Regional and Global Land Data Assimilation Systems: Innovations, Challenges, and Prospects

Youlong XIA[1], Zengchao HAO[2*], Chunxiang SHI[3], Yaohui LI[4], Jesse MENG[1], Tongren XU[5], Xinying WU[2], and Baoqing ZHANG[6]

1 *I.M. Systems Group at Environmental Modeling Center (EMC), National Centers for Environmental Prediction (NCEP), National Oceanic and Atmospheric Administration (NOAA), College Park, MD* 20740, *USA*
2 *College of Water Sciences, Beijing Normal University, Beijing* 100875, *China*
3 *National Meteorological Information Center, China Meteorological Administration, Beijing* 100081, *China*
4 *Institute of Arid Meteorology, China Meteorological Administration, Lanzhou* 730020, *China*
5 *State Key Laboratory of Earth Surface Processes and Resource Ecology, School of Natural Resources, Faculty of Geographical Science, Beijing Normal University, Beijing* 100875, *China*
6 *Key Laboratory of Western China's Environmental Systems of Ministry of Education, College of Earth and Environmental Sciences, Lanzhou University, Lanzhou* 730000, *China*

(Received November 9, 2018; in final form February 26, 2019)

ABSTRACT

Since the North American and Global Land Data Assimilation Systems (NLDAS and GLDAS) were established in 2004, significant progress has been made in development of regional and global LDASs. National, regional, project-based, and global LDASs are widely developed across the world. This paper summarizes and overviews the development, current status, applications, challenges, and future prospects of these LDASs. We first introduce various regional and global LDASs including their development history and innovations, and then discuss the evaluation, validation, and applications (from numerical model prediction to water resources management) of these LDASs. More importantly, we document in detail some specific challenges that the LDASs are facing: quality of the in-situ observations, satellite retrievals, reanalysis data, surface meteorological forcing data, and soil and vegetation databases; land surface model physical process treatment and parameter calibration; land data assimilation difficulties; and spatial scale incompatibility problems. Finally, some prospects such as the use of land information system software, the unified global LDAS system with nesting concept and hyper-resolution, and uncertainty estimates for model structure, parameters, and forcing are discussed.

Key words: land data assimilation system (LDAS), regional and global LDASs, in-situ observation, satellite retrieval, land surface model (LSM)

Citation: Xia, Y. L., Z. C. Hao, C. X. Shi, et al., 2019: Regional and global land data assimilation systems: Innovations, challenges, and prospects. *J. Meteor. Res.*, **33**(2), 159–189, doi: 10.1007/s13351-019-8172-4.

1. Introduction

The accuracy of a land surface model (LSM) simulation is largely affected by many errors that exist in surface meteorological forcing data (that drive the model), model parameters (e.g., soil and hydrologic parameters), and model structure (e.g., specific soil layer vs. varied soil layers). Therefore, it is a challenging task for an LSM to produce relatively accurate initial conditions for numerical weather and climate prediction and to simulate reasonable energy budget, water cycle, and biogeochemical cycle. With the increase of in-situ observations and remote sensing data, the accuracy of the LSM can be improved by assimilating these data into the model (Kumar et al., 2014). Based on this concept, the land data assimilation system (LDAS) was developed by integrating the advanced LSM with data assimilation (DA) techniques. The LDAS uses gauge-based and remotely-

Supported by the US Environmental Modeling Center (EMC) Land Surface Modeling Project (granted to Youlong Xia) and National Natural Science Foundation of China (51609111, granted to Baoqing Zhang).

*Corresponding author: haozc@bnu.edu.cn.

sensed observations and/or reanalysis land surface meteorological forcing data to drive the LSM, and uses DA techniques to ingest ground-measured or satellite-based data into the LSM to generate the merging products.

In the beginning, the North American and Global LDASs were mainly used to provide optimal initial states and fluxes to parent atmospheric models to enhance the prediction skill of land–atmosphere coupled numerical weather/climate systems (Mitchell et al., 2004; Rodell et al., 2004). The LDASs can be classified as national, regional, and global ones according to the target domain size. Depending on whether there exist interactions between LDASs and parent atmospheric models, LDASs can then be divided into uncoupled and coupled ones (or weakly coupled ones; see details in Section 2). Normally, the LDAS products include surface energy fluxes (e.g., sensible heat flux, latent heat flux, net radiation, ground heat flux, etc.), water fluxes (e.g., evaporation, transpiration, surface runoff and baseflow, and sublimation), and land state variables (soil moisture, soil temperature, snow water equivalent, canopy water storage, groundwater, and surface temperature). With the development and expansion of various LDASs, more useful products have been generated for a variety of applications, such as drought and flood monitoring, agricultural crop management and planning, wildfire early warning, and water resources management.

There are many uncoupled regional systems, such as the European LDAS (Jacobs et al., 2008), North American LDAS (NLDAS; Mitchell et al., 2004; Xia et al., 2012a, b), South American LDAS (de Goncalves et al., 2006), Central Asia and East/Southern/West Africa LDAS (McNally et al., 2017), and so on. Besides the regional systems, a number of uncoupled national LDASs also exist, such as the Canadian LDAS (CaLDAS; Carrera et al., 2015), China LDAS (CLDAS; Li et al., 2007), China Meteorological Administration (CMA) LDAS (Shi et al., 2011), and Korean LDAS (Lim et al., 2012); and on top of these is the global LDAS (GLDAS; Rodell et al., 2004). Two examples of the global weakly coupled LDASs are the NCEP GLDAS (Meng et al., 2012) and ECMWF LDAS (de Rosnay et al., 2014; de Rosnay, 2017; Albergel et al., 2018). The regional/national uncoupled LDASs are widely and rapidly developed as the surface meteorological forcing data, such as gauge-observed precipitation, air temperature, and radiation, can be more easily and accurately obtained for the national/regional domains than on a global scale (which involves difficulty in sharing data among countries). Furthermore, the development of regional/national LDASs has been effectively accelerated due to more affordable computational resources for higher spatial resolution LDASs at regional/national scales, more homogeneous datasets of climate, hydrology, vegetation, and soil characteristics, and more available high-quality observations, as well as increased governmental funding support in various countries. The advance in project-based LDASs depends heavily on their funding agencies and resources; however, they contribute significantly to the development of the regional, national, and global LDASs.

In the early stage, all LDAS systems do not assimilate any satellite information to improve the state variables due to the lack and low quality of satellite retrievals. With the increasing availability of ground-based and remotely-sensed data and advances in DA techniques, LDAS is becoming an efficient tool to improving LSM simulations and products, and to enhancing numerical forecast skills at various timescales (daily, weekly, subseasonal, and seasonal). In this study, we overview various LDAS systems including national, regional, project-based, and global LDASs and their initiatives, developments, and current status. The applications of LDAS products to weather forecast and subseasonal-to-seasonal (S2S) prediction, such as drought and flood monitoring, agricultural crop management, wild fire monitoring, and water resources management, are also introduced. This review covers general topics on uncoupled and weakly coupled LDASs, without specific comparison of various LDASs' performance (e.g., strengths and weaknesses). Detailed analysis of the shortcomings and strengths of various LDASs require in-depth knowledge of the forcing data (e.g., reanalysis data, observations, satellite retrievals, etc.), different LSMs and related physical parameterizations, and the DA algorithms and assimilated variables. They are beyond the scope of the present paper, but will be assessed by ongoing works and the LDAS collaborators and developers in the future.

In this paper, Sections 2 and 3 summarize regional, national, project-based, and global LDAS systems, respectively. Section 4 illustrates how to comprehensively evaluate LDAS products, in particular, using NLDAS as an example. The applications of LDAS products to numerical weather and seasonal climate models, drought monitoring, and agricultural crop and water resource management are discussed in Section 5. The challenges, discussions, and conclusions are presented in Sections 6, 7, and 8, respectively.

2. Regional, national, and project-based LDASs

The LDASs can be divided into various categories according to different classification standards such as the

domain size (e.g., national, regional, and global), spatial resolution (high vs. coarse resolution), application purpose (e.g., operational, research-based, project-based, etc.), and coupling mode between the LSM and parent atmospheric model (e.g., uncoupled and coupled). Note that following Penny et al. (2017) and Penny and Hamill (2017), the coupled LDAS here in this paper actually refers to the weakly coupled LDAS. A schematic diagram illustrating the differences between the uncoupled and weakly coupled LDASs is given in Fig. 1. As indicated in Fig. 1a, the uncoupled LDASs use given meteorological forcing data [derived from an atmospheric model, e.g., a General Circulation Model (GCM)] as well as soil and vegetation parameters to drive LSMs; and the outputs from the LDASs do not feed back in any way to the atmospheric model. As shown in Fig. 1b, for the weakly coupled LDASs, the LSMs are run with forecast forcing data from the atmospheric model and with some of the forecast variables (e.g., precipitation) being replaced by observations; the products of the LDASs such as moisture and temperature are then used as initial and boundary conditions to re-run the coupled system (atmospheric model + LDAS); and the outputs of the coupled system will then be fed back to the LDASs in the next time step. The direct impacts of observations on the analysis are limited to the region where the observations exist. Cross-region impacts are produced as a secondary effect by the integration of the coupled model forecast. Here, we mainly discuss regional, national, and project-based LDASs regardless of resolution, application, and coupling methodology (Tables 1a, 1b).

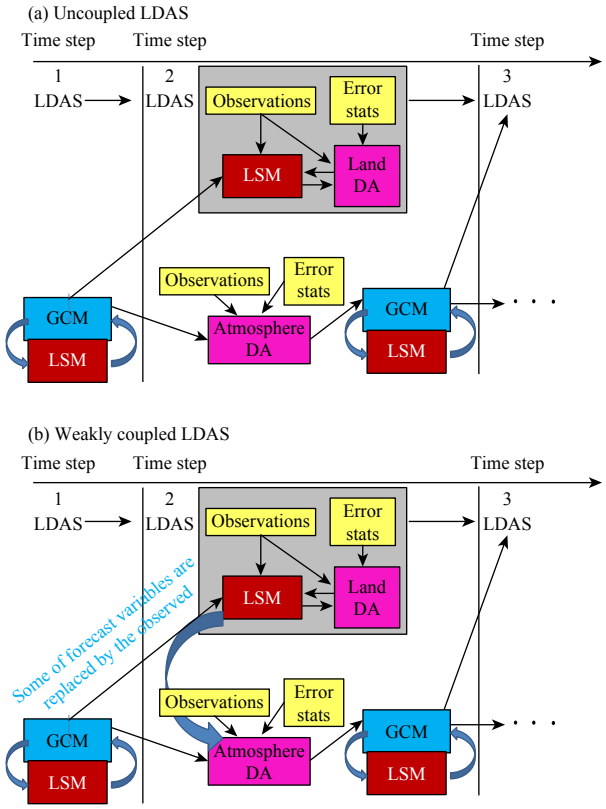

Fig. 1. Schematic diagrams describing the interaction between an LDAS and a coupled GCM model: (a) uncoupled LDAS and (b) weakly coupled LDAS. The LDAS uses some DA techniques (e.g., simple insertion, ensemble Kalman filter/smoother, etc.) to integrate past forecasts with observations to improve numerical weather and climate prediction skill (GCM: General Circulation Model; DA: Data Assimilation; LDAS: Land DA System; LSM: Land Surface Model).

The domain of a regional LDAS can extend over several countries. The NLDAS, which includes the continental United States, southern Canada, and northern Mexico, was initiated in 1998 as a multi-institution joint project among several government agencies and universities (Mitchell et al., 1999) and is one of the most successful uncoupled LDASs. After five-year collaborative efforts, the NLDAS configuration was established in 2004 (Mitchell et al., 2004). Some overview papers summarized the generation and validation of NLDAS forcing (Cosgrove et al., 2003a; Luo et al., 2003), production and validation of NLDAS streamflow and water budget (Lohmann et al., 2004), evaluation of energy budget, soil moisture, temperature, and snowpack (Pan et al., 2003; Robock et al., 2003; Sheffield et al., 2003; Schaake et al., 2004), and model spin-up testing (Cosgrove et al., 2003b). Four LSMs are executed in uncoupled mode in NLDAS, including the Noah (Ek et al., 2003), Mosaic (Koster and Saurez, 1994), Sacramento Soil Moisture Accounting (SAC-SMA; Burnash et al., 1973), and Variable Infiltration Capacity (VIC; Liang et al., 1994). These models are run at 0.125° spatial resolution and 1-h temporal resolution to produce water fluxes, energy fluxes, and state variables.

The purpose of the first phase of NLDAS (NLDAS-1) was to establish an NLDAS framework, and thus it was run for only a 3-yr period from 1 October 1996 to 30 September 1999. Although such a short-term run may be sufficient for evaluating forcing and model outputs, it is insufficient for practical applications such as drought monitoring, long-term trend analysis, and crop and water resources management. To overcome this weakness, the second phase of NLDAS (NLDAS-2; Xia et al., 2012a, b) was developed, based on the NLDAS-1 framework and was improved with a long-term run period from 1 January 1979 to present. The precipitation data derived from a daily gauge-only precipitation analysis are bias-corrected for consideration of the impacts of topography using a Parameter-elevation Regression on Independent Slopes Model (PRISM). The daily precipitation

Table 1a. Summary of various LDAS systems and their products over the world (CaLDAS: Canadian LDAS; CLDAS: China LDAS; CMA: China Meteorological Administration; CONUS: Continental United States; ELDAS: European LDAS; ERA-Interim: ECMWF Re-Analysis Interim data; ERA-5: The fifth generation ECMWF Re-Analysis; FLDAS: Famine Early Warning Systems Network LDAS; GLDAS: Global LDAS; HRLDAS: High-Resolution LDAS; IFS: Integrated Forecast System; NASA: National Aeronautics and Space Administration; NCA: National Climate Assessment; NLDAS: North American LDAS; SLDAS: South American LDAS)

Name	Domain, resolution	Data period	Reference	Website
Regional LDAS systems				
ELDAS	Europe, 0.2°	1 May–31 October 2000	Jacobs et al., 2008	N/A
NLDAS	CONUS, southern Canada, northern Mexico, 0.125°	1 January 1979–present	Mitchell et al., 2004; Xia et al., 2012a, b	EMC: http://www.emc.ncep.noaa.gov/mmb/nldas/ NASA: https://ldas.gsfc.nasa.gov/nldas/
SALDAS	South America, 40 km	17 March 2001–16 March 2002	de Goncalves et al., 2006	N/A
National LDAS systems				
CaLDAS	Canada, 1 km, 40 km	1 June–30 September 2009	Carrera et al., 2015	N/A
CLDAS	China, 0.25°	1 October 2002–31 December 2006	Li et al., 2007	N/A
CMA LDAS	China, 0.06°	1 January 2008–present	Shi et al., 2011	N/A
Korean LDAS	East Asia, 10 km	1 August 2004–12 December 2006	Lim et al., 2012	N/A
Project-based LDAS systems				
NASA FLDAS	Central Asia (1 km) and East/Southern/West Africa, 0.1°–0.25°	1 January 1981–present	McNally et al., 2017	https://lis.gsfc.nasa.gov/projects/fewsnet
NCA LDAS	NLDAS domain, 0.125°	1 January 1979–31 December 2018	Kumar et al., 2018	https://ldas.gsfc.nasa.gov/NCA-LDAS/
NCAR HRLDAS	Central US, 4 km, 12 km	1 January 2001–1 July 2002	Chen et al., 2007	https://ral.ucar.edu/solutions/products/high-resolution-land-data-assimilation-system-hrldas
Global LDAS systems				
NASA/GLDAS	Globe, 0.25°, 1.0°	1 January 1979–present for GLDAS-1; 1 January 1948–31 December 2008 for GLDAS-2	Rodell et al., 2004	https://ldas.gsfc.nasa.gov/gldas/
NCEP/GLDAS	Globe, ~38 km	1 January 1979–present	Meng et al., 2012	https://climatedataguide.ucar.edu/climate-data/climate-forecast-system-reanalysis-cfsr
ECMWF/GLDAS	Globe, 9 km, 32 km, 79 km	Real-time for 9-km operational IFS; 1 January 1950–present for 32-km ERA-5, 1 January 1979–30 November 2018 for 79-km ERA-Interim	de Rosnay et al., 2014; de Rosnay, 2017	https://software.ecmwf.int/wiki/display/LDAS/LDAS+Home

Note: Many LDASs used in the numerical weather prediction (NWP) systems of various operational centers and research institutes are not included in this table. Details of these systems are provided in Section 4, as well as relevant references.

analysis is temporally configured to produce hourly precipitation by using the NCEP Stage II hourly radar/gauge precipitation. The NLDAS-2 non-precipitation forcing data are generated from the North American Regional Reanalysis (NARR; Mesinger et al., 2006). The NARR downward shortwave radiation is corrected by using a ratio-based bias correction technique (Berg et al., 2003) and the University of Maryland's Surface Radiation Budget (SRB) dataset (Pinker et al., 2003). This is the classical uncoupled LDAS concept (Fig. 1a) where LSMs are run by using given surface meteorological forcing and the outputs from LDAS do not feed back to the atmosphere DA. Inspired by the development of NLDAS, many regional LDASs, such as the European, South America, Central Asia, and East/Southern/West Africa LDASs, were developed in the same way. More details about the development and applications of NLDAS are available at the websites of Environmental Modeling Center (EMC) (www.emc.ncep.noaa.gov/mmb/nldas/) and NASA (https://ldas.gsfc.nasa.gov/nldas).

