A., & Kistler, M. (2004). Realistic Initialization of Land Surface States: Impacts on Subseasonal Forecast Skill. *Journal of Hydrometeorology*, *5*(6), 1049–1063. doi: 10.1175/JHM-387.1

Kotsuki, S., Miyoshi, T., Kondo, K., & Potthast, R. (2022). A Local Particle Filter and Its Gaussian Mixture Extension Implemented with Minor Modifications to the LETKF. *Geoscientific Model Development Discussions*, 1–38.

Kumar, S., Dong, J., Peters-Lidard, C., Mocko, D., & Gomez, B. (2017). Role of forcing uncertainty and background model error characterization in snow data assimilation. *Hydrology and Earth System Sciences*, *21*(6), 2637–2647.

Kumar, S., Getirana, A., Libonati, R., Hain, C., Mahanama, S., & Andela, N. (2022). Changes in land use enhance the sensitivity of tropical ecosystems to fire-climate extremes. *Scientific Reports*, *12*(1), 964. doi: 10.1038/s41598-022-05130-0

Kumar, S., Holmes, T., Andela, N., Dharssi, I., Vinodkumar, H., & C. (2021). The 2019-2020 Australian Drought and Bushfires Altered the Partitioning of Hydrological Fluxes. *Geophysical Research Letters*, *48*(1), 1–10. doi: 10.1029/2020GL091411

Kumar, S., Holmes, T., Bindlish, R., Jeu, R., & Peters-Lidard, C. (2020a). Assimilation of vegetation optical depth retrievals from passive microwave radiometry. *Hydrol. Earth Syst. Sci*, *24*(7), 3431–3450. doi: 10.5194/hess-24-3431-2020

Kumar, S., Jasinski, M., Mocko, D., Rodell, M., Borak, J., & Li, B. (2018). NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment. *Journal of Hydrometeorology*. doi: 10.1175/JHM-D-17-0125.1

Kumar, S., Mocko, D., Vuyovich, C., & Peters-Lidard, C. (2020b). Impact of Surface Albedo Assimilation on Snow Estimation. *Remote Sensing*. Retrieved from `https://doi.org/10.3390/rs12040645` doi: 10.3390/rs12040645

Kumar, S., Mocko, M., D., W., S., P.-L., D., C., & Borak, J. (2019). Assimilation of Remotely Sensed Leaf Area Index into the Noah-MP Land Surface Model: Impacts on Water and Carbon Fluxes and States over the Continental United States. *Journal of Hydrometeorology*, *20*(7), 1359–1377. doi: 10.1175/JHM-D-18-0237.1

Kumar, S., Peters-Lidard, C., Arsenault, K., Getirana, A., Mocko, D., & Liu, Y. (2015a). Quantifying the Added Value of Snow Cover Area Observations in Passive Microwave Snow Depth Data Assimilation. *Journal of Hydrometeorology*, *16*(4), 1736–1741. doi: 10.1175/JHM-D-15-0021.1

Kumar, S., Peters-Lidard, C., Mocko, D., Reichle, R., Liu, Y., & Arsenault, K. (2014). Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation. *Journal of Hydrometeorology*, *15*(6), 2446–2469. doi: 10.1175/JHM-D-13-0132.1

Kumar, S., Peters-Lidard, C., Santanello, J., Reichle, R., Draper, C., & Koster, R. (2015b). Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. *Hydrology and Earth System Sciences*, *19*(11), 4463–4478. doi: 10.5194/hess-19-4463-2015

Kumar, S., Peters-Lidard, C., Tian, Y., Houser, P., Geiger, J., & Olden, S. (2006). Land information system: An interoperable framework for high resolution land surface modeling. *Environmental Modelling and Software*, *21*(10), 1402–1415. doi: 10.1016/j.envsoft.2005.07.004

Kumar, S., Reichle, R., Harrison, K., Peters-Lidard, C., Yatheendradas, S., & Santanello, J. (2012). A comparison of methods for a priori bias correction in soil moisture data assimilation. *Water Resources Research*, *48*(3), W03515.

Kumar, S., Reichle, R., Koster, R., Crow, W., & Peters-Lidard, C. (2009). Role of Subsurface Physics in the Assimilation of Surface Soil Moisture Observations. *Journal of Hydrometeorology*, *10*(6), 1534–1547. doi: 10.1175/2009JHM1134.1

Kumar, S., Reichle, R., Peters-Lidard, C., Koster, R., Zhan, X., & Crow, W. (2008). A land surface data assimilation framework using the land information system: Description and applications. *Advances in Water Resources*, *31*(11). doi: 10.1016/j.advwatres.2008.01.013

Kumar, S., Zaitchik, B., Peters-Lidard, C., Rodell, M., Reichle, R., & Li, B. (2016). Assimilation of gridded GRACE terrestrial water storage estimates in the North American Land Data Assimilation System. *Journal of Hydrometeorology*, *17*(7), 1951–1972. doi: 10.1175/JHM-D-15-0157.1

Kwon, Y., Forman, B., Ahmad, J., Kumar, S., & Yoon, Y. (2019). Exploring the utility of machine learning-based passive microwave brightness temperature data assimilation over terrestrial snow in high mountain Asia. *Remote Sensing*, *11*(19), 2265.

Kwon, Y., Yoon, Y., Forman, B., Kumar, S., & Wang, L. (2021). Quantifying the observational requirements of a space-borne LiDAR snow mission. *Journal of Hydrol-*

ogy, *601*, 126709. doi: 10.1016/j.jhydrol.2021.126709

Lahmers, T. M., Kumar, S., Rosen, D., Dugger, A., Gochis, D. J., Santanello, J. A., ... Dunlap, R. (2022). Assimilation of NASA's Airborne Snow Observatory Snow Measurements for Improved Hydrological Modeling: A Case Study Enabled by the Coupled LIS/WRF-Hydro System. *Water Resources Research*, *58*(3), e2021WR029867. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021WR029867` (e2021WR029867 2021WR029867) doi: https://doi.org/10.1029/2021WR029867

Lahoz, W., & De Lannoy, G. (2014). Closing the Gaps in Our Knowledge of the Hydrological Cycle over Land: Conceptual Problems. *Surveys in Geophysics*, *35*(3), 623–660. doi: 10.1007/s10712-013-9221-7

