A global near-real-time soil moisture index monitor for food security using integrated SMOS and SMAP

Sara Sadri[a,*], Ming Pan[a], Yoshihide Wada[b], Noemi Vergopolan[a], Justin Sheffield[c], James S. Famiglietti[e], Yann Kerr[d], Eric Wood[a]

[a] *Princeton University, 59 Olden St, Princeton, NJ 08540, USA*
[b] *International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, Laxenburg A-2361, Austria*
[c] *University of Southampton, University Road, Southampton SO17 1BJ, UK*
[d] *Centre d'Etudes Spatiales de la BIOsphère (CESBIO), Toulouse 31400, France*
[e] *University of Saskatchewan, Global Institute for Water Security, SK, Canada*

ARTICLE INFO

ABSTRACT

Soil Moisture (SM) is a direct measure of agricultural drought. While there are several global SM indices, none of them directly use SM observations in a near-real-time capacity and as an operational tool. This paper presents a near-real-time global SM index monitor based on integrated SMAP (Soil Moisture Active Passive) and SMOS (Soil Moisture and Ocean Salinity) remote sensing data. We make use of the short period (2015–2018) of SMAP datasets in combination with two approaches—Cumulative Distribution Function Mapping (CDFM) and Bayesian conditional process—and integrate them with SMOS data in a way that SMOS data is consistent with SMAP. The integrated SMOS and SMAP (SMOS/SMAP) has an increased global revisit frequency and a period of record from 2010 to the present. A four-parameter Beta distribution was fitted to the SMOS/SMAP dataset for each calendar month of each grid cell at ~36 km resolution for the period from 2010 to 2018. We used an asymptotic method that guarantees the values of the bounding parameters of the Beta distribution will envelop both the smallest and largest observed values. The Kolmogorov-Smirnov (KS) test showed that more grids globally will pass if the integrated dataset is from the Bayesian conditional approach. A daily global SM index map is generated and posted online based on translating each grid's integrated SM value for that day to a corresponding probability percentile relevant to the particular calendar month from 2010 to 2018. For validation, we use the Canadian Prairies Ecozone (CPE). We compare the integrated SM with the SMAP core validation and RISMA sites from ISMN, compare our indices with other models (VIC, ESA's CCI SM v04.4 integrated satellite data, and SPI-1), and make a two-by-two comparison of candidate indices using heat maps and summary CDF statistics. Furthermore, we visually compare our global SM-based index maps with those produced by other organizations. Our Global SM Index Monitor (GSMIM) performed, in many tests, similarly to the CCI's product SM index but with the advantage of being a near-real-time tool, which has applications for identifying evolving drought for food security conditions, insurance, policymaking, and crop planning especially for the remote parts of the globe.

1. Introduction

Surface soil moisture (SM) is an Essential Climate Variable (Liu et al., 2011) because it is a culmination of many hydrological, energy, and carbon cycle processes representing variables such as precipitation (Koster et al., 2009; Berg et al., 2017), runoff (Wagner et al., 2003), infiltration (Wagner et al., 2013), temperature (Lintner et al., 2014), evapotranspiration (ET) (Seneviratne et al., 2010), and other water supply processes (Entekhabi et al., 1996). Reduced supplies of precipitation, and subsequently SM for crops, leads to agricultural drought that impacts crop yield, consequently inflicting enormous economic impacts in developed countries and suffering by millions of people in less developed regions of the world (Sheffield and Wood, 2007; Wilhite et al., 2007; EM-DAT, 2020). As the global population is increasing to reach an estimated 9 billion people by 2050, the demand for agricultural produce is putting pressure on already scarce agricultural

lands, especially in developed countries (UN/ISDR, 2007; FAO, 2009). The food production, including biofuels, required for this growing population is estimated to increase by 52% by 2050. Accordingly, to meet food security by 2050, the world's arable land will need to be expanded by 70 million hectares (net balance of 120 ha in the developing countries such as Africa and South America, and a contraction of 50 million hectares in developed countries) (FAO, 2009). Additionally, the expected overall increase of drought events under climate change (Sheffield and Wood, 2008b) will impact agriculture and water resources over these large areas and can have tremendous economic, social, and humanitarian consequences for food and water security (Wilhite, 2000; Sheffield et al., 2014).

Although the severity of such impacts are closely linked to the vulnerability of a population to adverse conditions (Haile, 2005; FAO, 2009) and how their political sectors respond within the constraints of economies (Hill and Pittman, 2013), monitoring SM and providing timely assessments of intensity, duration, and spatial extent of its anomalous conditions (Seneviratne and Nicholls, 2018) are essential for identifying evolving agricultural drought, flood monitoring, crop planning, food security, insurance, policymaking, and resilience-building especially in parts of the world where livelihoods and economy are closely intertwined with climate variability (Knutson, 2004; Sheffield and Wood, 2011; Nijssen et al., 2013; Al Bitar et al., 2017). Despite the fact that SM availability plays such important roles for agricultural drought monitoring and numerical weather predictions, unfortunately, it is among the least observed variables globally, especially in the developing world (Sheffield and Wood, 2011; Reichle, 2017).