Examples of national LDASs include the CaLDAS, CLDAS, and CMA LDAS (Table 1). The CaLDAS is an uncoupled operational LDAS with an external land surface modeling system, which provides optimal initial conditions to its parent atmospheric model. The CaLDAS uses the ensemble Kalman filter (EnKF) to assimilate the observed precipitation to improve precipitation forcing. The improved precipitation and other forcing

Table 1b. Summary of various LDAS systems over the world for their forcings, coupling ways, and application purposes (DLWR: Downward Longwave Radiation; DSWR: Downward Shortwave Radiation; CLM2: Community Land Model version 2; CLM3.5: Community Land Model version 3.5; CLSM: Catchment Land Surface Model; CMAP: CPC Merged Analysis of Precipitation; CMORPH: CPC MORPHing technique; CoLM: Common Land Model; CPC: Climate Prediction Center; EDAS: Eta Data Assimilation System; EnKF: Ensemble Kalman Filter; ERA-40: 40-yr ECMWF Re-Analysis; GDAPS: Global Data Assimilation and Prediction System; GDAS: Global Data Assimilation System; GFS: Global Forecast System; GLDAS-1: Global LDAS phase 1; GLDAS-2: Global LDAS phase 2; GOES: Geostationary Operational Environmental Satellite system; ISBA: Interactions between Soil–Biosphere–Atmosphere; MERRA-2: Modern-Era Retrospective analysis for Research and Applications Version 2; Mosaic: NASA Land Surface Model; NARR: North American Regional Reanalysis; Noah: NCEP Land Surface Model (Noah2.5.2, Noah2.7.1, Noah2.8, and Noah3.3 indicate different versions); Noah-MP: Noah Multiphysics; P: precipitation; p_s: surface pressure; q_{2m}: 2-m specific humidity; RH_{2m}: 2-m relative humidity; SAC: Sacramento Soil Moisture Accounting Model; SIB2: Simple Biosphere model version 2; SSiB: Simplified Simple Biosphere model; T_{2m}: 2-m air temperature; TESSEL: Tiled ECMWF Scheme for Surface Exchanges over Land; H-TESSEL: TESSEL with a revised land surface hydrology); U_{10m}: 10-m x-direction wind speed; V_{10m}: 10-m y-direction wind speed; VIC: Variable Infiltration Capacity, VIC4.0.3, VIC4.0.4, and VIC4.1.2 indicate different versions)

Name	Meteorological forcing	Coupling strength and DA method	Land surafce model	Application purpose
Regional LDAS systems				
ELDAS	ERA-40 reanalysis (P, DSWR, DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s)	Uncoupled, no DA	TESSEL	Research
NLDAS	NARR reanalysis (DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s), CPC gauge-based P, bias corrected DSWR	Uncoupled, no DA	Noah2.8, Mosaic, SAC, VIC4.0.3	Operation
SALDAS	NCEP GDAS (P, DSWR, DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s)	Uncoupled, no DA	SSiB	Research
National LDAS systems				
CaLDAS	Canadian precipitation analysis with observations, NWP model DSWR, DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s	Weakly; assimilate T_{2m}, RH_{2m}, snow depth and L-band brightness temperature with EnKF, 6-hourly assimilation cycle	ISBA	Operation
China LDAS	Observed P, DSWR, T_{2m}, NASA GLDAS (DLWR, q_{2m}, U_{10m}, V_{10m}, p_s)	Uncoupled; assimilate soil moisture and snow with EnKF	CoLM, SiB2	Research
CMA LDAS	P blended with gauge observations and CMORPH, analyzed T_{2m} from stations and GFS product, Satellite-retrieved DSWR, and GFS DLWR, q_{2m}, U_{10m}, V_{10m}, p_s	Uncoupled, no DA	CoLM, Noah-MP, CLM3.5	Operation
Korean LDAS	Observed P and DSWR, Korean GDAPS (DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s)	Uncoupled, no DA	Noah2.5.2	Research
Project-based LDAS systems				
NASA FLDAS	NCEP GDAS and NASA MERRA-2 reanalysis (DSWR, DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s), P blended with gauge observations and satellite retrievals	Uncoupled, no DA	Noah3.3, VIC4.1.2	Research
NCA LDAS	NARR reanalysis (DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s), CPC gauge-based P, bias corrected DSWR	Uncoupled, no DA		Research
NCAR HRLDAS	Bias-corrected radar precipitation, GOES radiation, and EDAS reanalysis	Uncoupled, no DA	Noah2.7.1	Research
Global LDAS systems				
NASA/GLDAS	NCEP GDAS (DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s), P from disaggregated CMAP, and satellite-retrieved DSWR and DLWR (the Air Force Weather Agency) for GLDAS-1; bias-corrected reanalysis (P, DSWR, DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s) generated from Princeton University (Sheffield et al., 2006)	Uncoupled, no DA	CLM2, Mosaic, Noah2.7.1, and VIC4.0.4 for GLDAS-1; CLM3.5, CLSM, Noah2.7.1, and VIC4.0.4 for GLDAS-2	Research
NCEP/GLDAS	NCEP GDAS (DSWR, DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s), blended P using gauge-based observations, satellite-retrievals, and GDAS P	Weakly, direct insert method for snow depth, 6-hourly assimilation cycle	Noah2.7.1	Operation
ECMWF/GLDAS	ECMWF Integrated Forecasting System (P, DSWR, DLWR, T_{2m}, q_{2m}, U_{10m}, V_{10m}, p_s)	Weakly, use 2-D optimal interpolation for T_{2m}, RH_{2m}, and snow depth and cover analysis, assimilate soil moisture with simplified extended Kalman filter at 6-hourly assimilation cycle	H-TESSEL	Operation

data generated from the Environment Canada numerical weather model are used to drive the Interactions between Soil–Biosphere–Atmosphere (ISBA) LSMs. The CaLDAS also assimilates L-band passive brightness air temperature by coupling the LSM with a microwave radiative transfer model to reduce surface and root zone soil moisture errors (Balsamo et al., 2007; Carrera et al., 2015).

The CLDAS is a research-based system, which was developed by Chinese scientists from the Northwest Institute of Eco-Environment and Resources and Center for Excellence in Tibetan Plateau Earth Sciences of Chinese Academy of Sciences (CAS) (Li et al., 2007; Yang et al., 2007, 2009). As an uncoupled LDAS system, the CLDAS runs multiple LSMs with multiple DA algorithms. Along with the development of CLDAS, the scientists from CMA and the CAS Institute of Atmospheric Physics developed an operational CMA LDAS to support CMA operational drought monitoring. The uncoupled CMA LDAS runs the Community Land Model version 3.5 (CLM3.5; Oleson et al., 2008), Common Land Model (CoLM; Dai et al., 2003), and Noah LSM with multiple physics options (NoahMP; Niu et al., 2011) using observed precipitation, 2-m air temperature, and other blended forcing data (from reanalysis, satellite retrievals, model outputs, etc.) at a 1/16 degree spatial resolution in China. The CMA LDAS versions 1 and 2 were operationally implemented at the CMA National Meteorological Information Center in 2013 and 2016, respectively, and have been used for agricultural drought monitoring, wild fire early warning, and water resources management. The surface meteorological forcing data have been evaluated against independent observations, and the results show good performance of CMA LDAS precipitation, compared with NASA GLDAS forcing and observations (Yang et al., 2017).

The project-based LDAS is developed for specific applications and purposes. Three examples of this type of LDASs include the NASA Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System (FLDAS; McNally et al., 2017), NASA National Climate Assessment-Land Data Assimilation System (NCA-LDAS; Kumar et al., 2018), and NCAR High-Resolution Land Data Assimilation System (HRLDAS; Chen et al., 2007). The FLDAS uses various reanalysis products as meteorological forcing data and the Land Information System framework (LIS; Kumar et al., 2006) to drive multiple LSMs to generate soil moisture, evapotranspiration (ET), runoff, and other variables for developing countries for food security assessment in data-sparse regions. The NCA-LDAS adopts the NLDAS framework (Xia et al., 2017, 2018) to run multiple LSMs to produce water fluxes, energy fluxes, and state variables. Its purpose is to use terrestrial water storage and fluxes from NCA-LDAS to evaluate selected water indicators. NCAR HRLDAS was developed for providing initial state variables for the coupled Weather Research and Forecasting (WRF) model in western US and testing weather forecast skill improvement by using optimal initial conditions. The uncoupled HRLDAS uses the same grid, Noah LSM, land use, soil texture, terrain height, time-varying vegetation fields, and model parameters as those in the coupled WRF to ensure the climatology of similar state variables. Radar-based and reanalysis meteorological forcing data are used by HRLDAS to drive the Noah LSM as they have smaller errors when compared with WRF forecast products (Chen et al., 2007). It should be noted that after the scope of these projects is over, only a few are still maintained and most of them are not further developed or maintained.

3. Global Land Data Assimilation System (GLDAS)

GLDAS is the LDAS that extends the study domain to global scale. GLDAS includes the NASA GLDAS (Rodell et al., 2004), NCEP GLDAS (Meng et al., 2012), and ECMWF GLDAS (de Rosnay et al., 2014; de Rosnay, 2017; Albergel et al., 2018). The objective of GLDAS is to use advanced DA techniques to assimilate satellite- and ground-based observational products into LSMs to generate optimal fields of land surface states and fluxes, although neither satellite-based data nor in-situ observations have been assimilated yet.

The NASA GLDAS is an uncoupled system, which has an infrastructure similar to the NLDAS but over a global domain. NASA GLDAS phase 1 (GLDAS-1) uses a combination of NCEP's Global Data Assimilation System (GDAS) atmospheric analysis fields, spatially and temporally disaggregated NOAA Climate Prediction Center Merged Analysis of Precipitation (CMAP) pentad dataset, and observation-based downward shortwave and longwave radiation fields derived from the Air Force Weather Agency's AGRicultural METeorological modeling system (AGRMET). This combination forcing is used to run the CLM, Noah, Mosaic, and VIC models (Rodell et al., 2004). The spatial resolution is 1 degree and the time period is from 1979 to present. NASA GLDAS phase 2 (GLDAS-2) has two components: one is forced entirely by the Princeton meteorological forcing data (Sheffield et al., 2006), and the other is forced by a

combination of model and observation based forcing datasets as used in phase 1. The second component uses Global Precipitation Climatology Project (GPCP) precipitation field to replace CMAP precipitation with an improved disaggregation scheme, and quality control for the AGRMET dataset. More details of the NASA GLDAS are available at https://ldas.gsfc.nasa.gov/gldas/. A recent evaluation study in China has shown that the use of different data sources for GLDAS-1 causes a temporal discontinuity, while the use of forcing data from Princeton University in GLDAS-2 has overcome this problem (Wang et al., 2016). The temporal discontinuity is a very important issue for drought monitoring and water resources management as this will lead to a sudden change in anomaly and percentile. Normally, a consistent climatology is needed for calculation of anomaly and percentile. The products from NASA GLDAS have been used to support global drought analysis and water resources management for government agencies, academia, and private sectors all over the world.

The NCEP GLDAS is a weakly coupled system, aiming to provide optimal initial conditions for the NCEP Climate Forecast System Reanalysis (CFSR; Saha et al., 2010) and Climate Forecast System version 2 (Saha et al., 2014). It uses the CFSR global atmospheric DA system and observed precipitation data (CPC unified global daily gauge analysis and pentad data of CMAP) to drive the Noah LSM on a T382 global Gaussian grid (about 38 km). It uses six streams to run 6 yr simultaneously with a 1-yr spin-up period to ensure the product completion on a manageable schedule. All six streams are connected to produce a long-term product for the period from 1979 to 2009.

An evaluation shows that the correlation between the NCEP GLDAS simulated and observed soil moisture anomaly has largely increased in Illinois (Meng et al., 2012). The errors (bias and root mean square error) between GLDAS simulated and observed soil moisture are largely reduced when compared to the NCEP–NCAR reanalysis and NLDAS soil moisture simulation (Meng et al., 2012). However, the design of the six-stream run leads to large discontinuity for many state variables, such as soil moisture, at the connection points (e.g., 31 December 1986, 31 December 1989, 31 December 1994, 31 March 1999, and 31 March 2005), because 1-yr spin-up time is not sufficient (Cosgrove et al., 2003b), compared to the recent GLDAS one-stream run and NLDAS-2 Noah simulations (Fig. 2). Recently, the NCEP LDAS team is integrating the NLDAS with GLDAS to develop the NCEP unified LDAS (NULDAS) by leveraging the advantages of NLDAS and GLDAS and the community LDAS development. The NULDAS is a multiple-model LDAS with a soil moisture and snowpack DA of 0.04 degree globally. Details of NLDAS, GLDAS, and NULDAS development plans are available in the LDAS white paper (see https://www.emc.ncep.noaa.gov/mmb/nldas/White_Paper_for_Next_Phase_LDAS_final.pdf).

The ECMWF LDAS is a weakly coupled global LDAS in the ECMWF Integrated Forecast System (IFS; de Rosnay al., 2014; de Rosnay, 2017; Albergel et al., 2018). It includes snow data analysis and soil moisture DA (Fig. 3). The purpose of the ECMWF LDAS is to provide reasonable initial conditions for IFS. The surface meteorological forcing data produced by IFS, combined with 2-m air temperature and relative humidity data from surface synoptic observations (SYNOP), are used for soil moisture DA and a snow data analysis to generate updated soil moisture, snow water equivalent, and snow temperature. These state variables are used as initial conditions for IFS for the next analysis cycle. The LDAS runs separately from the upper air analysis of IFS every six-hour cycle. The state variables will feed back into the upper air analysis and affect the short-term forecast in the next cycle. In turn, the 4D-Var atmospheric analysis affects the LDAS through the short-term forecast from one cycle to the next. More details about the ECMWF LDAS system, including publications and presentations, are provided at https://software.ecmwf.int/wiki/display/LDAS/LDAS+Home.

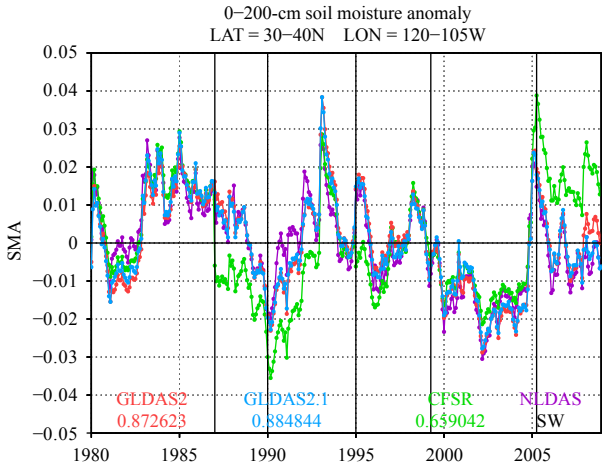

Fig. 2. Comparison of top 2-m soil moisture anomaly (SMA; $m^3\ m^{-3}$) simulated by Noah model in the NCEP GLDAS and NLDAS. Climate Forecast System Reanalysis (CFSR) uses GLDAS-1 with a six-stream run (vertical line), GLDAS-2 uses a single-stream to rerun GLDAS-1, and GLDAS-2.1 uses the newly released global 0.25° daily precipitation data from the NCEP Climate Prediction Center. The numbers are anomaly correlation coefficients between top 2-m soil moisture simulations. The domain of the southwestern U.S. covers a range of 30°–40°N, 120°–105°W.

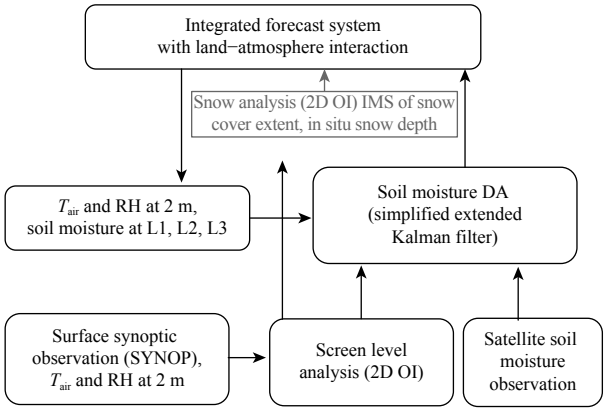

Fig. 3. A schematic diagram of the ECMWF LDAS (T_{air}: 2-m air temperature; L1, L2, L3: three soil layers; 2D OI: two-dimensional optimal interpolation; IMS: interactive multisensor snow; RH: relative humidity; SYNOP: surface synoptic observation).

4. Validation and evaluation of regional and global LDAS products

After discussing the development of national, regional, and global LDASs in Sections 2 and 3, here we will use the NLDAS as an example to briefly introduce LDAS products' validation and evaluation. In NLDAS phase 1, Mitchell et al. (2004) summarized the validation methods, tools, and datasets of NLDAS products, such as radiation (Luo et al., 2003), sensible heat flux, latent heat flux, soil temperature, soil moisture (Robock et al., 2003; Schaake et al., 2004), snow cover extent (Sheffield et al., 2003), snow water equivalent (Pan et al., 2003), land skin temperature (Mitchell et al., 2004), and streamflow (Lohmann et al., 2004). The evaluation metrics includes bias, relative bias, root-mean-square-error (RMSE), and Nash-Sutcliffe efficiency, based on different temporal scales (hourly, daily, and yearly) and different spatial scales (in-situ sites and grids, watershed, and the continent). For NLDAS phase 2, long-term data were generated and more metrics such as anomaly correlation, Taylor skill score (Taylor, 2001), and Normalized Information Contribution (NIC; Kumar et al., 2009) have also been included for the validation. Additional validation work for NLDAS-2 can be seen at the websites of NLDAS validation (https://ldas.gsfc.nasa.gov/nldas/NLDAS2valid.php) and publications (https://ldas.gsfc.nasa.gov/nldas/NLDASpublications.php).

Many datasets from in-situ observations, satellite retrievals, and reanalysis have been used to evaluate the LDAS products. These datasets are summarized in Table 2. The in-situ observations include tower flux measurements (radiation, sensible and latent heat fluxes, etc.), gauge measurements (such as streamflow, soil moisture, snow depth, and snow water equivalent). The satellite retrievals include precipitation, evapotranspiration, land surface temperature, soil moisture, snow cover, and snow water equivalent. The reanalysis products include the NCEP–NCAR Reanalysis 1 (R1; Kalnay et al., 1996), NCEP–DOE (Department of Energy) Reanalysis 2 (R2; Kanamitsu et al., 2002), CFSR (Saha et al., 2010), Twentieth-Century Reanalysis (20CR; Compo et al., 2011), MERRA (Rienecker et al., 2011), MERRA-2 (Draper et al., 2018), ECMWF Interim Re-Analysis (ERA-Interim; Dee et al., 2014), the fifth generation ECMWF atmospheric reanalysis ERA-5 (Hersbach and Dee, 2016), Japanese 25-yr (JRA-25; Onogi et al., 2007) and 55-yr Reanalysis Project (JRA-55; Kobayashi et al., 2015), and so on (see Table 2). These datasets have been widely used to assess the quality and reliability of the LDAS products. After a comprehensive evaluation procedure, the LDAS products can be used for various applications (research and operation) in academia, governmental agencies, and private enterprises. These applications will be discussed in the next section.

The evaluation and validation of the LDAS products have not only supported various applications, but also contributed to the improved quality of surface meteorological forcing data and development of LSMs' physics processes. For example, spatial discontinuity may be caused by replacing the missing GOES-derived downward shortwave radiation with the model data; thus, a bias correction method has been used to fix this problem in NLDAS-2 (Xia et al., 2013a). Furthermore, NLDAS-1 validation has improved the performance of the Noah and VIC models by tuning model parameters and upgrading model physics (Troy et al., 2008; Livneh et al., 2010; Wei et al., 2013), and so the improved Noah and VIC models are used in NLDAS-2 to generate more accurate LDAS products. In certain cases, the weakness and failure found in the LDAS validation procedure cannot be solved immediately due to a lack of the understanding of multiple factors, such as physical processes or model parameters. However, such findings can motivate the development of LSMs and the improvement of LDAS systems, including surface meteorological forcing, model physics processes, and model structures and parameters.