Lahoz, W., & Schneider, P. (2014). Data assimilation: making sense of Earth Observation. *Frontiers in Environmental Science*. Retrieved from `https://www.frontiersin.org/article/10.3389/fenvs.2014.00016`

Lakshmi, V. (2000). A simple surface temperature assimilation scheme for use in land surface models. *Water Resources Research*, *36*(12), 3687–3700. doi: 10.1029/2000WR900204

Lakshmi, V. (2014). *Remote sensing of the terrestrial water cycle* (Vol. 206). John Wiley & Sons.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

L'Ecuyer, T., Beaudoing, H. K., Rodell, M., Olson, W., Lin, B., Kato, S., ... Hilburn, K. (2015). The observed state of the energy budget in the early twenty-first century. *Journal of Climate*, *28*(21), 8319–8346. doi: 10.1175/JCLI-D-14-00556.1

Lettenmaier, D., Alsdorf, D., Dozier, J., Huffman, G., Pan, M., & Wood, E. (2015). Inroads of remote sensing into hydrologic science during the WRR era. *Water Resources Research*, *51*(9), 7309–7342. doi: 10.1002/2015WR017616

Lewis, P., Gomez-Dans, J., Kaminski, T., Settle, J., Quaife, T., & Gobron, N. (2012). An Earth Observation Land Data Assimilation System (EO-LDAS). *Remote Sensing of Environment*, *120*, 219–235. doi: 10.1016/j.rse.2011.12.027

Li, B., Rodell, M., Kumar, S., Beaudoing, H., Getirana, A., & Zaitchik, B. (2019). Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges. *Water Resources Research*, *55*(9), 7564–7586. doi:

10.1029/2018WR024618

Li, B., Rodell, M., Zaitchik, B., Reichle, R., Koster, R., & Dam, T. (2012). Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe. *Journal of Hydrology*, *446–447*, 103–115. doi: 10.1016/j.jhydrol.2012.04.035

Li, D., Lettenmaier, D., Margulis, S., & Andreadis, K. (2019). The Value of Accurate High-Resolution and Spatially Continuous Snow Information to Streamflow Forecasts. *Journal of Hydrometeorology*, *20*(4), 731–749. doi: 10.1175/JHM-D-18-0210.1

Li, F., Crow, W., & Kustas, W. (2010). Towards the estimation root-zone soil moisture via the simultaneous assimilation of thermal and microwave soil moisture retrievals. *Advances in Water Resources*, *33*(2), 201–214. doi: 10.1016/j.advwatres.2009.11.007

Lievens, H., De Lannoy, G., Al Bitar, A., Drusch, M., Dumedah, G., & Franssen, H.-J. (2016). Assimilation of SMOS soil moisture and brightness temperature products into a land surface model. *Remote Sensing of Environment*, *180*, 292–304.

Lievens, H., Martens, B., Verhoest, N., Hahn, S., Reichle, R., & Miralles, D. (2017). Assimilation of global radar backscatter and radiometer brightness temperature observations to improve soil moisture and land evaporation estimates. *Remote Sensing of Environment*, *189*, 194–210.

Lievens, H., Tomer, S., Al Bitar, A., De Lannoy, G., Drusch, M., & Dumedah, G. (2015). SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. *Remote Sensing of Environment*, *168*, 146–162. doi: 10.1016/j.rse.2015.06.025

Lin, L.-F., & Pu, Z. (2020). Improving near-surface short-range weather forecasts using strongly coupled land-atmosphere data assimilation with GSI-EnKF. *Monthly Weather Review*, *148*(7), 2863–2888.

Liu, Y., Peters-Lidard, C., Kumar, S., Foster, J., Shaw, M., Tian, Y., & Fall, G. (2013). Assimilating satellite-based snow depth and snow cover products for improving snow predictions in Alaska. *Advances in Water Resources*, *54*, 208–227.

Liu, Y., Weerts, A., Clark, M., Hendricks Franssen, H.-J., Kumar, S., & Moradkhani, H. (2012). Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrology and Earth*

System Sciences, *16*(10), 3863–3887.

Llop, J., Roberts, P., Hao, Z., Tomas, L., & Beauplet, V. (2014). Very low earth orbit mission concepts for earth observation: Benefits and challenges. In *Reinventing Space Conference* (pp. 18–21).

Lo, M.-H., Famiglietti, J. S., Reager, J. T., Rodell, M., Swenson, S., & Wu, W.-Y. (2016). Grace-based estimates of global groundwater depletion. In *Terrestrial water cycle and climate change* (p. 135-146). American Geophysical Union (AGU). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/9781118971772.ch7` doi: https://doi.org/10.1002/9781118971772.ch7

Loizu, J., Massari, C., Alvarez-Mozos, J., Tarpanelli, A., Brocca, L., & Casali, J. (2018). On the assimilation set-up of ASCAT soil moisture data for improving streamflow catchment simulation. *Advances in Water Resources*, *111*, 86–104. doi: 10.1016/j.advwatres.2017.10.034

Lu, Y., Steele-Dunne, S., Farhadi, L., & Giesen, N. (2017). Mapping Surface Heat Fluxes by Assimilating SMAP Soil Moisture and GOES Land Surface Temperature Data. *Water Resources Research*, *53*(12), 10858–10877. doi: 10.1002/2017WR021415

MacBean, N., Bacour, C., Raoult, N., Bastrikov, V., Koffi, E. N., Kuppel, S., … Peylin, P. (2022). Quantifying and Reducing Uncertainty in Global Carbon Cycle Predictions. In *Lessons and Perspectives From 15 Years of Data Assimilation Studies with the ORCHIDEE Terrestrial Biosphere Model. Global Biogeochemical Cycles* (p. 2021 007177). doi: 10.1029/2021GB007177

Magnusson, J., Gustafsson, D., Husler, F., & Jonas, T. (2014). Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods. *Water Resources Research*, *50*(10), 7816–7835. doi: 10.1002/2014WR015302

Mahfouf, J.-F. (1991). Analysis of soil moisture from near-surface parameters: A feasibility study. *Journal of Applied Meteorology and Climatology*, *30*(11), 1534–1547.

Malanotte-Rizzoli, P. (1997). Modern approaches to data assimilation in ocean modeling. *Oceanographic Literature Review*, *5*(44), 530.