Traditionally, drought indices are required to monitor drought conditions and are built upon long-term hydrological variables, mostly precipitation. Precipitation is the most spatially and temporally observed hydrologic variable (McKee et al., 1993). The Standardized Precipitation Index (SPI) (McKee et al., 1993) is a precipitation-based index recognized by the World Meteorological Organization (WMO) as the global index for meteorological drought. Another well-used index is the Palmer Drought Severity Index (PDSI) which is based on evapotranspiration and runoff (Palmer, 1965). Both SPI and PDSI are long-term drought indicators. Shorter-term indices such as CMI (Crop Moisture Index) exist but have limitations in assuming that parameters like land use/land cover and soil properties are uniform for all climatic regions (Narasimhan and Srinivasan, 2005). Other indices include the SM deficit index (SMDI) (Sánchez et al., 2016), the ET deficit index (ETDI) based on the SWAT model (Narasimhan and Srinivasan, 2005), the NLDAS-2 index based on four Land Surface Models (LSMs) of Variable Infiltration Capacity (VIC), Noah, Mosaic, and Sacramento (SAC) (Sheffield et al., 2014), and European Space Agency (ESA)'s ESSMI index (Carrão et al., 2016). All of these indices can be useful in support of decision-making systems but they are mostly regional tools based on estimated values from climatic variables, reanalysis, or hydrological modeling (Hunt et al., 2009; Sánchez et al., 2016). None of these indices directly use measured SM in near-real-time, in a global extent, and as an operational tool. Furthermore, they cannot be built toward drought prediction.

In recent years, a number of near-real-time drought monitoring systems have been developed based on such indices (Pozzi, 2013), but so far, the ability to monitor and predict the development and occurrence of extreme droughts at a global scale and in near-time has remained limited. One of the earliest attempts at developing a global drought monitor was by Kogan and Sullivan (1993) who used satellite information as the base of vegetation indices. Lloyd-Hughes and Saunders (2007) operated a global drought monitor from 2007 to 2013, which was updated on a monthly basis and used station-based precipitation from the Global Precipitation Climatology Centre (GPCC) (Schneider et al., 2014), air temperature from the European Centre for Medium-Range Weather Forecasts (ECMWF), SPI, and PDSI. AghaKouchak and Nakhjiri (2012) developed a global drought monitor

that uses near-real-time precipitation data from the Tropical Rainfall Measuring Mission (TRMM; (Huffman et al., 2007)) and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN; (Sorooshian et al., 2000)) to develop SPI indices. The output is then bias-corrected with historical precipitation from the GPCC (Adler et al., 2003) through the use of a Bayesian correction algorithm. Their efforts have led to the development of the Global Integrated Drought Monitoring and Prediction System (GIDMaPS) (Hao et al., 2014). The Global Drought Monitor Portal is based on the Standardized Precipitation-Evapotranspiration Index (SPEI). It is in near real-time, with a 1° spatial resolution and a monthly time resolution. It provides information about drought conditions on a global scale and is part of the Global Framework for Climate Services of the WMO and can act as a reference for global drought information from multiple sources (Heim and Brewer, 2012). Nijssen et al. (2013) described the implementation of a multimodel drought monitoring system or the Global Drought Information System (GDIS), which provides near-real-time estimates of surface moisture storage for the global land areas between 50°S and 50°N with a time lag of about 1 day at a spatial resolution of 0.5°. GDIS combines hydrological simulations from multiple LSMs (VIC, Noah, and Sacramento) to provide daily estimates of near-real-time SM conditions. Sheffield et al. (2014) developed an African Drought Monitor (AFDM) which uses VIC LSM whose accuracy comes from the availability of rain gages combined with temperature from the National Centers for Environmental Prediction's (NCEP) Global Forecast System (GFS) analysis fields. Drought conditions are reported as percentiles relative to a long-term climatology.

Since the 1980s, using microwave remote sensing data with spatially and temporally continuous operation over large areas, especially where ground measuring is not possible, has become an attractive option for drought monitoring (Sheffield et al., 2004; Bolten et al., 2010), flood forecasting (Wanders et al., 2014; Alvarez-Garreton et al., 2016), precipitation estimates (Wanders et al., 2015; Zhan et al., 2015), crop yield (Manzoni et al., 2013), and weather forecasting (Drusch et al., 2009). However, the use of satellite SM data for agricultural drought monitoring has remained limited. The reasons have been either that the short-term availability of satellite SM poses a statistical challenge for meaningful estimation of drought indices, or that the long-term data are not produced in near-real-time. For example, the ESA's ECV SM, better known as CCI (Climate Change Initiative) SM product, which merged 11 active and passive microwave satellite sensors, covers a 40-year period between November 1978 and June 2018 with a daily temporal resolution (European Space Agency, 2012). The CCI dataset is producing an updated SM product once or twice a year. In this situation, even if an index can be defined, it cannot be developed into an operational tool. For example, Carrão et al. (2016) developed an empirical standardized SM index for agricultural drought assessment based on CCI SM dataset product, but the index has not led to near-real-time applications. The Copernicus Climate Change Service (C3S) is an online service developed by ESA's CCI SM project as an R&D product to provide key indicators on climate change including the global SM gridded data from 1978 to present on a near-real-time (minimum 10 days and maximum 20 days after sensing) basis (C3S, 2020). Unlike the CCI data, the C3S data are not released as packages, however, they are still based on CCI and between 10 and 20 days behind real-time (de Jeu et al., 2018; Climate Change Service, 2018).