5. Applications of regional and global LDASs

The products from regional and global LDASs, such as soil moisture and temperature, snowpack, streamflow, and evapotranspiration, have been widely used in numer-

Table 2. Summary of in-situ observations, satellite retrievals, and reanalysis products (AmeriFlux: American Flux network; CERES: Clouds and the Earth's Radiant Energy System; Fluxnet: Flux Network; GOES: Geostationary Operational Environmental Satellite system; GEWEX/SRB: Global Energy and Water Exchange project/Surface Radiation Budget; ISMN: International Soil Moisture Network; ISTI: International Surface Temperature Initiative; LST: Land Surface Temperature; MODIS: Moderate Resolution Imaging Spectroradiometer; NCDC: National Climate Data Center; NASMD: North American Soil Moisture Database; SCAN: Soil and Climate Analysis Network; SNOTEL: Snow Telemetry; SWE: Snow Water Equivalent; USGS: U.S. Geological Survey; USCRN: U.S. Climate Reference Network)

In-situ observations				
Data name	Domain	Time step	Reference	Website
AmeriFlux	North America	Hourly, daily, monthly	Novick et al., 2018	http://ameriflux.lbl.gov/
Global FluxNet	Globe	Daily, monthly	Baldocchi et al., 2001	http://fluxnet.fluxdata.org/
USGS streamflow	US	Daily, monthly	Xia et al., 2012b	https://nwis.waterdata.usgs.gov/nwis/
Global streamflow	Globe	Daily, monthly	Beck et al., 2015	https://www.bafg.de/GRDC/EN/Home/homepage_node.html
NASMD soil moisture	North America	Daily, monthly	Quiring et al., 2016	http://soilmoisture.tamu.edu/Data/
ISMN soil moisture	Globe	Daily, monthly	Dorigo et al., 2011	https://ismn.geo.tuwien.ac.at/
ISTI LST	Globe	Daily, monthly	Rennie et al., 2014	ftp://ftp.ncdc.noaa.gov/pub/data/globaldatabank, https://www.ncdc.noaa.gov/gosic
SCAN, SNOTEL, USCRN, NCDC soil temperature	US	Hourly, daily, monthly	Hu and Feng, 2003;Schaefer et al., 2007; Bell et al., 2013; Xia et al., 2013b	http://www.wcc.nrcs.usda.gov/, https://www.ncdc.noaa.gov/crn/
SNOTEL SWE and snow depth	US	Daily	Pan et al., 2003	https://www.wcc.nrcs.usda.gov/snow/snotel-wereports.html
Global Cryosphere Watch (snow depth, SWE)	Globe	Daily, monthly	See website	https://globalcryospherewatch.org/reference/snow_inventory.php
Satellite retrievals				
Data name	Domain	Temporal and spatial resolutions	Reference	Website
Precipitation	Globe	Hourly to monthly, 0.04° to 5°	Sun et al., 2018	N/A
LandFlux-EVAL ET	Globe	Monthly, varied spatial resolution	Jiménez et al., 2011	http://www.iac.ethz.ch/group/land-climate-dynamics/research/landflux-eval.html
Satellite-derived LST (GOES, MODIS, Landsat)	Globe	GOES (hourly, 0.125°), MODIS (daily, 8-day, monthly, 1 km, 0.05°), Landsat (daily, 30 m)	Yu et al., 2009; Li et al., 2013; Wan, 2014; Parastatidis et al., 2017	https://www.ospo.noaa.gov/Products/land/glst/, https://modis.gsfc.nasa.gov/data/dataprod/mod11.php
Satellite-derived snow products	Globe	Daily, varied spatial resolution	Frei et al., 2012	https://globalcryospherewatch.org/reference/snow_inventory.php
Satellite-derived soil moisture	Globe	Daily, 30 m–50 km	Brocca et al., 2011	https://ismn.geo.tuwien.ac.at/
Satellite-derived radiation (GEWEX/SRB, CERES, GOES-R)	Globe	SRB (3-houly, daily, monthly; 1°), CERES (monthly; 1°), GOES-R (hourly; 0.25°, 0.5°)	Pinker et al., 2003; Zhang et al., 2013; 2015; Kato et al., 2018	https://gewex-srb.larc.nasa.gov/, http://ceres.larc.nasa.gov/, https://www.star.nesdis.noaa.gov/goesr/product_sw.php
Reanalysis products				
Data name	Domain	Temporal and spatial resolutions	Reference	Website
Atmospheric reanalysis	Globe	Sub-daily, daily, monthly 32 km to 2.5°	Bromwich and Wang, 2005; Uppala et al., 2005; Mesinger et al., 2006; Onogi et al., 2007; Saha et al., 2010; Rienecker et al., 2011; Dee et al., 2014; Laloyaux et al., 2016	https://climatedataguide.ucar.edu/data-type/atmospheric-reanalysis, https://climatedataguide.ucar.edu/climate-data/atmospheric-reanalysis-overview-comparison-tables

ical weather and climate prediction models (Case et al., 2011; Saha et al., 2014), water resources management (Zaitchik et al., 2010), drought monitoring and prediction (Xia et al., 2014a, b; Hao et al., 2016a), wild fire monitoring (Lewis et al., 2012), flood monitoring and management of dams (Munier et al., 2015; McNally et al., 2017), and agricultural crop management (Liu et al., 2017). A schematic diagram briefly summarizing scientific and practical applications of regional and global LDAS products, which can provide support to operational forecasters, decision makers, and research scientists in governmental agencies, academia, and the private

sectors, is shown in Fig. 4. More details are discussed below.

5.1 *Numerical weather forecast and seasonal climate prediction*

The land surface is a low-level boundary condition of the atmosphere, which affects lower-level atmospheric activities, such as redistribution of energy and water vapor, boundary layer height growth, and convective cloud formation. An LSM is usually used to simulate land surface processes, such as the continental hydrological cycle and the interaction between atmosphere and land surface on various temporal and spatial scales. The LSM is a component of the coupled GCM, which is the main tool to providing numerical weather and seasonal climate forecasts. Therefore, accurate initial states, such as soil moisture and temperature, snow water equivalent, snow cover, snow temperature, and snow density, can improve numerical weather and climate prediction. The atmosphere and land surface can interact by exchange of energy (i.e., sensible and latent fluxes) and water vapor (i.e., evapotranspiration). The LSM is driven by low-level atmospheric forcing data, such as precipitation, downward shortwave and longwave radiation, 2-m air temperature and specific humidity, 10-m wind speed, and surface pressure. As the coupled GCM produces substantial biases for precipitation and downward shortwave radiation when compared with observations, these biases often yield significant errors and drifts in soil moisture, soil temperature, and energy and water fluxes. These errors and drifts in turn affect atmospheric circulation, redistribution of energy and water vapor, as well as convective activity. The major purpose of regional and global LDASs is to ingest in-situ and remotely sensed data into the LSM model to enhance the accuracy of initial states and energy/water fluxes to improve land–atmosphere interaction and the GCM forecast/prediction.

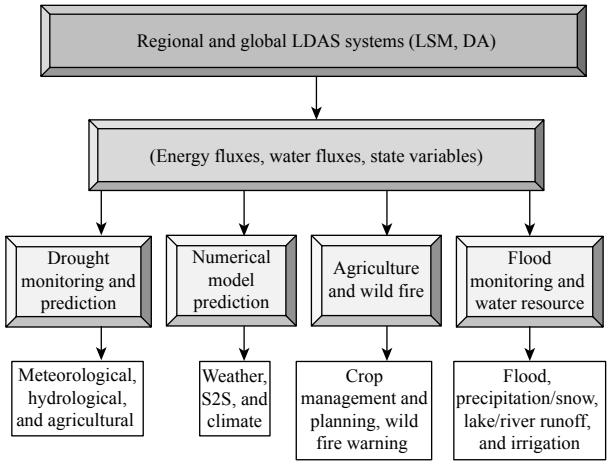

Fig. 4. A schematic diagram of regional and global LDAS outputs and their applications (S2S: subseasonal to seasonal).

Several operational centers have used their national and global LDAS systems to enhance the prediction skill of their coupled weather and seasonal climate forecast systems. The NCEP ingested the soil moisture and temperature generated from GLDAS (Meng et al., 2012) into the CFSR reanalysis (Saha et al., 2010) and CFSv2 forecast (Saha et al., 2014) to enhance seasonal simulation and forecast skill. The overall results in the CFSR show large improvements when compared with the NCEP–DOE R2 analysis and the observations. Sensitivity tests show that initial states generated from GLDAS improve CFS forecast skill when compared with initial states generated from R2 (Yang et al., 2011). The CaLDAS was developed to provide better representative land surface initial states for the environmental prediction and assimilation systems of Environmental Canada to enhance their forecast skill (Carrera et al., 2015).

The ECMWF LDAS system includes a 2-m air temperature and relative humidity analysis, snow depth and temperature analysis, soil temperature analysis, and soil moisture DA. It first uses a two-dimensional OI method to produce 2-m air temperature and relative humidity, and then the two fields are used as forcing in the DA of snow depth and soil moisture. Next, the data generated from previous steps are used for the soil and snow temperature analysis. Finally, all initial states (i.e., soil moisture and temperature, snow water equivalent, snow temperature, snow density, etc.) generated from the LDAS are provided to the ECMWF IFS (de Rosnay, 2017). In the research perspective, Case et al. (2011) improved predictions of summertime precipitation over southeastern U.S. by using LDAS-generated land surface states. Osuri et al. (2017) used LDAS-generated more realistic initial conditions to drive an NWP model to improve the prediction of severe thunderstorms over the Indian monsoon region. Santanello et al. (2016) investigated the impact of soil moisture assimilation on coupled land–atmosphere prediction, and the results showed the potential for higher-resolution soil moisture assimilation applications in weather and climate research. More details about the ECMWF LDAS system including relevant publications and presentations can be found at https://software.ecmwf.int/wiki/display/LDAS/LDAS+Home. Note that the current NCEP and ECMWF GLDASs are coupled to atmospheric GCMs every six hours rather than at each time step. Strictly speaking, such a coupling is generally weak and can be called quasi coupling. The coupling at

each time step is expected in the future.

Although the weakly coupled LDAS system has been used in numerical weather and seasonal climate systems, its application is still at an initial stage when compared with the development and applications of uncoupled LDAS. More experiments and investigations about the impact of LDAS-generated initial states on weather and climate predictions are needed in the future.

5.2 *Drought monitoring and prediction*

Drought is generally defined as a prolonged period of water deficit or water imbalance. Its characterization is commonly based on different drought indices. Drought can be categorized into four types in terms of its impact: meteorological, agricultural, hydrologic, and socioeconomic (Heim, 2002). Monitoring of different types of drought is achieved by computing drought indicators for a relatively long record. Traditional drought monitoring is generally based on in-situ meteorological observations including precipitation and temperature. Drought indicators, such as the Palmer Drought Severity Index (PDSI; Palmer, 1965) or the standardized precipitation index (SPI; McKee et al., 1993), have been developed based on these observations and employed for drought monitoring. In recent decades, advancements in remote sensing (Mu et al., 2013; AghaKouchak et al., 2015) and land surface modeling (Nijssen et al., 2014; Xia et al., 2014a) have advanced drought monitoring, which provides important monitoring information at regional or even global scales based on a variety of land surface variables (soil moisture, runoff, etc.). However, certain challenges still exist in drought monitoring based on remote sensing (e.g., errors in the estimations) or land surface simulations (e.g., uncertainties in the model simulations even with the same forcing products) (Hao et al., 2017).

To address these challenges, DA has been commonly used to merge in-situ observations, remote sensing products, LSM simulations, and climate forecasts for drought monitoring and prediction (Hao et al., 2018). A variety of LDASs developed in recent decades, including the NLDAS and GLDAS, provide multiple state and flux variables that have been widely applied to drought monitoring (Xia et al., 2014a, b; Hao et al., 2016a). Meanwhile, data products from LDAS also provide useful information for drought prediction by providing predictors or initial conditions (Hao et al., 2016b). For example, based on the recently developed Coupled Land and Vegetation Data Assimilation System (CLVDAS), initial conditions for both soil moisture and leaf area index (LAI) can be obtained for the prediction of eco-hydrological droughts (Sawada and Koike, 2016).

5.3 *Agricultural crop management*

Crop yields are important for socioeconomic development in different regions around the world, and accurate estimation of crop yields plays an important role in informed decision-making. Traditionally, agricultural crop management, including crop condition monitoring and crop yield predictions, is generally based on crop models or remote sensing observations. For the crop model, the status of crop growth can be simulated for timely crop management and decision-making to maximize crop yields. Remote sensing can offer comparatively accurate and reliable input variables for the crop model to improve simulation results.

However, it is difficult for the single crop model to expand from a single point to regional scale due to the difficulty of parameters needed, while remote sensing based products may fall short in time continuity and agriculture rationale (Li et al., 2011). The DA technique provides an important tool to integrating crop models and remotely sensed observations, thus addressing these two challenges with improved model simulations. The LDAS is mainly applied to agricultural practice in two aspects, crop growth monitoring (Ma et al., 2008; Machwitz et al., 2014) and crop yield forecasting (de Wit and van Diepen, 2007; Dente et al., 2008; Li et al., 2014), both of which are crucial for agricultural operations and management, as they can provide valuable information for precise irrigation, fertilization, irrigation, and forecast of agricultural productivity.

A commonly adopted DA scheme is to assimilate different types of observations from remote sensing products into crop models. In crop growth monitoring, many researchers have assimilated remote sensing data such as LAI (Pellenq and Boulet, 2004), biomass (Machwitz et al., 2014), and soil water (Nouvellon et al., 2001), to improve the ability of crop models in simulating crop growth. Similarly, the crop yield estimates have been improved by assimilating vegetation indices, such as LAI (Xie et al., 2017; Mokhtari et al., 2018), biomass (Jin et al., 2017), and soil moisture (Chakrabarti et al., 2017), into crop models. A simple example of using the LDAS to aid agricultural crop planning is shown in Fig. 5.

In this process of combining the assimilation of remote sensing data with crop models, both remote sensing techniques and crop models play a vital role. To simulate crop growth conditions accurately, researchers have made progress in improving crop models such as the WOFOST (World Food Studies; van Diepen et al., 1989) and DSSAT (Decision Support System for Agrotechnology Transfer; Jones et al., 2003). Meanwhile, some new

satellite products (such as Gaofen-1; Feng et al., 2016 and SPOT-6; Henry et al., 2017) as well as new tools for gaining information on crop dynamic change (e.g., the Unmanned Aerial Vehicles) have emerged. In addition, the development of radar technology has also provided data on crop growth status for assimilation. Generally, although combining DA with crop models has meant some significant improvements in the accuracy of crop models in crop growth simulation and yield prediction, there remains much scope for further improvements. For instance, how to combine different crop models and to acquire remote sensing data of higher spatial and temporal resolutions are still important for more accurate simulations of crop growth and yields.

5.4 *Flood warning and water resources management*

Hydrological models have been among the most valuable tools in simulating hydrologic processes at regional and global scales. However, these models are subject to a suite of uncertainties in the forcing data and model structures or parameters. Quantifying and reducing uncertainties are thus needed to obtain accurate and reliable hydrological forecast. Traditionally, the DA technique provides a promising potential to improve the hydrological model with higher accuracy and the quantification of uncertainty through integrating model simulations with observations (Liu et al., 2012; Khaki et al., 2018).

As one of the popular DA methods, Kalman filtering was introduced into flood forecasting around the 1970s, which was an early use of DA methods in hydrologic studies. Thereafter, the Kalman filter received extensive attention in the application of hydrologic forecasting (e.g., Kitanidis and Bras, 1980). Along with the development of ensemble theory, the ensemble Kalman filter method was developed in the late 1990s and has been widely used in flood warning and other hydrologic practices (Komma et al., 2008; Fan et al., 2017). Many of

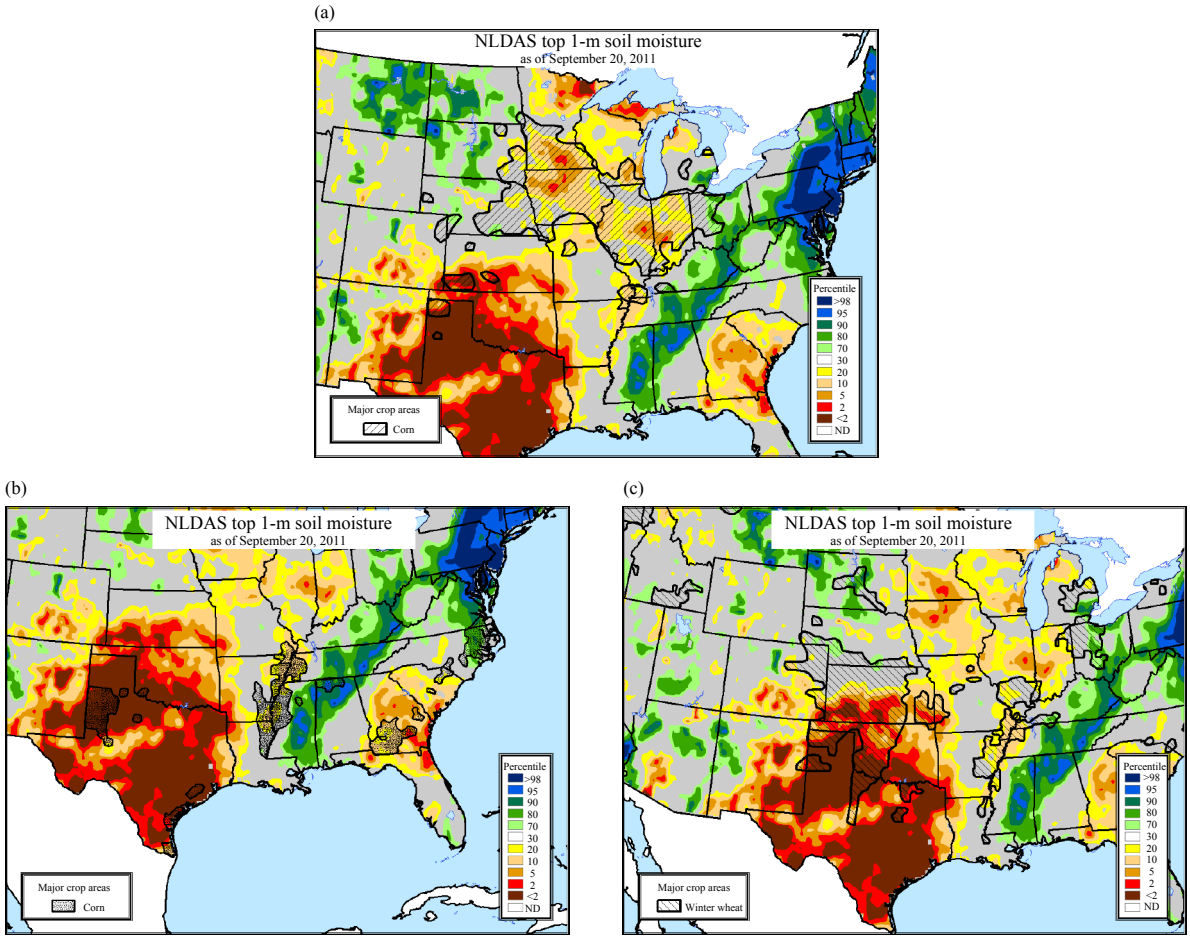

Fig. 5. Applications of three-model averaged NLDAS top 1-m soil moisture for agricultural crop planning in the USDA. (a) Used in conjunction with USDA crop-area shape files for crop-weather assessment to assess the dryness depicted in the Corn Belt of US, (b) adverse conditions afflicting cotton from Texas' drought to excessive wetness in the mid-Atlantic, and (c) dire planting prospects for winter wheat on the southern Plains of US.

these applications involve assimilation of gauge observations into hydrological models to improve the streamflow prediction. The advancement in satellite remote sensing, geographic information system development, and rain detection with digital radar has facilitated the assimilation of remote sensing products (precipitation, snow cover, soil moisture, etc.) into hydrological models (Crow and Wood, 2003; Reichle and Koster, 2005; Liu et al., 2017). Though hydrologic DA has been shown to be a useful tool for reducing predictive uncertainty and improving prediction accuracy, its application to operational hydrologic forecasting and water resources management is rather limited, partly due to lack of a clear mechanism to quantify the uncertainty in observations and forecast models and to merge data and models in an efficient and transparent way for forecasters (Liu et al., 2012).

Due to the errors/uncertainties in the atmospheric forcing (initial conditions), model structure, and parameters, the ensemble forecasting approach has gained much popularity and provided a new prospect for flood forecasting and water resources management, as this approach could prolong the effective prediction period and improve the forecast accuracy (Cloke and Pappenberger, 2009; Doycheva et al., 2017). However, a number of challenges also exist in ensemble hydrologic prediction, such as providing consistent ensemble forcing and improving uncertainty estimation in observations, which call for advances in many elements of the ensemble forecast system (Seo et al., 2014). DA has been integrated with the ensemble hydrologic forecasting for a variety of applications including the uncertainty quantification. For example, Wang et al. (2018) reduced the uncertainties of hydrologic ensemble forecasting by pre-processing and post-processing assimilation data. Recently, new assimilation schemes have been developed to address certain limitations in DA and improve hydrologic forecasting. For example, to address the challenge of computational demand with high dimensional systems in hydrologic applications, Khaki et al. (2018) proposed a nonparametric DA scheme (i.e., Kalman–Takens filter), which successfully improved the estimated water storage with less computational time. A review of the progress, challenges, and opportunities of DA in hydrologic forecasting has been provided by Liu et al. (2012).

6. Challenges and future prospects in LDAS

The national, regional, and global LDASs have been successfully developed over the past two decades and their products have been widely used in a variety of applications including NWP and drought and flood monitoring. However, there are still many challenges that may limit their further development and applications. These challenges come from poor data quality, a lack of scientific understanding of model physical processes and parameters, and spatial scale mismatches between in-situ observations and model grids. These challenges are discussed as follows.