Malik, M., Velde, R., Vekerdy, Z., & Su, Z. (2012). Assimilation of Satellite-Observed Snow Albedo in a Land Surface Model. *Journal of Hydrometeorology*, *13*(3), 1119–1130. doi: 10.1175/JHM-D-11-0125.1

Margulis, S., Fang, Y., Li, D., Lettenmaier, D., & Andreadis, K. (2019). The Utility of Infrequent Snow Depth Images for Deriving Continuous Space-Time Estimates of Seasonal Snow Water Equivalent. *Geophysical Research Letters*, *46*(10), 5331–5340. doi: 10.1029/2019GL082507

Margulis, S., Girotto, M., Cortes, G., & Durand, M. (2015). A Particle Batch Smoother Approach to Snow Water Equivalent Estimation. *Journal of Hydrometeorology*, *16*(4), 1752–1772. doi: 10.1175/JHM-D-14-0177.1

Margulis, S., McLaughlin, D., Entekhabi, D., & Dunne, S. (2002). Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. *Water Resources Research*, *38*(12), 18–35. doi: 10.1029/2001WR001114

Mass, C. (2012). Nowcasting: The Promise of New Technologies of Communication, Modeling, and Observation. *Bulletin of the American Meteorological Society*, *93*(6), 797–809. doi: 10.1175/BAMS-D-11-00153.1

Massari, C., Modanesi, S., Dari, J., Gruber, A., De Lannoy, G., & Girotto, M. (2021). A Review of Irrigation Information Retrievals from Space and Their Utility for Users. *Remote Sensing*. doi: 10.3390/rs13204112

McCabe, M., Rodell, M., Alsdorf, D., Miralles, D., Uijlenhoet, R., & Wagner, W. (2017). The future of Earth observation in hydrology. *Hydrology and Earth System Sciences*, *21*(7), 3879–3914.

McDonald, R., Green, P., Balk, D., Fekete, B., Revenga, C., Todd, M., & Montgomery, M. (2011). Urban growth, climate change, and freshwater availability. *Proceedings of the National Academy of Sciences*, *108*(15), 6312 – 6317. doi: 10.1073/pnas.1011615108

McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., & Wang, S. (2017). A land data assimilation system for sub-Saharan Africa food and water security applications. *Scientific Data*, *4*(1), 170012. doi: 10.1038/sdata.2017.12

Mehran, A., AghaKouchak, A., Nakhjiri, N., Stewardson, M., Peel, M., & Phillips, T. (2017). Compounding Impacts of Human-Induced Water Stress and Climate Change on Water Availability. *Scientific Reports*, *7*(1), 6282. doi: 10.1038/s41598-017-06765-0

Milly, P., Dunne, K., & Vecchia, A. (2005). Global pattern of trends in streamflow and water availability in a changing climate. *Nature*, *438*(7066), 347–350. doi: 10

.1038/nature04312

Mitchell, K., Lohmann, D., Houser, P., Wood, E., Schaake, J., & Robock, A. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research: Atmospheres*, *109*(D7).

Miyoshi, T., & Yamane, S. (2007). Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Monthly Weather Review*, *135*(11), 3841–3861.

Mladenova, I., Bolten, J., Crow, W., Sazib, N., Cosh, M., Tucker, C., & Reynolds, C. (2019). Evaluating the Operational Application of SMAP for Global Agricultural Drought Monitoring. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(9), 3387–3397. doi: 10.1109/JSTARS.2019.2923555

Mocko, D., Kumar, S., Peters-Lidard, C., & Wang, S. (2021). Assimilation of Vegetation Conditions Improves the Representation of Drought over Agricultural Areas. *Journal of Hydrometeorology*, *22*(5), 1085–1098. doi: 10.1175/JHM-D-20-0065.1

Montavon, G., Samek, W., & Muller, K.-R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, *73*, 1–15.

Montzka, C., Pauwels, V., Franssen, H.-J., Han, X., & Vereecken, H. (2012). Multivariate and Multiscale Data Assimilation in Terrestrial Systems: A Review. *Sensors*. doi: 10.3390/s121216291

Moore, A., Martin, M., Akella, S., Arango, H., Balmaseda, M., & Bertino, L. (2019). Synthesis of Ocean Observations Using Data Assimilation for Operational, Real-Time and Reanalysis Systems: A More Complete Picture of the State of the Ocean. *Frontiers in Marine Science*. Retrieved from `https://www.frontiersin.org/article/10.3389/fmars.2019.00090`

Moradkhani, H., Hsu, K., Gupta, H., & Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research*, *41*(5).

Mucia, A., Bonan, B., Albergel, C., Zheng, Y., & Calvet, J.-C. (2021). Assimilation of passive microwave vegetation optical depth in LDAS-Monde: a case study over the continental US. In *Biogeosciences Discussions* (pp. 1–45).

Munoz-Sabater, J., Dutra, E., Agusti-Panareda, A., Albergel, C., Arduini, G., & Balsamo, G. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383.

Munoz-Sabater, J., Lawrence, H., Albergel, C., de Rosnay, P., Isaksen, L., & Mecklenburg, S. (2019). Assimilation of SMOS brightness temperatures in the ECMWF Integrated Forecasting System. *Quarterly Journal of the Royal Meteorological Society*, *145*(723), 2524–2548. doi: 10.1002/qj.3577

Navon, I. (2009). *Data Assimilation for Numerical Weather Prediction: A Review BT - Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications* (S. Park & L. Xu, Eds.). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-540-71056-1_2

Neal, J., Atkinson, P., & Hutton, C. (2007). Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. *Journal of Hydrology*, *336*(3–4), 401–415.

Nearing, G., Crow, W., Thorp, K., Moran, M., Reichle, R., & Gupta, H. (2012). Assimilating remote sensing observations of leaf area index and soil moisture for wheat yield estimates: An observing system simulation experiment. *Water Resources Research*, *48*(5). doi: 10.1029/2011WR011420

Nearing, G., Yatheendradas, S., Crow, W., Zhan, X., Liu, J., & Chen, F. (2018). The efficiency of data assimilation. *Water Resources Research*, *54*(9), 6374–6392.

Nichols, N. (2010). Mathematical concepts of data assimilation. In *Data assimilation* (pp. 13–39). Springer.

Nie, W., Zaitchik, B., Rodell, M., Kumar, S., Arsenault, K., Li, B., & Getirana, A. (2019). Assimilating GRACE Into a Land Surface Model in the Presence of an Irrigation-Induced Groundwater Trend. *Water Resources Research*, *55*(12), 11274–11294. doi: 10.1029/2019WR025363

Oaida, C., Reager, J., Andreadis, K., David, C., Levoe, S., & Painter, T. (2019). A High-Resolution Data Assimilation Framework for Snow Water Equivalent Estimation across the Western United States and Validation with the Airborne Snow Observatory. *Journal of Hydrometeorology*, *20*(3), 357–378. doi: 10.1175/JHM-D-18-0009.1

Oki, T., & Kanae, S. (2006). Global Hydrological Cycles and World Water Resources. *Science*, *313*(5790), 1068–1072. doi: 10.1126/science.1128845

Paiva, R., Collischonn, W., Bonnet, M.-P., Goncalves, L., Calmant, S., Getirana, A.,

& Silva, J. (2013). Assimilating in situ and radar altimetry data into a large-scale hydrologic-hydrodynamic model for streamflow forecast in the Amazon. *Hydrology and Earth System Sciences*, *17*(7), 2929–2946.

Pan, M., & Wood, E. (2006). Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter. *J. Hydrometeor*, *7*, 534–547.

Papagiannopoulou, C., Miralles, D., Decubber, S., Demuzere, M., Verhoest, N., Dorigo, W., & Waegeman, W. (2017). A non-linear Granger-causality framework to investigate climate-vegetation dynamics. *Geoscientific Model Development*, *10*(5), 1945–1960.

Parajka, J., Naeimi, V., Bloschl, G., Wagner, W., Merz, R., & Scipal, K. (2006). Assimilating scatterometer soil moisture data into conceptual hydrologic models at the regional scale. *Hydrology and Earth System Sciences*, *10*(3), 353–368.

Park, J., Forman, B., & Kumar, S. (2021). Estimation of snow mass information via assimilation of C-band synthetic aperture radar backscatter observations into an advanced land surface model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Parrens, M., Calvet, J.-C., de Rosnay, P., & Decharme, B. (2014). Benchmarking of L-band soil microwave emission models. *Remote Sensing of Environment*, *140*, 407–419. doi: 10.1016/j.rse.2013.09.017

Pauwels, V., & De Lannoy, G. (2009). Ensemble-based assimilation of discharge into rainfall-runoff models: A comparison of approaches to mapping observational information to state space. *Water Resources Research*, *45*(8). doi: 10.1029/2008WR007590

Pauwels, V., De Lannoy, G., Hendricks Franssen, H.-J., & Vereecken, H. (2013). Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid Kalman filter. *Hydrology and Earth System Sciences*, *17*(9), 3499–3521.

Penny, S., & Hamill, T. (2017). Coupled data assimilation for integrated Earth system analysis and prediction. *Bull. Amer. Meteor. Soc*, *98*, 169– 172. doi: 10.1175/BAMS-D-17-0036.1

Peters-Lidard, C., Kumar, S., Mocko, D., & Tian, Y. (2011). Estimating evapotranspiration with land data assimilation systems. *Hydrological Processes*, *25*(26), 3979–3992.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., & Kuppel, S. (2016). A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle. *Geoscientific Model Development*, *9*(9), 3321–3346.

Phan, X., Ferro-Famil, L., Gay, M., Durand, Y., Dumont, M., & Morin, S. (2014). 1D-Var multilayer assimilation of X-band SAR data into a detailed snowpack model. *The Cryosphere*, *8*(5), 1975–1987.

Pinnington, E., Amezcua, J., Cooper, E., Dadson, S., Ellis, R., & Peng, J. (2021). Improving soil moisture prediction of a high-resolution land surface model by parameterising pedotransfer functions through assimilation of SMAP satellite data. *Hydrology and Earth System Sciences*, *25*(3), 1617–1641.

Pinnington, E., Quaife, T., & Black, E. (2018). Impact of remotely sensed soil moisture and precipitation on soil moisture prediction in a data assimilation system with the JULES land surface model. *Hydrology and Earth System Sciences*, *22*(4), 2575–2588.

Pinnington, E., Quaife, T., Lawless, A., Williams, K., Arkebauer, T., & Scoby, D. (2020). The land variational ensemble data assimilation framework: LAVENDAR v1.0.0. *Geoscientific Model Development*, *v1*(1), 0 0.

Pitman, A. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, *23*(5), 479–510. doi: 10.1002/joc.893

Pongratz, J., Dolman, H., Don, A., Erb, K.-H., Fuchs, R., & Herold, M. (2018). Models meet data: Challenges and opportunities in implementing land management in Earth system models. *Global Change Biology*, *24*(4), 1470–1487. doi: 10.1111/gcb.13988

Powell, T., Galbraith, D., Christoffersen, B., Harper, A., Imbuzeiro, H., & Rowland, L. (2013). Confronting model predictions of carbon fluxes with measurements of Amazon forests subjected to experimental drought. *New Phytologist*, *200*(2), 350–365.

Pullen, S., Jones, C., & Rooney, G. (2011). Using Satellite-Derived Snow Cover Data to Implement a Snow Analysis in the Met Office Global NWP Model. *Journal of Applied Meteorology and Climatology*, *50*(5), 958–973. doi: 10.1175/2010JAMC2527.1

Quaife, T., Lewis, P., Kauwe, M., Williams, M., Law, B., Disney, M., & Bowyer, P. (2008). Assimilating canopy reflectance data into an ecosystem model with an Ensemble Kalman Filter. *Remote Sensing of Environment*, *112*(4), 1347–1364. doi: 10.1016/j.rse.2007.05.020

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.-F., Minh, D. H. T., … Williams, M. (2019). The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space. *Remote Sensing of Environment*, *227*, 44-60. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425719301233 doi: https://doi.org/10.1016/j.rse.2019.03.032

Rabier, F., & Liu, Z. (2003). Variational data assimilation: theory and overview. In *Proc. ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean* (pp. 29–43). Reading, UK.