6.1 *In-situ and remotely sensed data and reanalysis products*

In this section, we first discuss general issues in commonly used data products in hydrometeorology, as they directly affect LDAS development and the product evaluation/validation. The in-situ measurements and remotely-sensed data are the basis for LDASs and especially important for enhancing the LSM, the DA scheme, product evaluation, and LSM parameter calibration. High-quality observations are crucial for LDAS development and applications. However, the quality of observed data is affected by missing data, malfunctioning instruments and sensors, different measurement standards, and limitations of measurement techniques. In particular, the measurement error and uncertainty are unknown for ground measurements.

6.1.1 *In-situ observations*

(a) Soil moisture

For measured soil moisture, Dorigo et al. (2013) and Quiring et al. (2016) summarized various causes of data quality issues and developed automated quality control methods to flag erroneous daily data measurements (Xia et al., 2015b; Liao et al., 2019). Gruber et al. (2016) used a triple collocation method to estimate the average network error and sensor error with very limited data. These efforts support the community goal to reasonably incorporate in-situ soil moisture observations into LDAS. For soil temperature, a quality control method (Hu et al., 2002) is used to flag erroneous daily soil temperature data for 292 U.S. cooperative stations (Hu and Feng, 2003). The quality-controlled data have been used to evaluate soil temperature simulations in NLDAS-2 (Xia et al., 2013b). It should be noted that the International Soil Moisture Network (ISMN) has integrated most of the in-situ soil moisture observational network measurements into a single dataset, which is very helpful for global LDAS development and applications. However, in-situ soil temperature measurements are still archived in different countries and by different organizations. Therefore, an effort to integrate these in-situ soil temperature measurements into a single dataset as ISMN is needed in the future.

(b) Fluxes

There are very limited tower flux measurements over the globe, with a lot of missing data due to instrument/sensor failures, in particular during the cold season (September to April). In addition, because net radiation, sensible and latent heat fluxes, and ground heat flux are measured separately, their energy budgets are not balanced (Wilson et al., 2002; Foken, 2008). Even though the energy balance closure at some sites has been corrected on the monthly timescale (Jung et al., 2010), the correction and reduction of this closure error are still uncertain for all timescales, in particular for hourly and daily timescales (Shuttleworth, 2007). Some corrected and uncorrected flux data have been used to crudely evaluate NLDAS-2 ET simulations (Xia et al., 2015a).

(c) Streamflow

Streamflow measured from gauges has been widely used to evaluate/calibrate LSMs and hydrological models (Lohmann et al., 2004; Troy et al., 2008; Xia et al., 2012b). In general, the measurement errors come from observations of stream stage, periodic measurements of streamflow, and development of a rating curve to convert stage to discharge (Hamilton and Moore, 2012). Harmel et al. (2006) briefly summarized the uncertainty range in streamflow measurements by various methods for small watersheds (varying from 4% to over 100%). Furthermore, as it is an integrated variable for a given watershed/basin, the data quality is largely affected by many human activities such as dam building, regulation of water use, and irrigation. In particular, for long-term (> 30 yr) measurement, these human activities also result in data discontinuity before and after the dams are built. Although NLDAS developers have identified over 1000 small to medium size basins ($< 10^4$ km^2) with small human effects, there are still data from gauges that have some human effects (Xia et al., 2012b). To solve this problem, it is a reasonable choice that either the data error is adjusted based on known human activities or these activities are included in the LSM as physical processes.

(d) Snowfall, snow depth, and snow water equivalent

In-situ snow depth and snow water equivalent are two important variables in the development of regional and global LDASs. Systematic measurement errors caused by wind include wetting loss, evaporative loss, capping of gauge orifice, and blowing snow (Rasmussen et al., 2012). A double fence intercomparison experiment organized by World Meteorological Organization (WMO) has shown that the catch efficiency for most widely used gauges has been reduced from 100% to 20% (depending on the gauge type) when wind speed increases from 0 to 8 m s^{-1} (Yang et al., 1998). Therefore, wind bias adjustments for the measurement of snowfall in windy environments are the first step. In general, wind causes an undercatch of snowfall measurement. In contrast, blowing snow can result in overestimation of snow depth and snow water equivalent. Recently, bias correction for wind has been applied to adjust gauge precipitation, snowfall, snow depth, and snow water equivalent (Ungersböck et al., 2001; Yang et al, 2005) to reduce the systematic bias of these measured data.

6.1.2 *Remotely sensed data*

Although in-situ data are measured directly by various instruments, these data are station dependent. There are few stations in remote regions such as Africa and the polar areas. Remotely sensed data related to the water cycle (precipitation, snowpack, soil moisture, ET, etc.) are grid-based data with temporal and spatial continuity (Table 2). However, most remotely sensed data are not directly measured from satellites and they need to be retrieved from satellite images by using either empirical or physical-based algorithms. Therefore, errors come from not only satellite image errors but also uncertainties in the parameters used in these algorithms, which need to be calibrated against limited in-situ observational data. Derin and Yilmaz (2014) indicated that remotely sensed precipitation generally has difficulty in representing high spatiotemporal variability in areas with complex topography when the precipitation is controlled by orography. Infrared-based retrievals cannot capture light precipitation events and underestimate orographic rainfall amounts, whereas passive-microwave based retrievals have difficulty in detecting orographic precipitation, especially during the cold season when the precipitation is snowfall. A comparative analysis for global precipitation products, including satellite retrievals, was performed and the details about the assessment of satellite-retrieved precipitation quality can be seen in Sun et al. (2018).

(a) Snowpack

During cold season in mountainous high-latitude regions, snowpack variations manifested by snow accumulation, snowmelt and refreezing processes, and snow sublimation play a key role in the regional and global water and energy cycles. Snow depth, snow cover fraction (SCF), and snow water equivalent (SWE) are three variables retrieved mainly from satellites to monitor snowpack variations. Dietz et al. (2012) discussed various problems and drawbacks of different sensors and methods to retrieve snow cover, SCF, and SWE. The main error sources lie in the effects of cloud cover, land cover (forest, water, ice), and terrain conditions (eleva-

tion, slope, and aspect). The clouds obscure the surface and so optical sensors cannot effectively detect snow cover. Furthermore, the sensor can misinterpret between clouds and snow cover. Water bodies and lake ice can lead to false snow-cover recognition. In addition, the sensors identify densely forested regions as land cover even when ground surface is covered by snow. In mountainous regions, snow-cover and snow-free areas vary greatly due to orographic effects, and sensors have difficulty in accurately detecting snow covered areas (Dietz et al., 2012). Overall, the data quality depends on the satellites, algorithms, and variables retrieved.

(b) Soil moisture

As in-situ soil moisture observations are few and geographically sparse, remotely sensed soil moisture is playing a critical role in numerical weather and seasonal climate forecasts, agricultural crop planning, drought monitoring, and water resources management. Compared to in-situ measurements, the limitation is that only the surface soil moisture is able to be retrieved from remote sensing data (Kerr, 2007; Wagner et al., 2007; Crow et al., 2012). It is still difficult to estimate soil moisture at the root zone depth using remote sensing methods, although the remotely sensed surface soil moisture can be vertically extrapolated to constrain root zone soil moisture estimates with the use of land DA techniques (Reichle et al., 2008). In addition, soil moisture in particular is a variable with extremely high heterogeneous characteristics as it is controlled by many factors with a highly heterogeneous spatial distribution (e.g., precipitation, topography, soil texture, hydrologic conductivity, vegetation, and other surface meteorological conditions). The main error sources for remotely sensed soil moisture include vegetation density and surface roughness (Al-Yaari et al., 2014), different overpass times and retrieval algorithms, distinctive error characteristics for each sensor (Yilmaz et al., 2012), downscaling method, and related vegetation, albedo and orographic data uncertainty (Peng et al., 2017). Cui et al. (2018) evaluated 10 satellite retrieved soil moisture products (e.g., SMAP, SMOS, FY3B, AMSR2, ESA CCI, etc.) using ground-based soil moisture observations in the Little Washita Watershed network of the U.S. and the REMEDHUS network of Spain. Overall results show that the passive and enhanced passive soil moisture products of SMAP are superior to other products and have a great potential for practical applications (Cui et al., 2018).

(c) ET

For the regional and global water and energy cycles, ET is an important component that affects precipitation partitioning (ET, runoff, and terrestrial water storage) and net radiation partitioning (sensible, latent, and ground heat flux). As discussed above, the ET data measured by flux towers are very limited. The satellite retrieved and model simulated ET data are good alternatives as they are grid-distributed. The global MODIS ET product has been widely used and its mean absolute error is 24.1% of the ET measured from towers, within the range (10%–30%) of the reported uncertainties in ET measurement (Mu et al., 2011). Two other widely used datasets are global satellite-based ET products derived from the Model Tree Ensemble (MTE) method (Jung et al., 2009) and the Global Land Evaporation Amsterdam Model (GLEAM; Miralles et al., 2011; Martens et al. 2017). Jiménez et al. (2011) compared 12 global ET products, including 5 satellite retrieved and several reanalysis and offline LSM products, and the results showed large uncertainty (the all-product global annual average is approximately 45 W m^{-2} with a spread of about 20 W m^{-2}). As each ET product has its own strengths and weaknesses, merging multiple ET datasets by the ensemble mean method can generate a better ET product, although the sources of uncertainty in remote sensing ET models remain unclear (Zhang et al., 2016). The uncertainty of ET estimation with remote sensing and various surface energy balance algorithms was reviewed by Liou and Kar (2014). Broadly speaking, ET uncertainty is generally associated with LST (land surface temperature obtained from remote sensing) uncertainty, differing ET retrieval methods, ignoring nocturnal transpiration and dew, errors in near-surface meteorological variables, inaccurate vegetation data (vegetation cover, plant height, etc.), and ignoring radiation diurnal variation (Liou and Kar, 2014). However, comprehensive analysis and diagnostics for ET uncertainty are still needed in the future to improve the scientific understanding of these satellite ET retrievals.

(d) Radiation

There are many remotely sensed variables related to the surface energy cycle (e.g., incoming and outgoing radiation, land surface temperature). In general, for the GEWEX SRB monthly shortwave and longwave radiation, the errors are within 10 W m^{-2} when evaluated against ground measurements obtained from the Baseline Surface Radiation Network (BSRN). Larger errors in the downward shortwave radiation appear over the African and South American regions where aerosols from biomass burning are not accounted for in the shortwave model (Konzelman et al., 1996). The error sources also include bias and random errors due to local meteorological conditions (e.g., atmospheric humidity and cloudiness). For CERES, the uncertainty is 3.0 (4.0) W m^{-2} for

the all-sky global annual mean upward (downward) shortwave radiation. The uncertainties in upward and downward longwave irradiance are 3.0 and 6.0 W m^{-2}, respectively. Overall uncertainty in the global annual mean surface shortwave and longwave net radiation is 8.0 W m^{-2} when compared with the observations from buoys and land surface sites (Kato et al., 2018). For GOES-R, the bias of downward shortwave and longwave radiation is 3.6 and –6.8 W m^{-2}, respectively, when compared with BSRN observations. The errors may come from inaccurate surface albedo in de-clouding processes and narrowband to broadband corrections in atmospheric anisotropy, and many other sources (Liang et al., 2010).

(e) LST

LST is one of the most important parameters in representing the physical processes of surface energy and water balance at local to global scales (Bateni and Entekhabi, 2012; Xu et al., 2019). Corresponding to station-based in-situ LST observations, satellite data cover a global range and have sufficiently high temporal and spatial resolutions. LST data are retrieved from satellite thermal infrared data (GOES, Landsat, MODIS, etc.), which are directly linked through the radiative transfer equation. Li et al. (2013) reviewed the status and perspectives of satellite retrieved LST and indicated the difficulties and problems in LST retrieval from space measurements. Error sources have been investigated for the LST retrieved from thermal infrared single channel remote sensing data (Jiménez-Muñoz and Sobrino, 2006). The LST error is 2.3, 1.5, and 2.0 K for GOES (Yu et al., 2009), Landsat (Parastatidis et al., 2017), and MODIS (Wan, 2014), respectively.

6.1.3 *Reanalysis products*

Another important data source is the atmospheric reanalysis produced by various operational centers and research institutes (Table 2). The atmospheric reanalysis is generated by using a DA scheme and the models that integrate all available observations every 6–12 h during the analysis period, as real-time observation network and data may change over the duration of each reanalysis product. The changing observation mix can generate artificial variability and spurious trends. However, the quality of various reanalysis products has been proven to be comparable with satellite retrievals when evaluated against independent observations. Sun et al. (2018) compared reanalysis precipitation products with satellite-retrieved and in-situ gauge-based precipitation data and the results showed that reanalysis datasets generally had larger discrepancies when compared with the satellite-retrieved and gauge-based data. The advantage of reanalysis products is that they cover a long-term period (> 30 yr) and have a complete spatiotemporal distribution. Beck et al. (2017a) also evaluated 22 global precipitation datasets using gauge-based observations and land surface modeling. These evaluations and intercomparisons give an overall assessment of data reliability and quality. In addition, many other variables including air temperature, radiation, evapotranspiration, energy fluxes, snow pack, etc., are intercompared and assessed against satellite retrievals and gauge-based observations (Troy and Wood, 2009; Jiménez et al., 2011; Wang and Zeng, 2013; Chaudhuri et al., 2014; Snauffer et al., 2016). The results show that these reanalysis products generally capture temporal and spatial variability when compared with the observations.

6.2 *Surface meteorological forcing data*

One of the most important LDAS components is the LSM, which simulates the energy and water cycles, including interactions between the atmosphere and land surface (Fig. 1). No matter how sophisticated the LSMs are in representing land surface processes, or how accurate the boundary and initial conditions are, such LSMs will not give realistic simulation results if their forcings are not accurate (Cosgrove et al., 2003b). The accuracy of surface meteorological forcing data strongly affects simulations of soil moisture, evapotranspiration, runoff, snowpack, and sensible and latent heat fluxes, particularly in uncoupled and weakly coupled LDASs. The standard forcings to drive LSMs include precipitation, 2-m air temperature and specific humidity, downward shortwave and longwave radiation, 10-m wind speed, and surface air pressure. Such a standard was first suggested and used in the Project for Intercomparison of Land-surface Parameterization Schemes (PILPS; Henderson-Sellers et al., 1995). Forcing data can be generated for various spatial scales from a small watershed to global scale. These forcing data can be obtained from atmospheric NWP models and atmospheric reanalysis products (Table 2). However, such models, which have biases and parameterization errors, produce forcing data with a large bias (especially precipitation and solar insolation). Substantial biases in forcing data will negatively affect LSM's simulations and outputs. A more practical and robust method is to obtain as much observational forcing data as possible. As a successful regional LDAS, the NLDAS uses the best available observations and model outputs to construct quality-controlled, spatially and temporally consistent forcing data to support its multi-LSM modeling activities (Cosgrove et al., 2003a). Using a

similar strategy, Rodell et al. (2004) and Sheffield et al. (2006) generated global surface meteorological forcing datasets.

National and regional LDASs were developed more rapidly than the global LDAS. One of the reasons is that accurate forcing data can be more easily obtained as each country has its own gauge-based precipitation observational network. When observational and modeled data are merged together, the challenging issue is how to generate forcing data that are temporally and spatially homogeneous. The NLDAS forcing data generation has been carefully performed. However, there is still a spatial discontinuity in precipitation along the boundary between the U.S. and Canada, or between the U.S. and Mexico as different precipitation products are blended together. With temporal variation, a precipitation forcing discontinuity also exists before and after 1996, as different hourly products are used to temporally disaggregate CPC gauge-based daily precipitation into hourly precipitation (Ferguson and Mocko, 2017). Temporal discontinuity is also found in the NASA GLDAS-1 forcing data (Wang et al., 2016) as various models and observation data are merged temporally. Generally speaking, the quality of national and regional LDAS forcing data is better than global LDAS forcing because of the use of more high-quality observational data. Therefore, mosaic-type global LDAS forcing data can be produced by merging various national and regional LDASs, forcing them together when all national and regional high-quality forcing data are included. The challenge here is that the boundary discontinuity issue between the various national and regional LDAS domains needs to be solved.

Generally speaking, atmospheric reanalysis products are usually used as a backbone forcing to drive LSMs for many LDAS systems as reanalysis products are free from missing data and cover a long-term period, although their spatial resolution is coarse. For example, NARR data are used for the NLDAS and CFSR data are used for the NCEP GLDAS. Among the seven variables of LDAS forcing, precipitation is the most important variable for LDAS development. Gauge-based precipitation plays a key role for surface meteorological forcing generation as its quality is the best where there are sufficient rain gauge stations (e.g., midlatitude). The satellite-retrieved precipitation is an important alternative, in particular for the tropics, where there are few rain gauge stations. Furthermore, the accuracy of reanalysis precipitation data is much lower over the tropics than over the midlatitudes (Meng et al., 2012; Lee and Biasutti, 2014). In the high latitude and Polar regions, both gauge-based precipitation (with few rain gauges) and satellite-retrieved precipitation are not reliable, and thus reanalysis precipitation data are used (Meng et al., 2012). The second most important variable is downward shortwave radiation at the land surface. The reanalysis downward shortwave radiation needs to be bias corrected against the satellite-retrieved product to reduce systematic biases (Berg et al., 2003; Sheffield et al., 2006; Xia et al., 2013a). The third most important variable is 2-m air temperature, which affects the exchanges of water and energy between the atmosphere and land surface. In addition, its zero value (0°C) is usually used to simply separate precipitation into rainfall and snowfall directly within the LSMs. Small air temperature errors may misclassify precipitation into rainfall or snowfall, which will strongly affect snowpack, runoff, and soil moisture simulations. However, precipitation phase is largely affected by many meteorological variables (e.g., air temperature, surface pressure) and topographic factors (e.g., elevation, slope) as it is a quite complex process (Dai, 2008; Ding et al., 2014). Therefore, a more accurate method to separate total precipitation into rainfall and snowfall is needed in the future.

There are many difficulties and challenges in generating global forcing fields, including temporal and spatial downscaling, orographic adjustment, bias correction, merging techniques, spatiotemporal discontinuity reduction, internal coherence adjustment, and data quality control (Sheffield et al., 2006). The Princeton University Terrestrial Hydrology Research Group generated 0.25°, 0.5°, and 1° global forcing data (http://hydrology.princeton.edu/data.pgf.php) and these data have been used in the NASA GLDAS-2 to improve its performance over GLDAS-1. As LDAS is developing toward hyper-resolution (approximately 1 km; Wood et al., 2011) and high resolution (approximately 4 km, NULDAS) global land surface modeling, high resolution and high-quality observational, reanalysis, and satellite-retrieved datasets, as well as advanced merging techniques and bias-correction methods are needed. Recently, Beck et al. (2018) used a multi-source weighted-ensemble approach to merge multiple reanalysis products with several gauge-based and satellite-retrieved products, to produce global 0.1° and 3-hourly global precipitation data covering a period from 1979 to present (http://www.gloh2o.org/), based on their previous work (Beck et al., 2017a, b). For regional domains, NOAA's National Water Center is producing 4-km high-quality surface meteorological forcing data over the continental United States (CONUS) from 1979 to present (David Kitzmiller, National Water Center, personal communication), and CMA's National Meteorological Information Center is producing 6-km high-quality surface meteorological forcing data over China from

2008 to present (Yang et al., 2017). The NCEP Land–Hydrology Group is initiating a project to produce global 0.04° and hourly surface meteorological forcing as a part of the NULDAS development. All these efforts are moving regional and global LDAS systems from 0.25°–1° down to approximately 0.04° spatial resolution.