Raczka, B., Hoar, T. J., Duarte, H. F., Fox, A. M., Anderson, J. L., Bowling, D. R., & Lin, J. C. (2021). Improving CLM5.0 Biomass and Carbon Exchange Across the Western United States Using a Data Assimilation System. *Journal of Advances in Modeling Earth Systems*, *13*(7), e2020MS002421. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002421 (e2020MS002421 2020MS002421) doi: https://doi.org/10.1029/2020MS002421

Raeder, K., Hoar, T. J., El Gharamti, M., Johnson, B. K., Collins, N., Anderson, J. L., … Coady, M. (2021). A new CAM6+ DART reanalysis with surface forcing from CAM6 to other CESM models. *Scientific Reports*, *11*(1), 1–24.

Raoult, N., Jupp, T., Cox, P., & Luke, C. (2016). Land-surface parameter optimisation using data assimilation techniques: the adJULES system V1.0. *Geoscientific Model Development*, *9*(8), 2833–2852.

Raoult, N., Ottle, C., Peylin, P., Bastrikov, V., & Maugis, P. (2021). Evaluating and Optimizing Surface Soil Moisture Drydowns in the ORCHIDEE Land Surface Model at In Situ Locations. *Journal of Hydrometeorology*, *22*(4), 1025–1043. doi: 10.1175/JHM-D-20-0115.1

Rayner, P. (2010). The current state of carbon-cycle data assimilation. *Current Opinion in Environmental Sustainability*, *2*(4), 289–296. doi: 10.1016/j.cosust.2010.05.005

Rayner, P., Michalak, A., & Chevallier, F. (2019). Fundamentals of data assimilation applied to biogeochemistry. *Atmospheric Chemistry and Physics*, *19*(22), 13911–13932.

Reichle, R. (2008a). Data assimilation methods in the Earth sciences. *Advances in Water Resources*, *31*(11), 1411–1418. doi: 10.1016/j.advwatres.2008.01.001

Reichle, R., Bosilovich, M., Crow, W., Koster, R., Kumar, S., Mahanama, S., & Zaitchik, B. (2009). Recent advances in land data assimilation at the NASA Global Modeling and Assimilation Office. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, 407–428.

Reichle, R., Crow, W., & Keppenne, C. (2008). An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resources Research*, *44*(3).

Reichle, R., Crow, W., Koster, R., Sharif, H., & Mahanama, S. (2008a). Contribution of soil moisture retrievals to land data assimilation products. *Geophysical Research Letters*, *35*(1). doi: 10.1029/2007GL031986

Reichle, R., De Lannoy, G., Forman, B., Draper, C., & Liu, Q. (2014). Connecting satellite observations with water cycle variables through land data assimilation: Examples using the NASA GEOS-5 LDAS. *Surveys in Geophysics*, *35*(3), 577–606.

Reichle, R., De Lannoy, G., Liu, Q., Koster, R., Kimball, J., & Crow, W. (2017). Global Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using Assimilation Diagnostics. *Journal of Hydrometeorology*, *18*(12), 3217–3237. doi: 10.1175/JHM-D-17-0130.1

Reichle, R., & Koster, R. (2004). Bias reduction in short records of satellite soil moisture. *Geophysical Research Letters*, *31*(19).

Reichle, R., Koster, R., De Lannoy, G., Forman, B., Liu, Q., Mahanama, S., & Toure, A. (2011). Assessment and Enhancement of MERRA Land Surface Hydrology Estimates. *Journal of Climate*, *24*(24), 6322–6338. doi: 10.1175/JCLI-D-10-05033.1

Reichle, R., Koster, R., Dong, J., & Berg, A. (2004). Global Soil Moisture from Satellite Observations, Land Surface Models, and Ground Data: Implications for Data Assimilation. *Journal of Hydrometeorology*, *5*(3), 430–442. doi: 10.1175/1525-7541(2004)005

Reichle, R., Kumar, S., Mahanama, S., Koster, R., & Liu, Q. (2010). Assimilation of

Satellite-Derived Skin Temperature Observations into Land Surface Models. *Journal of Hydrometeorology*, *11*(5), 1103–1122. doi: 10.1175/2010JHM1262.1

Reichle, R., Liu, Q., Ardizzone, J., Crow, W., De Lannoy, G., & Dong, J. (2021b). The Contributions of Gauge-Based Precipitation and SMAP Brightness Temperature Observations to the Skill of the SMAP Level-4 Soil Moisture Product. *Journal of Hydrometeorology*, *22*(2), 405–424. doi: 10.1175/JHM-D-20-0217.1

Reichle, R., Liu, Q., Koster, R., Crow, W., De Lannoy, G., & Kimball, J. (2019). Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3106–3130. doi: 10.1029/2019MS001729

Reichle, R., McLaughlin, D., & Entekhabi, D. (2002a). Hydrologic Data Assimilation with the Ensemble Kalman Filter. *Monthly Weather Review*, *130*(1), 103–114.

Reichle, R., Walker, J., Koster, R., & Houser, P. (2002b). Extended versus Ensemble Kalman Filtering for Land Data Assimilation. *Journal of Hydrometeorology*, *3*(6), 728–740. doi: 10.1175/1525-7541(2002)003

Reichle, R., Zhang, S., Liu, Q., Draper, C., Kolassa, J., & Todling, R. (2021a). Assimilation of SMAP Brightness Temperature Observations in the GEOS Land-Atmosphere Data Assimilation System. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 10628–10643.

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

Renzullo, L., Dijk, A., Perraud, J.-M., Collins, D., Henderson, B., & Jin, H. (2014). Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. *Journal of Hydrology*, *519*, 2747–2762. doi: 10.1016/j.jhydrol.2014.08.008

Ricci, S., Piacentini, A., Thual, O., Le Pape, E., & Jonville, G. (2011). Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting. *Hydrology and Earth System Sciences*, *15*(11), 3555–3575.

Ridler, M.-E., Zhang, D., Madsen, H., Kidmose, J., Refsgaard, J., & Jensen, K. (2017). Bias-aware data assimilation in integrated hydrological modelling. *Hydrology Research*, *49*(4), 989–1004. doi: 10.2166/nh.2017.117

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser,

P. R., ... Wood, E. F. (2015). The Observed State of the Water Cycle in the Early Twenty-First Century. *Journal of Climate*, *28*(21), 8289–8318. doi: 10.1175/JCLI-D-14-00555.1

Rodell, M., Famiglietti, J., Wiese, D., Reager, J., Beaudoing, H., Landerer, F., & Lo, M. (2018). Emerging trends in global freshwater availability. *Nature*, *557*(7707), 651–659. doi: 10.1038/s41586-018-0123-1

Rodell, M., & Houser, P. (2004). Updating a Land Surface Model with MODIS-Derived Snow Cover. *Journal of Hydrometeorology*, *5*(6), 1064–1075. doi: 10.1175/JHM-395.1

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., & Meng, C.-J. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*(3), 381–394.

Rodell, M., Velicogna, I., & Famiglietti, J. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, *460*(7258), 999–1002. doi: 10.1038/nature08238

Rodriguez-Fernandez, N., Aires, F., Richaume, P., Kerr, Y., Prigent, C., & Kolassa, J. (2015). Soil Moisture Retrieval Using Neural Networks: Application to SMOS. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(11), 5991–6007. doi: 10.1109/TGRS.2015.2430845

Rodriguez-Fernandez, N., de Rosnay, P., Albergel, C., Richaume, P., Aires, F., Prigent, C., & Kerr, Y. (2019). SMOS Neural Network Soil Moisture Data Assimilation in a Land Surface Model and Atmospheric Impact. *Remote Sensing*. Retrieved from `https://doi.org/10.3390/rs11111334` doi: 10.3390/rs11111334

Royer, A., Roy, A., Montpetit, B., Saint-Jean-Rondeau, O., Picard, G., Brucker, L., & Langlois, A. (2017). Comparison of commonly-used microwave radiative transfer models for snow remote sensing. *Remote Sensing of Environment*, *190*, 247–259. doi: 10.1016/j.rse.2016.12.020

Rudiger, C., Albergel, C., Mahfouf, J., Calvet, J., & Walker, J. (2010). Evaluation of the observation operator Jacobian for leaf area index data assimilation with an extended Kalman filter. *Journal of Geophysical Research: Atmospheres*, *115*(D9).

Sabater, J., Jarlan, L., Calvet, J.-C., Bouyssel, F., & Rosnay, P. (2007). From Near-Surface to Root-Zone Soil Moisture Using Different Assimilation Techniques.

Journal of Hydrometeorology, *8*(2), 194–206. doi: 10.1175/JHM571.1

Sabater, J., Rudiger, C., Calvet, J.-C., Fritz, N., Jarlan, L., & Kerr, Y. (2008). Joint assimilation of surface soil moisture and LAI observations into a land surface model. *Agricultural and Forest Meteorology*, *148*(8), 1362–1373. doi: 10.1016/j.agrformet.2008.04.003

Sandau, R. (2010). Status and trends of small satellite missions for Earth observation. *Acta Astronautica*, *66*(1), 1–12. doi: 10.1016/j.actaastro.2009.06.008

Santanello, J., Lawston, P., Kumar, S., & Dennis, E. (2019). Understanding the Impacts of Soil Moisture Initial Conditions on NWP in the Context of Land-Atmosphere Coupling. *Journal of Hydrometeorology*, *20*(5), 793–819. doi: 10.1175/JHM-D-18-0186.1

Sawada, Y., Koike, T., & Walker, J. (2015). A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics. *Journal of Geophysical Research: Atmospheres*, *120*(12), 5910–5930.

Schimel, D., Schneider, F., Participants, J., & E. (2019). Flux towers in the sky: global ecology from space. *New Phytologist*, *224*(2), 570–584.

Scholze, M., Kaminski, T., Knorr, W., Blessing, S., Vossbeck, M., Grant, J., & Scipal, K. (2016). Simultaneous assimilation of SMOS soil moisture and atmospheric CO2 in-situ observations to constrain the global terrestrial carbon cycle. *Remote Sensing of Environment*, *180*, 334–345. doi: 10.1016/j.rse.2016.02.058

Schurmann, G., Kaminski, T., Kostler, C., Carvalhais, N., Vobeck, M., & Kattge, J. (2016). Constraining a land-surface model with multiple observations by application of the MPI-Carbon Cycle Data Assimilation System V1.0. *Geoscientific Model Development*, *9*(9), 2999–3026.

Sellars, S., Nguyen, P., Chu, W., Gao, X., Hsu, K., & Sorooshian, S. (2013). Computational Earth Science: Big Data Transformed Into Insight. *Eos, Transactions American Geophysical Union*, *94*(32), 277–278. doi: 10.1002/2013EO320001

Sellers, P., Schimel, D., Berrien, M., Junjie, L., & Annmarie, E. (2018). Observing carbon cycle-climate feedbacks from space. *Proceedings of the National Academy of Sciences*, *115*(31), 7860–7868. doi: 10.1073/pnas.1716613115

Seo, D.-J., Cajina, L., Corby, R., & Howieson, T. (2009). Automatic state updating for operational streamflow forecasting via variational data assimilation. *Journal of Hydrology*, *367*(3–4), 255–275.

Seo, E., Lee, M.-I., & Reichle, R. (2021). Assimilation of SMAP and ASCAT soil moisture retrievals into the JULES land surface model using the Local Ensemble Transform Kalman Filter. *Remote Sensing of Environment*, *253*, 112222. doi: 10.1016/j.rse.2020.112222

Shahabadi, M., Belair, S., Bilodeau, B., Carrera, M., & Garand, L. (2019). Impact of Weak Coupling between Land and Atmosphere Data Assimilation Systems on Environment and Climate Change Canada's Global Deterministic Prediction System. *Weather and Forecasting*, *34*(6), 1741–1758. doi: 10.1175/WAF-D-19-0073.1

Shan, X., Steele-Dunne, S., Huber, M., Hahn, S., Wagner, W., Bonan, B., ... Georgievska, S. (2022). Towards constraining soil and vegetation dynamics in land surface models: Modeling ASCAT backscatter incidence-angle dependence with a Deep Neural Network. *Remote Sensing of Environment*, *279*, 113116. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0034425722002309` doi: https://doi.org/10.1016/j.rse.2022.113116

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Silva, J., Calmant, S., Seyler, F., Rotunno Filho, O., Cochonneau, G., & Mansur, W. (2010). Water levels in the Amazon basin derived from the ERS 2 and ENVISAT radar altimetry missions. *Remote Sensing of Environment*, *114*(10), 2160–2181.