6.3 *Soil and vegetation parameters*

Soil and vegetation parameters need to be provided simultaneously to LSMs as fixed fields. In NLDAS, 16 soil texture classes were derived from the 1-km STASGO (State Soil Geographic Database) database (Miller and White, 1998) over CONUS. Outside of CONUS, 16 soil texture classes were derived from the 5-min ARS (Agricultural Research Service) FAO (Food and Agriculture Organization) global soil database (Reynolds et al., 2000). In both NCEP GLDAS and NASA GLDAS-2, the soil texture map is a hybrid of 30-s STATSGO over CONUS and 5-min FAO elsewhere. This strategy leads to spatial discontinuity at the boundary of the two databases. Recently, two global harmonized soil databases have been developed. The harmonized world soil database was first developed by Fischer et al. (2008) and the data can be downloaded from the FAO website (www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/). Five years later, Wei et al. (2014) developed a comprehensive, global 1-km soil dataset for use in LSMs by merging various regional soil databases. A sensitivity test shows that soil type has the largest effect on evapotranspiration in semi-arid regions and on runoff in wet regions, when two soil databases are used in Noah-MP LSMs (Zheng and Yang, 2016). As some soil related parameters, such as hydrologic conductivity, directly affect soil water storage, the ET and runoff processes have large uncertainties and these parameters need to be calibrated through LSMs against the observations (see Section 6.4).

Vegetation parameters mainly include vegetation classification, vegetation greenness fraction, LAI, and vegetation albedo. In general, two satellite-based global land cover datasets are used in regional and global LDASs. One such dataset is the AVHRR (Advanced Very High Resolution Radiometer on NOAA polar satellites) based data at 1-km spatial resolution with 13-class vegetation at the global scale (Hansen et al., 2000). Another dataset is the modified IGBP MODIS 20-category vegetation (land-use) data (Friedl et al., 2010; Broxton et al., 2014). The use of MODIS land cover (in particular the addition of urban and build-up categories) has improved the simulation of daytime surface temperature and sea-breeze flows in southern California (Sequera et al., 2016). Global vegetation greenness fraction and LAI data can be obtained from the NASA GLDAS website (https://ldas.gsfc.nasa.gov/gldas/GLDASlaigreen.php). NESDIS/NOAA 0.15° monthly 5-yr climatology surface albedo can be obtained from the NCAR NOAH-MP website (https://ral.ucar.edu/solutions/products/noah-multiparameterization-land-surface-model-noah-mp-lsm). Although regional and global LDASs use the same vegetation and soil class, different LSMs may use different parameter values related to soil and vegetation class (root depth, root density, number and thickness of soil layers, canopy resistance, hydrological conductivity, etc.).

Recently, a weekly real-time vegetation parameter database has been generated from the MODIS and Visible Infrared Imager Radiometer Suite (VIIRS) sensors, and its impact on coupled and uncoupled land surface modeling has been tested by Case et al. (2014) and Fang et al. (2018). However, for drought monitoring and anomaly analysis, using real-time vegetation parameters may lead to temporal discontinuities. Therefore, for drought monitoring and water resources management, a multi-year averaged land cover database is more appropriate, while for coupled numerical weather and seasonal climate prediction (in which the LSM provides initial conditions), a real-time land cover database is more appropriate. How to hybridize various global land cover databases remains a pressing challenge. More efforts along this line are needed in the future. For example, some hyper-resolution soil and vegetation datasets need to be developed over regional scales (about 100 m) (Wood et al., 2011; Yuan et al., 2018), and some experiments have demonstrated large improvement in soil moisture, terrestrial water storage, and sensible heat flux simulations (Singh et al., 2015).

6.4 *Model physical processes and parameters*

LSMs are used to simulate the exchange of surface water, energy, and carbon fluxes at the soil–atmosphere interface. Table 3 summarizes the various LSMs generally used around the world. Some LSMs have relatively simple representations of physical processes, which are used in an NWP model for short-term weather forecasts (< 14 days). Some LSMs can describe moderately complicated physical processes and can be used in numerical models for S2S (between 15 days and 3 months) prediction. Some LSMs have very sophisticated physical processes and can be used in numerical models for climate prediction (> 3 months). Even with the most complic-

ated LSMs, certain physical processes such as irrigation and groundwater pumping are still lacking, in particular for operational applications, although some sensitivity tests have been performed (Leng et al., 2013, 2015; Lawston et al., 2015). The soil physics issues such as frozen soil physics, tiling soil texture varying from depth to depth, and direct use of observed soil and hydrologic parameters (conductivity, fraction of sand, loam, clay, etc.) need further investigation.

As model complexity increases, more model parameters need to be calibrated to produce reasonable results. Theoretically, when both simple and complex LSMs are calibrated with the same observations, the complicated LSMs would perform better than the simple ones (Xia et al., 2002). However, the complex models do not necessarily perform better than simple models for the overall simulation of energy and water fluxes and state variables (Best et al., 2015), as some physical processes (e.g., biochemistry) are not associated with these variables and unrelated model parameters may bring extra uncertainties and errors. In addition, calibration of model parameters may improve one variable and degrade others. For example, Troy et al. (2008) calibrated the VIC model using USGS observed streamflow to obtain model parameters. However, the calibrated parameters improved the water and energy cycle but degraded the soil moisture simulation (Xia et al., 2018). Therefore, the multi-criteria method developed by Gupta et al. (1999) is appropriate for obtaining calibrated model parameters by accommodating multiple variables instead. When LSMs lack or unreasonably represent some physical processes, calibrated model parameters can compensate for the impact of

Table 3. LSM suite and its focused timescales and coupled systems (ESL: Earth Research Laboratory, RUC: Rapid Refresh Cycle; NWS: National Weather Service, SAC-SMA: Sacramento Soil Moisture Accounting; NOAH-MP: Noah Multi-physics; CoLM: Common Land Model; UW: University of Washington at Seattle; VIC: Variable Infiltration Capacity; TESSEL: Tiled ECMWF Scheme for Surface Exchanges over Land; UKMO: United Kingdom Meteorological Office; JULES: Joint UK Land Environment Simulator; BATS: Biosphere–Atmosphere Transfer Scheme; GFDL: Geophysical Fluid Dynamics Laboratory; LM: Land Model; CLM: Community Land Model; SIB: Simple Biosphere model; ISBA: Interactions between Soil–Biosphere–Atmosphere; NWP: Numerical Weather Prediction)

LSM name	Complexity	Physical process	Reference	Country and website
Weather forecast and NWP				
ISBA	Simple	Energy, water	Noilhan and Planton, 1989	France, https://www.umr-cnrm.fr/isbadoc/model.html
NCEP/NCAR Noah	Simple	Energy, water	Chen et al., 1997	USA, https://ral.ucar.edu/solutions/products/unified-noah-lsm
NOAA/ESL RUC	Simple	Energy, water	Smirnova et al., 1997	USA, N/A
NWS SAC-SMA	Simple	Water	Burnash et al., 1973	USA, http://www.nws.noaa.gov/ohd/hrl/hsmb/hydrology/PBE_SAC-SMA/index.html
Seasonal and S2S prediction				
NOAH-MP	Medium	Energy, water, vegetation dynamics, carbon	Niu et al., 2011	USA, https://ral.ucar.edu/solutions/products/noah-multi parameterization-land-surface-model-noah-mp-lsm
China CoLM	Medium	Energy, water, vegetation dynamics, carbon	Dai et al., 2003	China, N/A
NASA Catchment	Medium	Energy, water	Koster et al., 2000	USA, N/A
UW/Princeton VIC	Medium	Energy, water	Liang et al., 1994	USA, http://vic.readthedocs.io/en/master/
ECMWF TESSEL	Medium	Energy, water, vegetation dynamics, carbon	van den Hurk et al., 2000	UK, https://confluence.ecmwf.int//display/OIFS/3.3+OpenIFS%3A+Surface+Model+HTESSEL
UKMO JULES	Medium	Energy, water, vegetation dynamics, carbon	Best et al., 2011; Clark et al., 2011	UK, http://jules.jchmr.org/, http://jules-lsm.github.io/vn4.3/
BATS	Medium	Energy, water, carbon	Dickinson et al., 1993	USA, N/A
Climate change and climate change prediction				
NOAA/GFDL LM	Complex	Energy, water, vegetation dynamics, carbon, biochemistry	Milly et al., 2014	USA, https://www.gfdl.noaa.gov/land-model/
NCAR CLM	Complex	Energy, water, vegetation dynamics, carbon, biochemistry	Lawrence et al., 2011	USA, http://www.cesm.ucar.edu/models/cesm2/land/
SIB	Complex	Energy, water, vegetation dynamics, carbon, phenology	Sellers et al., 1986	USA, http://biocycle.atmos.colostate.edu/research/models/sib3/

missing or misrepresented parameters to improve the results. It should be noted that such improvements may not be due to model physics upgrades/improvements. Therefore, a scientific understanding of land surface processes and the addition/improvement of these processes are a key for LSM development.

6.5 *Land data assimilation*

LSMs cannot simulate energy and water fluxes accurately, because of forcing data errors, inappropriate model structure/parameters, lack of physical processes (e.g., irrigation), and lack of scientific understanding of certain physical processes. DA becomes a preferred method to ingest in-situ and satellite observations into the LSM to enhance model output accuracy. For the NLDAS project, Kumar et al. (2014) used an ensemble Kalman smoother to assimilate remotely sensed soil moisture and snow depth into the Noah LSM. The results showed marginal improvement over the open-loop configuration. The Gravity Recovery and Climate Experiment (GRACE) based terrestrial water storage anomaly is assimilated into the NASA catchment model, and the results show marginal improvement in simulation of unconfined groundwater variability over eastern U.S. and in simulation of surface and root-zone soil moisture across the US. The land DA challenges include (a) uncertainty in the definition of observation operators due to huge climatological differences between satellite-retrieved and model-forecast surface states, (b) the complexity and diversity of modeling error sources, and (c) forcing uncertainty and the characterization of background model errors. Lahoz and Schneider (2014) summarized multiple aspects of land DA challenges including data quality, methods, and LSMs. These challenges hamper land DA greatly, and affect regional and global LDAS development.

6.6 *Spatial incompatibility problems*

Spatial scale mismatch is a critical issue for regional and global LDAS development. In-situ measurements have an extensive range, from several meters to several kilometers, depending on the measured variables. However, the LSM-simulated and satellite-retrieved products have a 10–200-km spatial resolution. Model output evaluation, model parameter calibration, and DA greatly suffer from such a spatial incompatibility problem. Spatial averaging is a simple way to reduce the impact of this problem and has been used in many validation projects for soil moisture (Entin et al., 2000; Robock et al., 2003, Xia et al., 2015b) and tower fluxes (Robock et al., 2003, Xia et al., 2015a), in particular, for homogeneous underlying surfaces (fluxes) and soil texture (soil moisture). Dirmeyer et al. (2006) suggested that anomaly correlation between in-situ measurements and simulations at a nearby grid point is an appropriate metric, as the spatial averaging has little effect. However, spatial averaging does have a large impact on error verification (Dirmeyer et al., 2006).

Another reasonable method is to upscale sparse in-situ tower fluxes and soil moisture observations to a coarse resolution to reduce the spatial mismatch impact. Many methods, from a simple area-weighted average to a deep machine learning technique, have been used to upscale in-situ tower measurements to a regional scale (Liu et al., 2016; Li et al., 2018; Xu et al., 2018), and recent advances in upscaling tower eddy covariance measurements are summarized by Xiao et al. (2012). For LSM-simulated and satellite-retrieved soil moisture, Crow et al. (2012) summarized many existing soil moisture upscaling strategies. These upscaling methods can be used to reduce the detrimental impact of spatial sampling errors on the reliability of these variables with validation using ground-based observations that are spatially sparse. Recently, many approaches such as random forests (Clewley et al., 2017), Bayesian data fusion (Gao et al., 2017), and ridge regression (Kang et al., 2018) have been used to upscale in-situ soil moisture to regional scales. The upscaled soil moisture data have been used for supporting regional and global LDAS development including LSM model validation and calibration, downscaled satellite retrievals for land DA, and LDAS product evaluation (Qin et al., 2013, 2015). Upscaling strategies have a great potential for future LDAS development, in particular for reducing the impact of scale-mismatch errors for DA and evaluation work.

6.7 *Future prospects*

Along with the wide application of the Land Information System (LIS; Kumar et al., 2006) software, which is developed by NASA scientists, in the LDAS community, some future regional and global LDASs will be run within the LIS framework. The NASA LIS software includes land data development, a land information system, and a land data verification toolkit, which provides a very convenient way to run LDAS. The software can easily set up a preprocessor for multiple LSMs including soil, vegetation, and orographic parameters, as well as surface meteorological forcing (e.g., downscaling and upscaling). A DA procedure (e.g., ensemble Kalman filter or smoother) and many validation datasets and metrics within the LIS software can be used to evaluate LDAS outputs. This will make LDAS more commonly used to reduce workloads needed for setting up an LDAS.

The regional LDAS will be nested in either a hyper- or high-resolution global LDAS to simplify the system. The patch-based global LDAS, including high-resolution regional LDASs, will be an intermediate solution in the near future where high-resolution meteorological forcing data remain unavailable in mountainous and remote regions, although there are spatial discontinuities between the boundaries of various regional LDAS domains. As the spatial resolution of global coupled forecast and reanalysis systems become finer and their reanalyzed and forecast data, including satellite-retrieved data, become more accurate, a more homogeneous global LDAS is expected in the long run. Enhancement in the data quality of forcing, in-situ observations, and satellite retrievals will greatly speed up the development of LDAS. Smaller forcing errors will directly reduce LSMs output errors, and high-quality observations from in-situ stations and satellites will improve LDAS products via the land DA procedure.

The progress in LSM development, including upgrade of land surface physical processes, addition of various new physics (e.g., irrigation; groundwater pumping; tiling soil texture and topography; urban, crop, vegetation dynamics; root growth; etc.), and regionalized model parameter calibration (Xu et al., 2011, 2015; Yang et al., 2016; Mizukami et al., 2017), will also speed up the LDAS development. A major difference between regionalized and small watershed model parameter calibration is whether the spatial discontinuities of model parameters are removed. In addition, better scientific understanding of model physical processes and representation of nonlinear relationship between variables and parameters will greatly help improve LSM simulations and finally improve LDAS products.

Uncertainty estimates for model structure, model parameters, and forcing will greatly help in understanding how to improve LDAS. Nearing et al. (2016) used information theory to investigate uncertainty caused by errors in forcing data, model parameters, and model structure within the NLDAS framework. The uncertainty of simulated soil moisture mainly comes from model parameter errors, and the uncertainty of simulated ET mainly comes from forcing errors. Clark et al. (2015a, b) developed a Structure of Unified Multiple Modeling Alternatives (SUMMA) to perform a systematic model analysis, which includes detecting reasons for inter-model differences and diagnosing weaknesses of individual models in different hydro-climatic regimes. Such a deep and systematic analysis will be definitely helpful for regional and global LDAS development.

Future uncoupled and weakly coupled LDAS systems will be added to the broader earth system modeling paradigm via the Joint Effort for Data Assimilation Integration (JEDI; Penny et al., 2017). Through many coupling tests and comprehensive evaluation work including land–atmosphere interaction (e.g., energy, water, and momentum fluxes) and forecast skills, some reasonable LDAS systems can be chosen to be tested, and eventually several optimal systems can be transitioned to operational weather and climate forecast systems in the future. Overall, satellite retrievals or in-situ observations will be truly assimilated into future LDASs.

7. Discussion

It should be noted that the discussion about true land DA is simple in this review as we mostly focus on uncoupled LDAS systems without a DA procedure (Mitchell et al., 2004; Rodell et al., 2004). The DA procedure is only simply briefed because the land DA approaches and their applications to LSMs will be addressed by another planned paper in this special issue (Ming Pan, Princeton University, personal communication). There are a variety of other LDASs that have been developed in the past few decades. For example, the land DA systems used in Numerical Weather Prediction (NWP) of various operational centers include the UK Met Office's NWP system (Dharssi et al., 2011), ECMWF's uncoupled land reanalysis (Balsamo et al., 2015), NASA Global Modeling and Assimilation Office's SMAP (Soil Moisture Active Passive) level 4 system (Reichle et al., 2017a), Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2; Reichle et al., 2017b; Draper et al., 2018), Meteo-France's uncoupled land modeling system (Albergel et al., 2010; Mahfouf, 2010), and Environment Canada's uncoupled LDAS (Bélair et al., 2003a, b; Balsamo et al., 2007; Carrera et al., 2015). Details of these systems can be found in the references above and are not covered in this study. Only ECMWF and NCEP global LDAS systems, as representatives of these operational systems, are discussed in previous sections.

Intercomparison of various LDAS systems is important to understanding of their performance including strengths and weaknesses. Two previous efforts, the Project for Intercomparison of Land surface Parameterization Schemes (PILPS; Henderson-Sellers et al., 1995; Bowling et al., 2003) and Global Soil Wetness Project (GSWP; Dirmeyer et al., 1999, 2006), have been made for such an intercomparison during the past 20 years. The PILPS focuses on the intercomparison of various LSMs when the same surface meteorological forcing and soil

and vegetation parameters are used for many in-situ sites (e.g., tower sites at various climate zones and vegetation covers) on regional scales. The GSWP, on the other hand, focuses on the global scale. Following the PILPS and GSWP, the intercomparison concept is used to develop NLDAS and GLDAS for various applications in both research and operation.

In this review, intercomparison of various LDAS systems is a challenge due to the systems' complexities. The LDAS systems discussed here use different surface meteorological forcings, land surface models, soil and vegetation parameters, spatial resolutions, and coupling methods. Furthermore, they may cover different spatial domains and time periods (see Tables 1a, 1b). Mo et al. (2012) investigated the impact of different forcing data when the same VIC model is used. They found that precipitation and radiation largely affected the VIC model simulations. Mizukami et al. (2016) studied the relative impact of different forcing data and models and indicated that the choice of forcing dataset was less important than the choice of LSMs for the runoff simulations. In contrast, for evapotranspiration simulation, forcing data errors have larger impact than LSM parameters and structures (Nearing et al., 2016). Even though all conditions are the same, the difference of model versions also results in large differences for water and energy cycle simulations (Xia et al., 2018). Therefore, the intercomparison among various LDAS systems is too difficult to perform. However, it is acknowledged that such an intercomparison is valuable and needs to be implemented in the near future through national and international community efforts.

This review is not intended for recommendation of better LDAS products for users, as this type of users demand usually depends on the application regions and purposes. In general, regional LDASs may have better performance (with higher spatial resolutions) than GLDASs for a given region as more local and regional information is incorporated. The uncoupled LDASs may have better performance than weakly coupled LDASs as more observed forcing data are integrated. However, this may not necessarily hold in some regions, in particular for those with few observations.

It is recommended that independent evaluation and validation be performed by using in-situ observations for a given small water basin or region before the LDAS products are used. For the regions without observations, a simple comparison of multiple datasets including the LDAS, satellite retrievals, reanalysis, and NWP model products may be needed to decide which LDAS product will be used for decision making. For some regions, the ensemble mean of several products should be used when an uncertainty estimate is given. Many applications have been performed along this line, which has advanced our understanding of the performance of different LDASs (Jiménez et al., 2011; Decker et al., 2012; Chen et al., 2013; Wang et al., 2016; Yang et al., 2017; Bai et al., 2018).

8. Conclusions

Since the setup of NLDAS and GLDAS in 2004, the development of regional and global LDASs has demonstrated significant progress. This study summarized the past development, current status, challenges, and future prospects of LDASs. The development of LDAS theory, method, and systems requires a better quantization of model and observation errors, integrating model calibration with data assimilation by optimally using available observation information, improving the surface meteorological forcing (in-situ observations, satellite retrievals, reanalysis, etc.), updating soil and vegetation parameter datasets, advancing land surface processes (such as vegetation and carbon dynamics, interaction between groundwater and deep soil, irrigation, boundary flow in soil, and ecosystems), development of statistical and dynamical downscaling techniques, and so on.

A solid approach is needed to establish research to operation (R2O) transition through efficient collaboration between academia, governmental agencies, and private enterprises. The R2O will speed up the application of LDAS outputs by producing and providing reliable and timely outputs to decision makers in government and private sectors. The O2R (operation to research) will speed up the improvement and upgrading of LDAS in operation and meanwhile leading and facilitating new research in academia and research community. The interaction between the research and operation communities will become an efficient way to further development of LDAS. Moreover, a pathway between operational centers and application sectors needs to be established. The feedback on the quality and reliability of LDAS outputs from end users and the public will be a great help to LDAS developers at universities and institutions. In turn, the end users will greatly benefit from the modeling approaches and high-quality outputs available from the development of regional and global LDAS systems.

Acknowledgment. The authors thank Eric Luebehusen of U.S. Department of Agriculture who helped us generate Fig. 5. We acknowledge Mary Hart for proofreading and editing our first draft, Holly Norton and

Roshan Shrestha for the EMC internal review, and three anonymous reviewers for valuable comments.

REFERENCES

AghaKouchak, A., A. Farahmand, F. S. Melton, et al., 2015: Remote sensing of drought: Progress, challenges and opportunities. *Rev. Geophy.*, **53**, 452–480, doi: 10.1002/2014RG000456.

Albergel, C., J.-C. Calvet, J.-F. Mahfouf, et al., 2010: Monitoring of water and carbon fluxes using a land data assimilation system: A case study for southwestern France. *Hydrol. Earth Syst. Sci.*, **14**, 1109–1124, doi: 10.5194/hess-14-1109-2010.

Albergel, C., E. Dutra, S. Munier, et al., 2018: ERA-5 and ERA-Interim driven ISBA land surface model simulations: Which one performs better? *Hydrol. Earth Syst. Sci.*, **22**, 3515–3532, doi: 10.5194/hess-22-3515-2018.

Al-Yaari, A., J.-P. Wigneron, A. Ducharne, et al., 2014: Global-scale comparison of passive (SMOS) and active (ASCAT) satellite based microwave soil moisture retrievals with soil moisture simulations (MERRA-Land). *Remote Sen. Environ.*, **152**, 614–626, doi: 10.1016/j.rse.2014.07.013.

Bai, W. K., X. L. Gu, S. L. Li, et al., 2018: The performance of multiple model-simulated soil moisture datasets relative to ECV satellite data in China. *Water*, **10**, 1384, doi: 10.3390/w10101384.

Baldocchi, D., E.Falge, L. H. Gu, et al., 2001: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Amer. Meteor. Soc.*, **82**, 2415–2434, doi: 10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2.

Balsamo, G., J.-F. Mahfouf, S. Bélair, et al., 2007: A land data assimilation system for soil moisture and temperature: An information content study. *J. Hydrometeor.*, **8**, 1225–1242, doi: 10.1175/2007JHM819.1.

Balsamo, G., C. Albergel, A. Beljaars, et al., 2015: ERA-interim/land: A global land surface reanalysis data set. *Hydrol. Earth Syst. Sci.*, **19**, 389–407, doi: 10.5194/hess-19-389-2015.

Bateni, S. M., and D. Entekhabi, 2012: Relative efficiency of land surface energy balance components. *Water Resour. Res.*, **48**, W04510, doi: 10.1029/2011WR011357.

Beck, H. E., A. de Roo, and A. I. J. M. van Dijk, 2015: Global maps of streamflow characteristics based on observations from several thousand catchments. *J. Hydrometeor.*, **16**, 1478–1501, doi: 10.1175/JHM-D-14-0155.1.

Beck, H. E., N. Vergopolan, M. Pan, et al., 2017a: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.*, **21**, 6201–6217, doi: 10.5194/hess-21-6201-2017.

Beck, H. E., A. I. J. M. van Dijk, V. Levizzani, et al., 2017b: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.*, **21**, 589–615, doi: 10.5194/hess-21-589-2017.

Beck, H. E., E. F. Wood, M. Pan, et al., 2018: MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bull. Amer. Meteor. Soc.* doi: 10.1175/BAMS-D-17-0138.1.

Bélair, S., L.-P. Crevier, J. Mailhot, et al., 2003a: Operational implementation of the ISBA land surface scheme in the Canadian regional weather forecast model. Part I: Warm season results. *J. Hydrometeor.*, **4**, 352–370, doi: 10.1175/1525-7541(2003)4<352:OIOTIL>2.0.CO;2.

Bélair, S., R. Brown, J. Mailhot, et al., 2003b: Operational implementation of the ISBA land surface scheme in the Canadian regional weather forecast model. Part II: Cold season results. *J. Hydrometeor.*, **4**, 371–386, doi: 10.1175/1525-7541(2003)4<371:OIOTIL>2.0.CO;2.

Bell, J. E., M. A. Palecki, C. B. Baker, et al., 2013: U.S. climate reference network soil moisture and temperature observations. *J. Hydrometeor.*, **14**, 977–988, doi: 10.1175/JHM-D-12-0146.1.

Berg, A. A., J. S. Famiglietti, J. P. Walker, et al., 2003: Impact of bias correction to reanalysis products on simulations of North American soil moisture and hydrological fluxes. *J. Geophys. Res. Atmos.*, **108**, 4490, doi: 10.1029/2002JD003334.

Best, M. J., M. Pryor, D. B. Clark, et al., 2011: The Joint UK Land Environment Simulator (JULES), model description–Part 1: Energy and water fluxes. *Geosci. Model Dev.*, **4**, 677–699, doi: 10.5194/gmd-4-677-2011.

Best, M. J., G. Abramowitz, H. R. Johnson, et al., 2015: The plumbing of land surface models: Benchmarking model performance. *J. Hydrometeor.*, **16**, 1425–1442, doi: 10.1175/JHM-D-14-0158.1.

Bowling, L. C., D. P. Lettenmaier, B. Nijssen, et al., 2003: Simulation of high latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(e): 1: Experiment description and summary intercomparisons. *Glob. Planet. Change*, **38**, 1–30, doi: 10.1016/S0921-8181(03)00003-1.

Brocca, L., S. Hasenauer, T. Lacava, et al., 2011: Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sens. Environ.*, **115**, 3390–3408, doi: 10.1016/j.rse.2011.08.003.

Bromwich, D. H., and S. H. Wang, 2005: Evaluation of the NCEP-NCAR and ECMWF 15- and 40-yr reanalyses using rawinsonde data from two independent Arctic field experiments. *Mon. Wea. Rev.*, **133**, 3562–3578, doi: 10.1175/MWR3043.1.

Broxton, P. D., X. B. Zeng, D. Sulla-Menashe, et al., 2014: A global land cover climatology using MODIS data. *J. Appl. Meteor. Climatol.*, **53**, 1593–1605, doi: 10.1175/JAMC-D-13-0270.1.

Burnash, R. J. C., R. L. Ferral, and R. A. McGuire, 1973: A Generalized Streamflow Simulation System-Conceptual Modeling for Digital Computer. Technical Report, Joint Fed.–State River Forecast Cent., U. S. Natl. Weather Serv. and California Dep. of Water Resoure, Sacramento, CA, USA, 204 pp.

Carrera, M. L., S. Bélair, and B. Bilodeau, 2015: The Canadian land data assimilation system (CaLDAS): Description and synthetic evaluation study. *J. Hydrometeor.*, **16**, 1293–1314, doi: 10.1175/JHM-D-14-0089.1.

Case, J. L, S. V. Kumar, J. Srikishen, et al., 2011: Improving numerical weather predictions of summertime precipitation over the southeastern United States through a high-resolution initializationof the surface state. *Wea. Forecasting*, **26**, 785–807, doi: 10.1175/2011WAF2222455.1.

Case, J. L., F. J. Lafontaine, J. R. Bell, et al., 2014: A real-time MODIS vegetation product for land surface and numerical weather prediction models. *IEEE Trans. Geosci. Remote*

Sens., **52**, 1772–1786, doi: 10.1109/TGRS.2013.2255059.

Chakrabarti, S., T. Bongiovanni, T. Judge, et al., 2017: Assimilation of SMOS soil moisture for quantifying drought impacts on crop yield in agricultural regions. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **7**, 3867–3879, doi: 10.1109/JSTARS.2014.2315999.

Chaudhuri, A. H., R. M. Ponte, and A. T. Nguyen, 2014: A comparison of atmospheric reanalysis products for the Arctic Ocean and implications for uncertainties in air–sea fluxes. *J. Climate*, **27**, 5411–5421, doi: 10.1175/JCLI-D-13-00424.1.

Chen, F., Z. Janjic, and K. Mitchell, 1997: Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. *Bound.-Layer Meteor.*, **85**, 391–421, doi: 10.1023/A:1000531001463.

Chen, F., K. W. Manning, M. A. LeMone, et al., 2007: Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system. *J. Appl. Meteor. Climatol.*, **46**, 694–713, doi: 10.1175/JAM2463.1.

Chen, Y. Y., K. Yang, J. Qin, et al., 2013: Evaluation of AMSR-E retrievals and GLDAS simulations against observations of a soil moisture network on the central Tibetan Plateau. *J. Geophys. Res. Atmos.*, **118**, 4466–4475, doi: 10.1002/jgrd.50301.

Clark, D. B., L. M. Mercado, S. Sitch, et al., 2011: The Joint UK Land Environment Simulator (JULES), model description. Part 2: Carbon fluxes and vegetation dynamics. *Geosci. Model Dev.*, **4**, 701–722, doi: 10.5194/gmd-4-701-2011.

Clark, M. P., B. Nijssen, J. D. Lundquist, et al., 2015a: A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resour. Res.*, **51**, 2498–2514, doi: 10.1002/2015WR017198.

Clark, M. P., B. Nijssen, J. D. Lundquist, et al., 2015b: A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resour. Res.*, **51**, 2515–2542, doi: 10.1002/2015WR017200.

Clewley, D., J. B. Whitcomb, R. Akbar, et al., 2017: A method for upscaling in situ soil moisture measurements to satellite footprint scale using random forests. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **10**, 2663–2673, doi: 10.1109/JSTARS.2017.2690220.

Cloke, H. L.,and F. Pappenberger, 2009: Ensemble flood forecasting: A review. *J. Hydrol.*, **375**, 613–626, doi: 10.1016/j.jhydrol.2009.06.005.

Compo, G. P., J. S. Whitaker, P. D. Sardeshmukh, et al., 2011: The Twentieth Century Reanalysis project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, doi: 10.1002/qj.776.

Cosgrove, B. A., D. Lohmann, K. E. Mitchell, et al., 2003a: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res. Atmos.*, **108**, 8842, doi: 10.1029/2002JD003118.

Cosgrove, B. A., D. Lohmann, K. E. Mitchell, et al., 2003b: Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS). *J. Geophys. Res. Atmos.*, **108**, 8845, doi: 10.1029/2002JD003316.

Crow, W. T., and E. F. Wood, 2003: The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. *Adv. Water Resour.*, **26**, 137–149, doi: 10.1016/S0309-1708(02)00088-X.

Crow, W. T., A. A. Berg, M. H. Cosh, et al., 2012: Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Rev. Geophys.*, **50**, RG2002, doi: 10.1029/2011RG000372.

Cui, C. Y., J. Xu, and J. Y. Zeng, 2018: Soil moisture mapping from satellites: An intercomparison of SMAP, SMOS, FY3B, AMSR2, and ESA CCI over two dense network regions at different spatial scales. *Remote Sens.*, **10**, 33, doi: 10.3390/rs10010033.

Dai, A. G., 2008: Temperature and pressure dependence of the rain–snow phase transition over land and ocean. *Geophys. Res. Lett.*, **35**, L12802, doi: 10.1029/2008GL033295.

Dai, Y. J., X. B. Zeng, R. E. Dickinson, et al., 2003: The common land model. *Bull. Amer. Meteor. Soc.*, **84**, 1013–1023, doi: 10.1175/BAMS-84-8-1013.

de Goncalves, L. G. G., W. J. Shuttleworth, E. J. Burke, et al., 2006: Toward a South America land data assimilation system: Aspects of land surface model spin-up using the simplified simple biosphere. *J. Geophys. Res. Atmos.*, **111**, D17110, doi: 10.1029/2005JD006297.

de Rosnay, P., 2017: Land Surface Data for Land Surface Analysis. ECMWF Data Assimilation Training Course, ECMWF, Reading, UK, 45 pp. Available at https://software.ecmwf.int/wiki/display/LDAS/LDAS+Home?preview=/27398058/76382811/Land_satellite_NWP_SAF_TC_2017.pdf.

de Rosnay, P., G. Balsamo, C. Albergel, et al., 2014: Initialization of land surface variables for numerical weather prediction. *Surv. Geophys.*, **35**, 607–621, doi: 10.1007/s10712-012-9207-x.

de Wit, A. J. W., and C. A. van Diepen, 2007: Crop model data assimilation with the ensemble Kalman filter for improving regional crop yield forecasts. *Agric. Forest Meteor.*, **146**, 38–56, doi: 10.1016/j.agrformet.2007.05.004.

Decker, M., M. A. Brunke, Z. Wang, et al., 2012: Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *J. Climate*, **25**, 1916–1944, doi: 10.1175/JCLI-D-11-00004.1.

Dee, D. P., M. Balmaseda, G. Balsamo, et al., 2014: Toward a consistent reanalysis of the climate system. *Bull. Amer. Meteor. Soc.*, **95**, 1235–1248, doi: 10.1175/BAMS-D-13-00043.1.

Dente, L., G. Satalino, F. Mattia, et al., 2008: Assimilation of leaf area index derived from ASAR and MERIS data into CERES-wheat model to map wheat yield. *Remote Sens. Environ.*, **112**, 1395–1407, doi: 10.1016/j.rse.2007.05.023.

Derin, Y., and K. K. Yilmaz, 2014: Evaluation of multiple satellite-based precipitation products over complex topography. *J. Hydrometeor.*, **15**, 1498–1516, doi: 10.1175/JHM-D-13-0191.1.

Dharssi, I., K. J. Bovis, B. Macpherson, et al., 2011: Operational assimilation of ASCAT surface soil wetness at the Met Office. *Hydrol. Earth Syst. Sci.*, **15**, 2729–2746, doi: 10.5194/hess-15-2729-2011.

Dickinson, R. E., A. Henderson-Sellers, and P. J. Kennedy, 1993: Biosphere–Atmosphere Transfer Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model. NCAR Technical Note NCAR/TN-387+STR, NCAR, Boulder, 72 pp, doi: 10.5065/D67W6959.

Dietz, A. J., C. Kuenzer, U. Gessner, et al., 2012: Remote sensing of snow–a review of available methods. *Int. J. Remote Sens.*, **33**, 4094–4134, doi: 10.1080/01431161.2011.640964.

Ding, B. H., K. Yang, J. Qin, et al., 2014: The dependence of precipitation types on surface elevation and meteorological con-

ditions and its parameterization. *J. Hydrol.*, **513**, 154–163, doi: 10.1016/j.jhydrol.2014.03.038.

Dirmeyer, P. A., A. J. Dolman, and N. Sato, 1999: The global soil wetness project. *Bull. Amer. Meteor. Soc.*, **80**, 851–878, doi: 10.1175/1520-0477(1999)080<0851:TPPOTG>2.0.CO;2.

Dirmeyer, P. A., X. Gao, M. Zhao, et al., 2006: GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bull. Amer. Meteor. Soc.*, **87**, 1381–1398, doi: 10.1175/BAMS-87-10-1381.

Dorigo, W. A., W. Wagner, R. Hohensinn, et al., 2011: The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrol. Earth Syst. Sci.*, **15**, 1675–1698, doi: 10.5194/hess-15-1675-2011.

Dorigo, W. A., A. Xaver, M. Vreugdenhil, et al., 2013: Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. *Vadose Zone Journal*, **12**, 1–21, doi: 10.2136/vzj2012.0097.

Doycheva, K., G. Horn, C. Koch, et al., 2017: Assessment and weighting of meteorological ensemble forecast members based on supervised machine learning with application to runoff simulations and flood warning. *Adv. Eng. Inform.*, **33**, 427–439, doi: 10.1016/j.aei.2016.11.001.

Draper, C. S., R. H. Reichle, and R. D. Koster, 2018: Assessment of MERRA-2 land surface energy flux estimates. *J. Climate*, **31**, 671–691, doi: 10.1175/JCLI-D-17-0121.1.

Ek, M. B., K. E. Mitchell, Y. Lin, et al., 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res. Atmos.*, **108**, 8851, doi: 10.1029/2002JD00 3296.

Entin, J. K., A. Robock, K. Y. Vinnikov, et al., 2000: Temporal and spatial scales of observed soil moisture variations in the extratropics. *J. Geophys. Res. Atmos.*, **105**, 11865–11877, doi: 10.1029/2000JD900051.

Fan, Y. R., G. H. Huang, B. W. Baetz, et al., 2017: Development of integrated approaches for hydrological data assimilation through combination of ensemble Kalman filter and particle filter methods. *J. Hydrol.*, **550**, 412–426, doi: 10.1016/j.jhydrol.2017.05.010.

Fang, L., X. W. Zhan, C. R. Hain, et al., 2018: Impact of using near real-time green vegetation fraction in Noah land surface model of NOAA NCEP on numerical weather predictions. *Adv. Meteor.,* doi: 10.1155/2018/9256396.

Feng, L., J. Li, W. S. Gong, et al., 2016: Radiometric cross-calibration of Gaofen-1 WFV cameras using Landsat-8 OLI images: A solution for large view angle associated problems. *Remote Sens. Environ.*, **174**, 56–68, doi: 10.1016/j.rse.2015.11.031.

Ferguson, C. R., and D. M. Mocko, 2017: Diagnosing an artificial trend in NLDAS-2 afternoon precipitation. *J. Hydrometeor.*, **18**, 1051–1070, doi: 10.1175/JHM-D-16-0251.1.

Fischer, G., F. Nachtergaele, S. Prieler, et al., 2008: Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy. Available at www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/. Accessed on 31 March 2019.

Foken, T., 2008: The energy balance closure problem: An overview. *Ecol. Appl.*, **18**, 1351–1367, doi: 10.1890/06-0922.1.

Frei, A., M. Tedesco, S. Lee, et al., 2012: A review of global satellite-derived snow products. *Adv. Space Res.*, **50**, 1007–1029, doi: 10.1016/j.asr.2011.12.021.

Friedl, M. A., D. Sulla-Menashe, B. Tan, et al., 2010: MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.*, **114**, 168–182, doi: 10.1016/j.rse.2009.08.016.

Gao, S. G., Z. L. Zhu, H. T. Weng, et al., 2017: Upscaling of sparse *in situ* soil moisture observations by integrating auxiliary information from remote sensing. *Int. J. Remote Sens.*, **38**, 4782–4803, doi: 10.1080/01431161.2017.1320444.

Gruber, A., C.–H. Su, W. T. Crow, et al., 2016: Estimating error cross-correlations in soil moisture data sets using extended collocation analysis. *J. Geophys. Res. Atmos.*, **121**, 1208–1219, doi: 10.1002/2015JD024027.

Gupta, H. V., L. A. Bastidas, S. Sorooshian, et al., 1999: Parameter estimation of a land surface scheme using multicriteria methods. *J. Geophys. Res. Atmos.*, **104**, 19491–19503, doi: 10.1029/1999JD900154.

Hamilton, A. S., and R. D. Moore, 2012: Quantifying uncertainty in streamflow records. *Can. Water Resour. J.*, **37**, 3–21, doi: 10.4296/cwrj370186.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, et al., 2000: Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sen.*, **21**, 1331–1364, doi: 10.1080/014311600210209.

Hao, Z. C., F. H. Hao, Y. L. Xia, et al., 2016a: A statistical method for categorical drought prediction based on NLDAS-2. *J. Appl. Meteor. Climatol.*, **55**, 1049–1061, doi: 10.1175/JAMC-D-15-0200.1.

Hao, Z. C., Y. Hong, Y. L. Xia, et al., 2016b: Probabilistic drought characterization in the categorical form using ordinal regression. *J. Hydrol.*, **535**, 331–339, doi: 10.1016/j.jhydrol.2016.01.074.

Hao, Z. C., X. Yuan, Y. L. Xia, et al., 2017: An overview of drought monitoring and prediction systems at regional and global scales. *Bull. Amer. Meteor. Soc.*, **98**, 1879–1896, doi: 10.1175/BAMS-D-15-00149.1.