Slater, A., & Clark, M. (2006). Snow Data Assimilation via an Ensemble Kalman Filter. *Journal of Hydrometeorology*, *7*(3), 478–493. doi: 10.1175/JHM505.1

Sluka, T., Penny, S., Kalnay, E., & Miyoshi, T. (2016). Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophysical Research Letters*, *43*(2), 752–759. doi: 10.1002/2015GL067238

Smith, W., Fox, A., MacBean, N., Moore, D., & Parazoo, N. (2020). Constraining estimates of terrestrial carbon uptake: new opportunities using long-term satellite observations and data assimilation. *New Phytologist*, *225*(1), 105–112. doi: 10.1111/nph.16055

Smyth, E., Raleigh, M., & Small, E. (2019). Particle Filter Data Assimilation of Monthly Snow Depth Observations Improves Estimation of Snow Density and SWE. *Water Resources Research*, *55*(2), 1296–1311. doi: 10.1029/2018WR023400

Stavros, E., Schimel, D., Pavlick, R., Serbin, S., Swann, A., & Duncanson, L. (2017). ISS observations offer insights into plant function. *Nature Ecology & Evolution*, *1*(7), 1–5.

Su, C.-H., Ryu, D., Dorigo, W., Zwieback, S., Gruber, A., & Albergel, C. (2016). Homogeneity of a global multisatellite soil moisture climate data record. *Geophysical Research Letters*, *43*(21), 11,211–11,252. doi: 10.1002/2016GL070458

Su, H., Yang, Z.-L., Dickinson, R., Wilson, C., & Niu, G.-Y. (2010). Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information. *Journal of Geophysical Research: Atmospheres*, *115*(D10). doi: 10.1029/2009JD013035

Su, H., Yang, Z.-L., Niu, G.-Y., & Dickinson, R. (2008). Enhancing the estimation of continental-scale snow water equivalent by assimilating MODIS snow cover with the ensemble Kalman filter. *Journal of Geophysical Research: Atmospheres*, *113*(D8). doi: 10.1029/2007JD009232

Sun, Y., Frankenberg, C., Wood, J. D., Schimel, D. S., Jung, M., Guanter, L., . . . Yuen, K. (2017). OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence. *Science*, *358*(6360), 5747. doi: 10.1126/science.aam5747

Syed, T., Famiglietti, J., Rodell, M., Chen, J., & Wilson, C. (2008). Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resources Research*, *44*(2). doi: 10.1029/2006WR005779

Tan, D. G. H., Andersson, E., Fisher, M., & Isaksen, L. (2007). Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM–Aeolus wind profiling mission. *Quarterly Journal of the Royal Meteorological Society*, *133*(623), 381-390. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.43` doi: https://doi.org/10.1002/qj.43

Tangdamrongsub, N., Han, S.-C., Yeo, I.-Y., Dong, J., Steele-Dunne, S., Willgoose, G., & Walker, J. (2020). Multivariate data assimilation of GRACE, SMOS, SMAP measurements for improved regional soil moisture and groundwater storage estimates. *Advances in Water Resources*, *135*, 103477. doi: 10.1016/j.advwatres.2019.103477

Thirel, G., Martin, E., Mahfouf, J.-F., Massart, S., Ricci, S., & Habets, F. (2010). A past discharges assimilation system for ensemble streamflow forecasts over France–Part 1: Description and validation of the assimilation system. *Hydrology and Earth System Sciences*, *14*(8), 1623–1637.

Thomas, B., & Famiglietti, J. (2019). Identifying Climate-Induced Groundwater Depletion in GRACE Observations. *Scientific Reports*, *9*(1), 4124. Retrieved from `https://doi.org/10.1038/s41598-019-40155-y` doi: 10.1038/s41598-019-40155-y

Tian, S., Tregoning, P., Renzullo, L., Dijk, A., Walker, J., Pauwels, V., & Allgeyer, S. (2017). Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals. *Water Resources Research*, *53*(3), 1820–1840.

Tollefson, J. (2017). Race to provide commercial weather data heats up. *Nature*, *542*(7639).

Toure, A., Reichle, R., Forman, B., Getirana, A., & De Lannoy, G. (2018). Assimilation of MODIS Snow Cover Fraction Observations into the NASA Catchment Land Surface Model. *Remote Sensing*. doi: 10.3390/rs10020316

Tremolet, Y., & Auligne, T. (2020). The Joint Effort for Data Assimilation Integration (JEDI). *JCSDA Q*, *66*, 1–5.

Trenberth, K., Dai, A., Rasmussen, R., & Parsons, D. (2003). The Changing Character of Precipitation. *Bulletin of the American Meteorological Society*, *84*(9), 1205–1218. doi: 10.1175/BAMS-84-9-1205

Trenberth, K., Dai, A., Schrier, G., Jones, P., Barichivich, J., Briffa, K., & Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, *4*(1), 17–22. doi: 10.1038/nclimate2067

Trenberth, K., Smith, L., Qian, T., Dai, A., & 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. doi: 10.1175/JHM600.1

Tsai, W.-P., Feng, D., Pan, M., Beck, H., Lawson, K., & Yang, Y. (2021). From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling. *Nature Communications*, *12*(1), 5988. Retrieved from `https://doi.org/10.1038/s41467-021-26107-z` doi: 10.1038/s41467-021-26107-z

van Leeuwen, P. J. (2009). Particle Filtering in Geophysical Systems. *Monthly*

Weather Review, *137*(12), 4089 - 4114. Retrieved from https://journals.ametsoc.org/view/journals/mwre/137/12/2009mwr2835.1.xml doi: 10.1175/2009MWR2835.1

van Leeuwen, P. J. (2015). Representation errors and retrievals in linear and nonlinear data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *141*(690), 1612-1623. Retrieved from https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2464 doi: https://doi.org/10.1002/qj.2464

van Leeuwen, P. J., Künsch, H. R., Nerger, L., Potthast, R., & Reich, S. (2019). Particle filters for high-dimensional geoscience applications: A review. *Quarterly Journal of the Royal Meteorological Society*, *145*(723), 2335-2365. Retrieved from https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3551 doi: https://doi.org/10.1002/qj.3551

Vargas Godoy, M., Markonis, Y., Hanel, M., Kysely, J., & Papalexiou, S. (2021). The Global Water Cycle Budget: A Chronological Review. *Surveys in Geophysics*, *42*(5), 1075–1107. doi: 10.1007/s10712-021-09652-6