Hao, Z. C., V. P. Singh, and Y. L. Xia, 2018: Seasonal drought prediction: Advances, challenges, and future prospects. *Rev. Geophys.*, **56**, 108–141, doi: 10.1002/2016RG000549.

Harmel, R. D., R. J. Cooper, R. M. Slade, et al., 2006: Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Transactions of the ASABE*, **49**, 689–701, doi: 10.13031/2013.20488.

Heim, Jr. R. R., 2002: A review of twentieth-century drought indices used in the United States. *Bull. Amer. Meteor. Soc.*, **83**, 1149–1166, doi: 10.1175/1520-0477-83.8.1149.

Henderson-Sellers, A., A. J. Pitman, P. K. Love, et al., 1995: The project for intercomparison of land surface parameterization schemes (PILPS): Phases 2 and 3. *Bull. Amer. Meteor. Soc.*, **76**, 489–504, doi: 10.1175/1520-0477(1995)076<0489:TPFIOL>2.0.CO;2.

Henry, F., D. E. Herwindiati, S. Mulyono, et al., 2017: Sugarcane land classification with satellite imagery using logistic regression model. *IOP Conference Series: Materials Science and Engineering*, **185**, 012024, doi: 10.1088/1757-899X/185/1/012024.

Hersbach, H., and D. Dee, 2016: ERA5 reanalysis is in production.

ECMWF Newsletter, **147**, 1–7.

Hu, Q., and S. Feng, 2003: A daily soil temperature dataset and soil temperature climatology of the contiguous United States. *J. Appl. Meteor.*, **42**, 1139–1156, doi: 10.1175/1520-0450(2003)042<1139:ADSTDA>2.0.CO;2.

Hu, Q., S. Feng, and G. Schaefer, 2002: Quality control for USDA NRCS SM-ST network soil temperatures: A method and a dataset. *J. Appl. Meteor.*, **41**, 607–619, doi: 10.1175/1520-0450(2002)041<0607:QCFUNS>2.0.CO;2.

Jacobs, C. M. J., E. J. Moors, H. W. Ter Maat, et al., 2008: Evaluation of European Land Data Assimilation System (ELDAS) products using in situ observations. *Tellus A*, **60**, 1023–1037, doi: 10.1111/j.1600-0870.2008.00351.x.

Jiménez, C., C. Prigent, B. Mueller, et al., 2011: Global intercomparison of 12 land surface heat flux estimates. *J. Geophys. Res. Atmos.*, **116**, D02102, doi: 10.1029/2010JD014545.

Jiménez-Muñoz, J. C., and J. A. Sobrino, 2006: Error sources on the land surface temperature retrieved from thermal infrared single channel remote sensing data. *Int. J. Remote Sens.*, **27**, 999–1014, doi: 10.1080/01431160500075907.

Jin, X. L., Z. H. Li, G. J. Yang, et al., 2017: Winter wheat yield estimation based on multi-source medium resolution optical and radar imaging data and the AquaCrop model using the particle swarm optimization algorithm. *ISPRS J. Photogram. Remote Sens.*, **126**, 24–37, doi: 10.1016/j.isprsjprs.2017.02.001.

Jones, J. W., G. Hoogenboom, C. H. Porter, et al., 2003: The DSSAT cropping system model. *Eur. J. Agron.*, **18**, 235–265, doi: 10.1016/S1161-0301(02)00107-7.

Jung, M., M. Reichstein, and A. Bondeau, 2009: Towards global empirical upscaling of FLUXNET eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, **6**, 2001–2013, doi: 10.5194/bg-6-2001-2009.

Jung, M., M. Reichstein, P. Ciais, et al., 2010: Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, **467**, 951–954, doi: 10.1038/nature09396.

Kalnay, E., M. Kanamitsu, R. Kistler, et al., 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437–472, doi: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, et al., 2002: NCEP-DOE AMIP-II reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1644, doi: 10.1175/BAMS-83-11-1631.

Kang, J., R. Jin, X. Li, et al., 2018: Spatial upscaling of sparse soil moisture observations based on ridge regression. *Remote Sens.*, **10**, 192, doi: 10.3390/rs10020192.

Kato, S., F. G. Rose, D. A. Rutan, et al., 2018: Surface irradiances of edition 4.0 clouds and the earth's radiant energy system (CERES) energy balanced and filled (EBAF) data product. *J. Climate*, **31**, 4501–4527, doi: 10.1175/JCLI-D-17-0523.1.

Kerr, Y. H., 2007: Soil moisture from space: Where are we? *Hydrogeol. J.*, **15**, 117–120, doi: 10.1007/s10040-006-0095-3.

Khaki, M., F. Hamilton, E. Forootan, et al., 2018: Nonparametric data assimilation scheme for land hydrological applications. *Water Resour. Res.*, **54**, 4946–4964, doi: 10.1029/2018WR022854.

Kitanidis, P. K., and R. L. Bras, 1980: Real-time forecasting with a conceptual hydrologic model: 2. Applications and results. *Water Resour. Res.*, **16**, 1034–1044, doi: 10.1029/WR016i006p01034.

Kobayashi, S., Y. Ota, Y. Harada, et al., 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan Ser. II*, **93**, 5–48, doi: 10.2151/jmsj.2015-001.

Komma, J., G. Blöschl, and C. Reszler, 2008: Soil moisture updating by ensemble Kalman filtering in real-time flood forecasting. *J. Hydrol.*, **357**, 228–242, doi: 10.1016/j.jhydrol.2008.05.020.

Konzelmann, T., D. R. Cahoon, and C. H. Whitlock, 1996: Impact of biomass burning in equatorial Africa on the downward surface shortwave irradiance: Observations versus calculations. *J. Geophys. Res. Atmos.*, **101**, 22833–22844, doi: 10.1029/96JD01556.

Koster, R. D., and M. J. Suarez, 1994: The components of a 'SVAT' scheme and their effects on a GCM's hydrological cycle. *Adv. Water Resour.*, **17**, 61–78, doi: 10.1016/0309-1708(94)90024-8.

Koster, R. D., M. J. Suarez, A. Ducharne, et al., 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res. Atmos.*, **105**, 24809–24822, doi: 10.1029/2000JD900327.

Kumar, S. V., C. D. Peters-Lidard, Y. Tian, et al., 2006: Land Information System—An interoperable framework for high resolution land surface modeling. *Environ. Model. Soft.*, **21**, 1402–1415, doi: 10.1016/j.envsoft.2005.07.004.

Kumar, S. V., R. H. Reichle, R. D. Koster, et al., 2009: Role of subsurface physics in the assimilation of surface soil moisture observations. *J. Hydrometeor.*, **10**, 1534–1547, doi: 10.1175/2009JHM1134.1.

Kumar, S. V., C. D. Peters-Lidard, D. Mocko, et al., 2014: Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation. *J. Hydrometeor.*, **15**, 2446–2469, doi: 10.1175/JHM-D-13-0132.1.

Kumar, S. V., M. Jasinski, D. Mocko, et al., 2018: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment. *J. Hydrometeor.,* doi: 10.1175/JHM-D-17-0125.1.

Lahoz, W. A., and P. Schneider, 2014: Data assimilation: Making sense of earth observation. *Front. Environ. Sci.*, **2**, 16, doi: 10.3389/fenvs.2014.00016.

Laloyaux, P., M. Balmaseda, D. Dee, et al., 2016: A coupled data assimilation system for climate reanalysis. *Quart. J. Roy. Meteor. Soc.*, **142**, 65–78, doi: 10.1002/qj.2629.

Lawrence, D. M., K. W. Oleson, M. G. Flanner, et al., 2011: Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.*, **3**, M03001, doi: 10.1029/2011MS00045.

Lawston, P. M., J. A. Santanello, Jr. B. F. Zaitchik, et al., 2015: Impact of irrigation methods on land surface model spinup and initialization of WRF forecasts. *J. Hydrometeor.*, **16**, 1135–1154, doi: 10.1175/JHM-D-14-0203.1.

Lee, D. E., and M. Biasutti, 2014: Climatology and variability of precipitation in the twentieth-century reanalysis. *J. Climate*, **27**, 5964–5981, doi: 10.1175/JCLI-D-13-00630.1.

Leng, G. Y., M. Y. Huang, Q. H. Tang, et al., 2013: Modeling the effects of irrigation on land surface fluxes and states over the conterminous United States: Sensitivity to input data and model parameters. *J. Geophys. Res. Atmos.*, **118**, 9789–9803,

doi: 10.1002/jgrd.50792.

Leng, G. Y., M. Y. Huang, Q. H. Tang, et al., 2015: A modeling study of irrigation effects on global surface water and groundwater resources under a changing climate. *J. Adv. Model. Earth Syst.*, **7**, 1285–1304, doi: 10.1002/2015MS000437.

Lewis, P., J. Gómez-Dans, T. Kaminski, et al., 2012: An earth observation land data assimilation system (EO-LDAS). *Remote Sens. Environ.*, **120**, 219–235, doi: 10.1016/j.rse.2011.12.027.

Li, R., C. J. Li, Y. Y. Dong, et al., 2011: Assimilation of remote sensing and crop model for LAI estimation based on ensemble Kalman filter. *Agric. Sci. China*, **10**, 1595–1602, doi: 10.1016/S1671-2927(11)60156-9.

Li, X., C. L. Huang, C. Tao, et al., 2007: Development of a Chinese land data assimilation system: Its progress and prospects. *Prog. Natural Sci.*, **17**, 163–173. (in Chinese)

Li, X., S. M. Liu, H. X. Li, et al., 2018: Intercomparison of six upscaling evapotranspiration methods: From site to the satellite pixel. *J. Geophys. Res. Atmos.*, **123**, 6777–6803, doi: 10.1029/2018JD028422.

Li, Y., Q. G. Zhou, J. Zhou, et al., 2014: Assimilating remote sensing information into a coupled hydrology–crop growth model to estimate regional maize yield in arid regions. *Ecological Modelling*, **291**, 15–27, doi: 10.1016/j.ecolmodel.2014.07.013.

Li, Z. L., B. H. Tang, H. Wu, et al., 2013: Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.*, **131**, 14–37, doi: 10.1016/j.rse.2012.12.008.

Liang, S. L., K. C. Wang, X. T. Zhou, et al., 2010: Review on estimation of land surface radiation and energy budgets from ground measurement, remote sensing and model simulations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **3**, 225–240, doi: 10.1109/JSTARS.2010.2048556.

Liang, X., D. P. Lettenmaier, E. F. Wood, et al., 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res. Atmos.*, **99**, 14415–14428, doi: 10.1029/94JD00483.

Liao, W. L., D. G. Wang, G. L. Wang, et al., 2019: Quality control and evaluation of the observed daily data in North American Soil Moisture Database. *J. Meteor. Res.*, **33**, , doi: 10.1007/s13351-019-8121-2.

Lim, Y.-J., K.-Y. Byun, T.-Y. Lee, et al., 2012: A land data assimilation system using the MODIS-derived land data and its application to numerical weather prediction in East Asia. *Asia–Pacific J. Atmos. Sci.*, **48**, 83–95, doi: 10.1007/s13143-012-0008-4.

Liou, Y.-A., and S. K. Kar, 2014: Evapotranspiration estimation with remote sensing and various surface energy balance algorithms—A review. *Energies*, **7**, 2821–2849, doi: 10.3390/en7052821.

Liu, S. M., Z. W. Xu, L. S. Song, et al., 2016: Upscaling evapotranspiration measurements from multi-site to the satellite pixel scale over heterogeneous land surfaces. *Agric. Forest Meteor.*, **230–231**, 97–113, doi: 10.1016/j.agrformet.2016.04.008.

Liu, X. M., T. T. Yang, K. Hsu, et al., 2017: Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau. *Hydrol. Earth Syst. Sci.*, **21**, 169–181, doi: 10.5194/hess-21-169-2017.

Liu, Y., A. H. Weerts, M. Clark, et al., 2012: Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrol. Earth Syst. Sci.*, **16**, 3863–3887, doi: 10.5194/hess-16-3863-2012.

Livneh, B., Y. L. Xia, K. E. Mitchell, et al., 2010: Noah LSM snow model diagnostics and enhancements. *J. Hydrometeor.*, **11**, 721–738, doi: 10.1175/2009JHM1174.1.

Lohmann, D., K. E. Mitchell, P. R. Houser, et al., 2004: Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project. *J. Geophys. Res. Atmos.*, **109**, D07S91, doi: 10.1029/2003JD003517.

Luo, L. F., A. Robock, K. E. Mitchell, et al., 2003: Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains. *J. Geophys. Res. Atmos.*, **108**, 8843, doi: 10.1029/2002JD003246.

Ma, Y. P., S. L. Wang, L. Zhang, et al., 2008: Monitoring winter wheat growth in North China by combining a crop model and remote sensing data. *Int. J. Appl. Earth Obs. Geoinfo.*, **10**, 426–437, doi: 10.1016/j.jag.2007.09.002.

Machwitz, M., L. Giustarini, C. Bossung, et al., 2014: Enhanced biomass prediction by assimilating satellite data into a crop growth model. *Environ. Model. Soft.*, **62**, 437–453, doi: 10.1016/j.envsoft.2014.08.010.

Mahfouf, J. F., 2010: Assimilation of satellite-derived soil moisture from ASCAT in a limited-area NWP model. *Quart. J. Roy. Meteor. Soc.*, **136**, 784–798, doi: 10.1002/qj.602.

Martens, B., D. G. Miralles, H. Lievens, et al., 2017: GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geosci. Model Dev.*, **10**, 1903–1925, doi: 10.5194/gmd-10-1903-2017.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, Am. Meteor. Soc., Anaheim, CA, USA, 179–184.

McNally, A., K. Arsenault, S. Kumar, et al., 2017: A land data assimilation system for sub-Saharan Africa food and water security applications. *Scientific Data*, **4**, 170012, doi: 10.1038/sdata.2017.12.

Meng, J., R. Q. Yang, H. L. Wei, et al., 2012: The land surface analysis in the NCEP climate forecast system reanalysis. *J. Hydrometeor.*, **13**, 1621–1630, doi: 10.1175/JHM-D-11-090.1.

Mesinger, F., G. DiMego, E. Kalnay, et al., 2006: North American regional reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, doi: 10.1175/BAMS-87-3-343.

Miller, D. A., and R. A. White, 1998: A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interaction*, **2**, 1–26, doi:10.1175/1087-3562(1998)002<0001:ACUSMS>2.3.CO;2.

Milly, P. C. D., S. L. Malyshev, E. Shevliakova, et al., 2014: An enhanced model of land water and energy for global hydrologic and earth-system studies. *J. Hydrometeor.*, **15**, 1739–1761, doi: 10.1175/JHM-D-13-0162.1.

Miralles, D. G., T. R. H. Holmes, R. A. M. de Jeu, et al., 2011: Global land-surface evaporation estimated from satellite-based observations. *Hydrol. Earth Syst. Sci.*, **15**, 453–469, doi: 10.5194/hess-15-453-2011.

Mitchell, K., P. Houser, E. Wood, et al., 1999: GCIP Land Data Assimilation System (LDAS) Project now underway.

GEWEX News, **9**, 3–6.

Mitchell, K. E., D. Lohmann, P. R. Houser, et al., 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res. Atmos.*, **109**, D07S90, doi: 10.1029/2003JD003823.

Mizukami, N., M. P. Clark, E. D. Gutmann, et al., 2016: Implications of the methodological choices for hydrologic portrayals of climate change over the contiguous United States: Statistically downscaled forcing data and hydrologic models. *J. Hydrometeor.*, **17**, 73–98, doi: 10.1175/JHM-D-14-0187.1.

Mizukami, N., M. P. Clark, A. J. Newman, et al., 2017: Towards seamless large-domain parameter estimation for hydrologic models. *Water Resour. Res.*, **53**, 8020–8040, doi: 10.1002/2017WR020401.

Mo, K. C., L. C. Chen, S. Shukla, et al., 2012: Uncertainties in North American land data assimilation systems over the contiguous United States. *J. Hydrometeor.*, **13**, 996–1009, doi: 10.1175/JHM-D-11-0132.1.

Mokhtari, A., H. Noory, and M. Vazifedoust, 2018: Improving crop yield estimation by assimilating LAI and inputting satellite-based surface incoming solar radiation into SWAP model. *Agric. Forest Meteor.*, **250–251**, 159–170, doi: 10.1016/j.agrformet.2017.12.250.

Mu, Q. Z., M. S. Zhao, and S. W. Running, 2011: Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.*, **115**, 1781–1800, doi: 10.1016/j.rse.2011.02.019.

Mu, Q. Z., M. S. Zhao, J. S. Kimball, et al., 2013: A remotely sensed global terrestrial drought severity index. *Bull. Amer. Meteor. Soc.*, **94**, 83–98, doi: 10.1175/BAMS-D-11-00213.1.

Munier, S., A. Polebistki, C. Brown, et al., 2015: SWOT data assimilation for operational reservoir management on the upper Niger River basin. *Water Resour. Res.*, **51**, 554–575, doi: 10.1002/2014WR016157.

Nearing, G. S., D. M. Mocko, C. D. Peters-Lidard, et al., 2016: Benchmarking NLDAS-2 soil moisture and evapotranspiration to separate uncertainty contributions. *J. Hydrometeor.*, **17**, 745–759, doi: 10.1175/JHM-D-15-0063.1.

Nijssen, B., S. Shukla, C. Y. Lin, et al., 2014: A prototype Global Drought Information System based on multiple land surface models. *J. Hydrometeor.*, **15**, 1661–1676, doi: 10.1175/JHM-D-13-090.1.

Niu, G. Y., Z. L. Yang, K. E. Mitchell, et al., 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res. Atmos.*, **116**, D12109, doi: 10.1029/2010JD015139.

Noilhan, J., and S. Planton, 1989: A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev.*, **117**, 536–549, doi: 10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2.

Nouvellon, Y., M. S. Moran, D. Lo Seen, et al., 2001: Coupling a grassland ecosystem model with Landsat imagery for a 10-year simulation of carbon and water budgets. *Remote Sens. Environ.*, **78**, 131–149, doi: 10.1016/S0034-4257(01)00255-3.

Novick, K. A., J. A. Biederman, A. R. Desai, et al., 2018: The AmeriFlux network: A coalition of the willing. *Agric. Forest Meteor.*, **249**, 444–456, doi: 10.1016/j.agrformet.2017.10.009.

Oleson, K. W., G.-Y. Niu, Z.-L. Yang, et al., 2008: Improvements to the Community Land Model and their impact on the hydrological cycle. *J. Geophys. Res. Biogeo.*, **113**, G01021, doi: 10.1029/2007JG000563.

Onogi, K., J. Tsutsui, H. Koide, et al., 2007: The JRA-25 reanalysis. *J. Meteor. Soc. Japan Ser. II*, **85**, 369–432, doi: 10.2151/jmsj.85.369.

Osuri, K. K., R. Nadimpalli, U. C. Mohanty, et al., 2017: Improved prediction of severe thunderstorms over the Indian monsoon region using high-resolution soil moisture and temperature initialization. *Scientific Reports*, **7**, 41377, doi: 10.1038/srep41377.

Palmer, W. C., 1965: Meteorological Drought. Research Paper No. 45, U.S. Weather Bureau, Washington, D. C., 58 pp.

Pan, M., J. Sheffield, E. F. Wood, et al., 2003: Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent. *J. Geophys. Res. Atmos.*, **108**, 8850, doi: 10.1029/2003JD003994.

Parastatidis, D., Z. Mitraka, N. Chrysoulakis, et al., 2017: Online global land surface temperature estimation from landsat. *Remote Sens.*, **9**, 1208, doi: 10.3390/rs9121208.

Pellenq, J., and G. Boulet, 2004: A methodology to test the pertinence of remote-sensing data assimilation into vegetation models for water and energy exchange at the land surface. *Agronomie*, **24**, 197–204, doi: 10.1051/agro:2004017.

Peng, J., A. Loew, O. Merlin, et al., 2017: A review of spatial downscaling of satellite remotely sensed soil moisture. *Rev. Geophys.*, **55**, 341–366, doi: 10.1002/2016RG000543.

Penny, S. G., and T. M. Hamill, 2017: Coupled data assimilation for integrated earth system analysis and prediction. *Bull. Amer. Meteor. Soc.*, **98**, ES169–ES172, doi: 10.1175/BAMS-D-17-0036.1.

Penny, S. G., S. Akella, O. Alves, et al., 2017: Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations. World Weather Research Programme (WWRP 2017–3), World Meteorological Organization, Geneva, Switzerland, 59 pp.

Pinker, R. T., J. D. Tarpley, I. Laszlo, et al., 2003: Surface radiation budgets in support of the GEWEX continental-scale international project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American land data assimilation system (NLDAS) project. *J. Geophys. Res. Atmos.*, **108**, 8844, doi: 10.1029/2002JD003301.

Qin, J., K. Yang, N. Lu, et al., 2013: Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia. *Remote Sens. Environ.*, **138**, 1–9, doi: 10.1016/j.rse.2013.07.003.

Qin, J., L. Zhao, Y. Y. Chen, et al., 2015: Inter-comparison of spatial upscaling methods for evaluation of satellite-based soil moisture. *J. Hydrol.*, **523**, 170–178, doi: 10.1016/j.jhydrol.2015.01.061.

Quiring, S .M., T. W. Ford, J. K. Wang, et al., 2016: The North American soil moisture database: Development and applications. *Bull. Amer. Meteor. Soc.*, **97**, 1441–1459, doi: 10.1175/BAMS-D-13-00263.1.

Rasmussen, R., B. Baker, J. Kochendorfer, et al., 2012: How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed. *Bull. Amer. Meteor. Soc.*, **93**, 811–829,

doi: 10.1175/BAMS-D-11-00052.1.

Reichle, R. H., and R. D. Koster, 2005: Global assimilation of satellite surface soil moisture retrievals into the NASA catchment land surface model. *Geophys. Res. Lett.*, **32**, L02404, doi: 10.1029/2004GL021700.

Reichle, R. H., W. T. Crow, R. D. Koster, et al., 2008: Contribution of soil moisture retrievals to land data assimilation products. *Geophys. Res. Lett.*, **35**, L01404, doi: 10.1029/2007GL031986.

Reichle, R. H., G. J. M. De Lannoy, Q. Liu, et al., 2017a: Assessment of the SMAP level-4 surface and root-zone soil moisture product using in situ measurements. *J. Hydrometeor.*, **18**, 2621–2645, doi: 10.1175/JHM-D-17-0063.1.

Reichle, R., Q. Liu, R. D. Koster, et al., 2017b: Land surface precipitation in MERRA-2. *J. Climate*, **30**, 1643–1664, doi: 10.1175/JCLI-D-16-0570.1.

Rennie, J. J., J. H. Lawrimore, B. E. Gleason, et al., 2014: The international surface temperature initiative global land surface databank: Monthly temperature data release description and methods. *Geosci. Data J.*, **1**, 75–102, doi: 10.1002/gdj3.8.

Reynolds, C. A., T. J. Jackson, and W. J. Rawls, 2000: Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions. *Water Resour. Res.*, **36**, 3653–3662, doi: 10.1029/2000WR900130.

Rienecker, M. M., M. J. Suarez, R. Gelaro, et al., 2011: MERRA: NASA's modern-era retrospective analysis for research and applications. *J. Climate*, **24**, 3624–3648, doi: 10.1175/JCLI-D-11-00015.1.

Robock, A., L. F. Luo, E. F. Wood, et al., 2003: Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season. *J. Geophys. Res. Atmos.*, **108**, 8846, doi: 10.1029/2002JD003245.

Rodell, M., P. R. Houser, U. Jambor, et al., 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.*, **85**, 381–394, doi: 10.1175/BAMS-85-3-381.

Saha, S., S. Moorthi, H.-L. Pan, et al., 2010: The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1058, doi: 10.1175/2010BAMS3001.1.

Saha, S., S. Moorthi, X. R. Wu, et al, 2014: The NCEP climate forecast system version 2. *J. Climate*, **27**, 2185–2208, doi: 10.1175/JCLI-D-12-00823.1.

Santanello, Jr. J. A., S. V. Kumar, C. D. Peters-Lidard, et al., 2016: Impact of soil moisture assimilation on land surface model spinup and coupled land–atmosphere prediction. *J. Hydrometeor.*, **17**, 517–540, doi: 10.1175/JHM-D-15-0072.1.

Sawada, Y., and T. Koike, 2016: Towards ecohydrological drought monitoring and prediction using a land data assimilation system: A case study on the Horn of Africa drought (2010–2011). *J. Geophy. Res. Atmos.*, **121**, 8229–8242, doi: 10.1002/2015JD024705.

Schaake, J. C., Q. Y. Duan, V. Koren, et al., 2004: An intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS). *J. Geophys. Res. Atmos.*, **109**, D01S90, doi: 10.1029/2002JD003309.

Schaefer, G. L., M. H. Cosh, and T. J. Jackson, 2007: The USDA natural resources conservation service soil climate analysis network (SCAN). *J. Atmos. Oceanic Technol.*, **24**, 2073–2077, doi: 10.1175/2007JTECHA930.1.

Sellers, P. J., Y. Mintz, Y. C. Sud, et al., 1986: A simple biosphere model (SIB) for use within general circulation models. *J. Atmos. Sci.*, **43**, 505–531, doi: 10.1175/1520-0469(1986)043<0505:ASBMFU>2.0.CO;2.

Seo, D.-J., Y. Q. Liu, H. Moradkhani, et al., 2014: Ensemble prediction and data assimilation for operational hydrology. *J. Hydrol.*, **519**, 2661–2662, doi: 10.1016/j.jhydrol.2014.11.035.

Sequera, P., J. E. González, K. McDonald, et al., 2016: Improvements in land-use classification for estimating daytime surface temperatures and sea-breeze flows in Southern California. *Earth Interaction*, **20**, 1–32, doi: 10.1175/EI-D-14-0034.1.

Sheffield, J., M. Pan, E. F. Wood, et al., 2003: Snow process modeling in the North American Land Data Assimilation System (NLDAS): 1. Evaluation of model-simulated snow cover extent. *J. Geophys. Res. Atmos.*, **108**, 8849, doi: 10.1029/2002JD003274.

Sheffield, J., G. Goteti, and E. F. Wood, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate*, **19**, 3088–3111, doi: 10.1175/JCLI3790.1.

Shi, C. X., Z. H. Xie, H. Qian, et al., 2011: China land soil moisture EnKF data assimilation based on satellite remote sensing data. *Sci. China Earth Sci.*, **54**, 1430–1440, doi: 10.1007/s11430-010-4160-3.

Shuttleworth, W. J., 2007: Putting the "vap" into evaporation. *Hydrol. Earth Syst. Sci.*, **11**, 210–244, doi: 10.5194/hess-11-210-2007.

Singh, R. S., J. T. Reager, N. L. Miller, et al., 2015: Toward hyper-resolution land-surface modeling: The effects of fine-scale topography and soil texture on CLM4.0 simulations over the Southwestern U.S. *Water Resour. Res.*, **51**, 2648–2667, doi: 10.1002/2014WR015686.

Smirnova, T. G., J. M. Brown, and S. G. Benjamin, 1997: Performance of different soil model configurations in simulating ground surface temperature and surface fluxes. *Mon. Wea. Rev.*, **125**, 1870–1884, doi: 10.1175/1520-0493(1997)125<1870:PODSMC>2.0.CO;2.

Snauffer, A. M., W. W. Hsieh, and A. J. Cannon, 2016: Comparison of gridded snow water equivalent products with in situ measurements in British Columbia, Canada. *J. Hydrol.*, **541**, 714–726, doi: 10.1016/j.jhydrol.2016.07.027.

Sun, Q., C. Miao, Q. Duan, et al., 2018: A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.*, **56**, 79–107, doi: 10.1002/rog.v56.1.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.*, **106**, 7183–7192, doi: 10.1029/2000JD900719.

Troy, T. J., and E. F. Wood, 2009: Comparison and evaluation of gridded radiation products across northern Eurasia. *Environ. Res. Lett.*, **4**, 045008, doi: 10.1088/1748-9326/4/4/045008.

Troy, T. J., E. F. Wood, and J. Sheffield, 2008: An efficient calibration method for continental-scale land surface modeling. *Water Resour. Res.*, **44**, W09411, doi: 10.1029/2007WR006513.

Ungersböck, M., F. Rubel, T. Fuchs, et al., 2001: Bias correction of global daily rain gauge measurements. *Phys. Chem. Earth B: Hydrol., Oceans Atmos.*, **26**, 411–414, doi: 10.1016/S1464-1909(01)00027-2.

Uppala, S. M., P. W. KÅllberg, A. J. Simmons, et al., 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, doi: 10.1256/qj.04.176.

van den Hurk, B. J. J. M., P. Viterbo, A. C. M. Beljaars, et al., 2000: Offline Validation of the ERA-40 Surface Scheme. ECMWF Tech. Memo., 295, European Center for Medium-Range Weather Forecasts, Reading, UK, 43 pp.

van Diepen, C. A., J. Wolf, H. van Keulen, et al., 1989: WOFOST: A simulation model of crop production. *Soil Use Manag.*, **5**, 16–24, doi: 10.1111/j.1475-2743.1989.tb00755.x.

Wagner, W., G. Blöschl, P. Pampaloni, et al., 2007: Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. *Hydrol. Res.*, **38**, 1–20, doi: 10.2166/nh.2007.029.

Wan, Z. M., 2014: New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sens. Environ.*, **40**, 36–45, doi: 10.1016/j.rse.2013.08.027.

Wang, A. H., and X. B. Zeng, 2013: Development of global hourly 0.5° land surface air temperature datasets. *J. Climate*, **26**, 7676–7691, doi: 10.1175/JCLI-D-12-00682.1.

Wang, S., B. C. Ancell, G. H. Huang, et al., 2018: Improving robustness of hydrologic ensemble predictions through probabilistic pre- and post-processing in sequential data assimilation. *Water Resour. Res.*, **54**, 2129–2151, doi: 10.1002/2018WR022546.

Wang, W., W. Cui, X. J. Wang, et al., 2016: Evaluation of GLDAS-1 and GLDAS-2 forcing data and Noah model simulations over China at the monthly scale. *J. Hydrometeor.*, **17**, 2815–2833, doi: 10.1175/JHM-D-15-0191.1.

Wei, H. L., Y. L. Xia, K. E. Mitchell, et al., 2013: Improvement of the Noah land surface model for warm season processes: Evaluation of water and energy flux simulation. *Hydrol. Process.*, **27**, 297–303, doi: 10.1002/hyp.9214.

Wei, S. G., Y. J. Dai, Q. Y. Duan, et al., 2014: A global soil data set for earth system modeling. *J. Adv. Model. Earth Syst.*, **6**, 249–263, doi: 10.1002/2013MS000293.

Wilson, K., A. Goldstein, E. Falge, et al., 2002: Energy balance closure at FLUXNET sites. *Agric. Forest Meteor.*, **113**, 223–243, doi: 10.1016/S0168-1923(02)00109-0.

Wood, E. F., J. K. Roundy, T. J. Try, et al., 2011: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resour. Res.*, **47**, W05301, doi: 10.1029/2010WR010090.

Xia, Y. L., A. J. Pitman, H. V. Gupta, et al., 2002: Calibrating a land surface model of varying complexity using multicriteria methods and the Cabauw dataset. *J. Hydrometeor.*, **3**, 181–194, doi: 10.1175/1525-7541(2002)003<0181:CALSMO>2.0.CO;2.

Xia, Y. L., K. E. Mitchell, M. B. Ek, et al., 2012a: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res. Atmos.*, **117**, D03109, doi: 10.1029/2011JD016048.

Xia, Y. L., K. E. Mitchell, M. B. Ek, et al., 2012b: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *J. Geophys. Res. Atmos.*, **117**, D03110, doi: 10.1029/2011JD016051.

Xia, Y. L., B. A. Cosgrove, M. B. Ek, et al., 2013a: Overview of the North American Land Data Assimilation System (NLDAS). *Land Surface Observation, Modeling and Data Assimilation*, S. L. Liang, X. Li, and X. H. Xie, Eds., World Scientific, Hackensack NJ, 337–377.

Xia, Y. L., M. B. Ek, J. Sheffield, et al., 2013b: Validation of Noah-simulated soil temperature in the North American Land Data Assimilation System phase 2. *J. Appl. Meteor. Climatol.*, **52**, 455–471, doi: 10.1175/JAMC-D-12-033.1.

Xia, Y. L., M. B. Ek, D. Mocko, et al., 2014a: Uncertainties, correlations, and optimal blends of drought indices from the NLDAS multiple land surface model ensemble. *J. Hydrometeor.*, **15**, 1636–1650, doi: 10.1175/JHM-D-13-058.1.

Xia, Y. L., M. B. Ek, C. D. Peters-Lidard, et al., 2014b: Application of USDM statistics in NLDAS-2: Optimal blended NLDAS drought index over the continental United States. *J. Geophys. Res. Atmos.*, **119**, 2947–2965, doi: 10.1002/2013JD020994.

Xia, Y. L., M. T. Hobbins, Q. Z. Mu, et al., 2015a: Evaluation of NLDAS-2 evapotranspiration against tower flux site observations. *Hydrol. Process.*, **29**, 1757–1771, doi: 10.1002/hyp.10299.

Xia, Y. L., M. B. Ek, Y. H. Wu, et al., 2015b: Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: Comparison and analysis. *J. Hydrometeor.*, **16**, 1962–1980, doi: 10.1175/JHM-D-14-0096.1.

Xia, Y. L., D. M. Mocko, M. Huang, et al., 2017: Comparison and assessment of three advanced land surface models in simulating terrestrial water storage components over the United States. *J. Hydrometeor.*, **18**, 625–649, doi: 10.1175/JHM-D-16-0112.1.

Xia, Y. L., D. M. Mocko, S. G. Wang, et al., 2018: Comprehensive evaluation of the variable infiltration capacity (VIC) model in the North American Land Data Assimilation System. *J. Hydrometeor.*, **17**, 1853–1879, doi: 10.1175/JHM-D-18-0139.1.

Xiao, J. F., J. Q. Chen, K. J. Davis, et al., 2012: Advances in upscaling of eddy covariance measurements of carbon and water fluxes. *J. Geophys. Res. Biogeo.*, **117**, G00J01, doi: 10.1029/2011JG001889.

Xie, Y., P. X. Wang, X. J. Bai, et al., 2017: Assimilation of the leaf area index and vegetation temperature condition index for winter wheat yield estimation using Landsat imagery and the CERES-Wheat model. *Agric. Forest Meteor.*, **246**, 194–206, doi: 10.1016/j.agrformet.2017.06.015.

Xu, T. R., S. L. Liang, and S. M. Liu, 2011: Estimating turbulent fluxes through assimilation of geostationary operational environmental satellites data using ensemble Kalman filter. *J. Geophys. Res. Atmos.*, **116**, D09109, doi: 10.1029/2010JD015150.

Xu, T. R., S. M. Liu, Z. W. Xu, et al., 2015: A dual-pass data assimilation scheme for estimating surface fluxes with FY3A-VIRR land surface temperature. *Sci. China Earth Sci.*, **58**, 211–230, doi: 10.1007/s11430-014-4964-7.

Xu, T. R., Z. X. Guo, S. M. Liu, et al., 2018: Evaluating different machine learning methods for upscaling evapotranspiration from flux towers to the regional scale. *J. Geophysics. Res. At-*

mos., **123**, 8674–8690, doi: 10.1029/2018JD028447.

Xu, T. R., X. L. He, S. M. Bateni, et al., 2019: Mapping regional turbulent heat fluxes via variational assimilation of land surface temperature data from polar orbiting satellites. *Remote Sens. Environ.*, **221**, 444–461, doi: 10.1016/j.rse.2018.11.023.

Yang, D. Q., B. E. Goodison, J. R. Metcalfe, et al., 1998: Accuracy of NWS 8" standard nonrecording precipitation gauge: Results and application of WMO intercomparison. *J. Atmos. Oceanic Technol.*, **15**, 54–68, doi: 10.1175/1520-0426(1998)015<0054:AONSNP>2.0.CO;2.

Yang, D. Q., D. Kane, Z. P. Zhang, et al., 2005: Bias corrections of long-term (1973–2004) daily precipitation data over the northern regions. *Geophys. Res. Lett.*, **32**, L19501, doi: 10.1029/2005GL024057.

Yang, F., H. Lu, K. Yang, et al., 2017: Evaluation of multiple forcing data sets for precipitation and shortwave radiation over major land areas of China. *Hydrol. Earth Syst. Sci.*, **21**, 5805–5821, doi: 10.5194/hess-21-5805-2017.

Yang, K., T. Watanabe, T. Koike, et al., 2007: Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget. *J. Meteor. Soc. Japan Ser. II*, **85**, 229–242.

Yang, K., T. Koike, I. Kaihotsu, et al., 2009: Validation of a dual-pass microwave land data assimilation system for estimating surface soil moisture in semiarid regions. *J. Hydrometeor.*, **10**, 780–793, doi: 10.1175/2008JHM1065.1.

Yang, K., L. Zhu, Y. Y. Chen, et al., 2016: Land surface model calibration through microwave data assimilation for improving soil moisture simulations. *J. Hydrol.*, **523**, 266–276, doi: 10.1016/j.jhydrol.2015.12.018.

Yang, R. Q., K. Mitchell, J. Meng, et al., 2011: Summer-season forecast experiments with the NCEP Climate Forecast System using different land models and different initial land states. *J. Climate*, **24**, 2319–2334, doi: 10.1175/2010JCLI3797.1.

Yilmaz, M. T., W. T. Crow, M. C. Anderson, et al., 2012: An objective methodology for merging satellite- and model-based soil moisture products. *Water Resour. Res.*, **48**, W11502, doi: 10.1029/2011WR011682.

Yu, Y. Y., D. Tarpley, J. L. Privette, et al., 2009: Developing algorithm for operational GOES-R land surface temperature product. *IEEE Trans. Geosci. Remote Sens.*, **47**, 936–951, doi: 10.1109/TGRS.2008.2006180.

Yuan, X., P. Ji, L. Y. Wang, et al., 2018: High-resolution land surface modeling of hydrological changes over the Sanjiangyuan region in the eastern Tibetan Plateau: 1. Model development and evaluation. *J. Adv. Model. Earth Syst.*, **10**, 2806–2828, doi: 10.1029/2018MS001412.

Zaitchik, B. F., M. Rodell, and F. Olivera, 2010: Evaluation of the Global Land Data Assimilation System using global river discharge data and a source-to-sink routing scheme. *Water Resour. Res.*, **46**, W06507, doi: 10.1029/2009WR007811.

Zhang, K., J. S. Kimball, and S. W. Running, 2016: A review of remote sensing based actual evapotranspiration estimation. *WIREs Water*, **3**, 834–853, doi: 10.1002/wat2.1168.

Zhang, T. P., P. W. Stackhouse, S. K. Gupta, et al., 2013: The validation of the GEWEX SRB surface shortwave flux data products using BSRN measurements: A systematic quality control, production and application approach. *J. Quant. Spectrosc. Radiat. Transfer*, **122**, 127–140, doi: 10.1016/j.jqsrt.2012.10.004.

Zhang, T. P., P. W. Stackhouse, J. S. Gupta, et al., 2015: The validation of the GEWEX SRB surface longwave flux data products using BSRN measurements. *J. Quant. Spectrosc. Radiat. Transfer*, **150**, 134–147, doi: 10.1016/j.jqsrt.2014.07.013.

Zheng, H., and Z. L. Yang, 2016: Effects of soil-type datasets on regional terrestrial water cycle simulations under different climatic regimes. *J. Geophys. Res. Atmos.*, **121**, 14,387–14,402, doi: 10.1002/2016JD025187.

Tech & Copy Editor: Lan YI