Vitart, F., Balsamo, G., Bidlot, J.-R., Lang, S., Tsonevsky, I., Richardson, D., & Alonso-Balmaseda, M. (2019, 02). *Use of ERA5 to Initialize Ensemble Re-forecasts* (No. 841). ECMWF. Retrieved from https://www.ecmwf.int/node/18872 doi: 10.21957/w8i57wuz6

Vorosmarty, C., & Sahagian, D. (2000). Anthropogenic Disturbance of the Terrestrial Water Cycle. *BioScience*, *50*(9), 753–765. doi: 10.1641/0006-3568(2000)050

Wainwright, C., Black, E., & Allan, R. (2021). Future Changes in Wet and Dry Season Characteristics in CMIP5 and CMIP6 Simulations. *Journal of Hydrometeorology*, *22*(9), 2339–2357. doi: 10.1175/JHM-D-21-0017.1

Weerts, A., & El Serafy, G. (2006). Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. *Water Resources Research*, *42*(9).

Wegiel, J., Kumar, S., Peters-Lidard, C., Kemp, E., Geiger, J., & Wang, S. (2020). Data assimilation enhancements to Air Force Weather's Land Information System. In *Conference on Hydrology.*

Weiss, M., de Beaufort, L., Baret, F., Allard, D., Bruguier, N., & Marloie, O. (2001). Mapping leaf area index measurements at different scales for the validation of large swath satellite sensors: first results of the VALERI project. In *Proceed-*

ings of the 8th international symposium on physical measurements and signatures in remote sensing (pp. 8–12).

Wood, A., Lettenmaier, D., & Palmer, R. (1997). Assessing climate change implications for water resources planning. In *Climate Change and Water Resources Planning Criteria* (pp. 203–228). Springer.

Wrzesien, M., Kumar, S., Vuyovich, C., Gutmann, E., Kim, R., & Forman, B. (2022). Development of a Nature Run for Observing System Simulation Experiments (OSSEs) for Snow Mission Development. *Journal of Hydrometeorology*, *23*(3), 351–375. doi: 10.1175/JHM-D-21-0071.1

Wu, W.-Y., Lo, M.-H., Wada, Y., Famiglietti, J., Reager, J., & Yeh, P.-F. (2020). Divergent effects of climate change on future groundwater availability in key mid-latitude aquifers. *Nature Communications*, *11*(1), 3710. Retrieved from `https://doi.org/10.1038/s41467-020-17581-y` doi: 10.1038/s41467-020-17581-y

Xue, Y., & Forman, B. (2017). Integration of satellite-based passive microwave brightness temperature observations and an ensemble-based land data assimilation framework to improve snow estimation in forested regions. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 311–314). IEEE.

Xue, Y., Forman, B., & Reichle, R. (2018). Estimating snow mass in North America through assimilation of AMSR-E brightness temperature observations using the Catchment land surface model and support vector machines. *Water Resources Research*, *54*(9), 6488–6509. doi: 10.1029/2017WR022219

Yilmaz, M., & Crow, W. (2013). The Optimality of Potential Rescaling Approaches in Land Data Assimilation. *Journal of Hydrometeorology*, *14*(2), 650–660. doi: 10.1175/JHM-D-12-052.1

Yilmaz, M., DelSole, T., & Houser, P. (2011). Improving Land Data Assimilation Performance with a Water Budget Constraint. *Journal of Hydrometeorology*, *12*(5), 1040–1055. doi: 10.1175/2011JHM1346.1

Yin, J., Zhan, X., Zheng, Y., Hain, C., Ek, M., & Wen, J. (2016). Improving Noah land surface model performance using near real time surface albedo and green vegetation fraction. *Agricultural and Forest Meteorology*, *218–219*, 171–183. doi: 10.1016/j.agrformet.2015.12.001

Ying, Y. (2020). Assimilating Observations with Spatially Correlated Errors Using a Serial Ensemble Filter with a Multiscale Approach. *Monthly Weather Review*, *148*(8), 3397 - 3412. Retrieved from https://journals.ametsoc.org/view/journals/mwre/148/8/mwrD190387.xml doi: 10.1175/MWR-D-19-0387.1

Zaitchik, B., & Rodell, M. (2009). Forward-Looking Assimilation of MODIS-Derived Snow-Covered Area into a Land Surface Model. *Journal of Hydrometeorology*, *10*(1), 130–148. doi: 10.1175/2008JHM1042.1

Zaitchik, B., Rodell, M., & Reichle, R. (2008). Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin. *Journal of Hydrometeorology*, *9*(3), 535–548. doi: 10.1175/2007JHM951.1

Zhang, L., Wang, J., Bai, Z., & Lv, C. (2015). Effects of vegetation on runoff and soil erosion on reclaimed land in an opencast coal-mine dump in a loess area. *Catena*, *128*, 44–53.

Zhang, Y.-F., Hoar, T. J., Yang, Z.-L., Anderson, J. L., Toure, A. M., & Rodell, M. (2014). Assimilation of MODIS snow cover through the Data Assimilation Research Testbed and the Community Land Model version 4. *Journal of Geophysical Research: Atmospheres*, *119*(12), 7091-7103. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD021329 doi: https://doi.org/10.1002/2013JD021329

Zhao, L., & Yang, Z.-L. (2018). Multi-sensor land data assimilation: Toward a robust global soil moisture and snow estimation. *Remote Sensing of Environment*, *216*, 13–27. doi: 10.1016/j.rse.2018.06.033

Zhou, J., Wu, Z., Crow, W., Dong, J., & He, H. (2020). Improving spatial patterns prior to land surface data assimilation via model calibration using SMAP surface soil moisture data. *Water Resources Research*, *56*(10), 2020 027770.

Zsoter, E., Cloke, H., Stephens, E., de Rosnay, P., Munoz-Sabater, J., Prudhomme, C., & Pappenberger, F. (2019). How Well Do Operational Numerical Weather Prediction Configurations Represent Hydrology? *Journal of Hydrometeorology*, *20*(8), 1533–1552. doi: 10.1175/JHM-D-18-0086.1

Zuo, H., Balmaseda, M., Tietsche, S., Mogensen, K., & Mayer, M. (2019). The ECMWF operational ensemble reanalysis-analysis system for ocean and sea ice: a description of the system and assessment. *Ocean Science*, *15*(3), 779–808.