Currently, there are two orbiting satellite missions that measure the global surface SM (nominally top 0–5 cm) in a near-real-time capacity. One is NASA's Soil Moisture Active and Passive (SMAP) (Entekhabi et al., 2010; Chan et al., 2016) and the other is ESA's Soil Moisture and Ocean Salinity (SMOS) (Kerr et al., 2010, 2016). SMOS was launched in November 2009 (Pan et al., 2010; Kerr et al., 2016) while SMAP is a more recent mission launched in January 2015, mapping the globe gridded at about ~36 km resolution every 2–3 days.

Of particular interest to NASA's SMAP mission is a global drought index product that monitors SM in terms of probability percentiles for

dry and wet conditions in near-real-time (Bindlish et al., 2018; JPL, 2020). However, the short period of SMAP and its global revisit frequency (every 2–3 days) pose a statistical challenge for meaningful assessment of its indices. Sadri et al. (2018) developed a near-real-time soil moisture index monitor for the Contiguous U.S. (CONUS) using SMAP at near-surface (~36 km resolution) and an assimilated version of it at root zone (~9 km resolution) depths. To overcome the short-term of SMAP record, data was integrated with the simulated SM data from VIC LSM using the Bayesian conditional process. For the global scale, the ability of VIC to simulate SM is not sufficient globally (Sadri et al., 2018) and hence it cannot be used. Bindlish et al. (2018) demonstrated the use of reprocessed constant-angle SMOS brightness temperatures in the SMAP SM retrieval algorithm to develop a consistent multi-satellite product. This approach, in practice, has become a very time-consuming task since the brightness temperatures from the two missions are not consistent and have a bias of about 2.7 K over land with respect to each other. Furthermore, the two missions use different retrieval algorithms and ancillary datasets which result in further inconsistencies between the SM products. Therefore, no integrated SMOS and SMAP dataset has become available.

In this paper, we use statistical approaches to integrate SMAP and SMOS (Kerr et al., 2010, 2016) as one dataset (we refer to the integrated SMOS and SMAP SM dataset as SMOS/SMAP from this point on). The SMOS/SMAP is consistent with the SMAP data but, with an increased global revisit frequency (1-day) and a period of record from 2010 to the present that would be unachievable by either one of the satellites alone. We describe the implementation of a Global Soil Moisture Index Monitor (GSMIM) between 55°S and 80°N with a time lag of about 7–8 days with a spatial resolution of ~36 km. Similar to the AFDM, in the GSMIM drought and pluvial conditions are reported as percentiles relative to long-term climatology. The system is an extension and improvement of its prototype system developed by Sadri et al. (2018) for the Contiguous U.S. Major differences to achieve this global system include: we integrated SMOS with SMAP datasets using Cumulative Distribution Function Mapping (CDFM) and Bayesian conditional process approaches and compared the two approaches to understand the efficiency of each method in defining a global drought index; we used the 4-parameter Beta distribution with an asymptotic approach but for extending the tails of the distribution, which guarantees that the value of the two bounding parameters will envelop both the smallest and largest observed values; and we did validation by comparing our results with the in-situ SM data, other indices, and SPI-1.

A key element of this study is the development of a web interface to view the daily SM index maps overlaid as a dynamic layer of Google Maps, which provides coordinate lookup, panning, zooming, and overlaying other geographical layers such as land use, land covers, and political borders. Users can access the archive of previous days, and download the maps.

In Section 1, we looked at the importance of monitoring SM and having a global near-real-time agricultural drought index. The rest of this paper is organized as follows: Section 2 explains our datasets and the availability of passive and active microwave soil moisture retrievals. Section 3 explains our methodologies. Section 4 is about validating the results in the Canadian Prairies. In Section 5, we explain the results and discuss the findings. Section 6 provides online references to the data used in this study, and Section 7 provides discussion and ideas for future work.

2. Datasets

The resources where we obtained global SM measurements include in-situ measurements, simulations from LSMs, and several Earth Observations that use active and/or passive measurements (Kerr and Njoku, 1990).

2.1. SMAP and SMOS

The SMAP mission, launched by NASA in January 2015, maps the globe (between 85.044°N and S) using a 1.4 GHz L-band microwave radiometer and delivers brightness temperature (TB) observations at a fixed 40 incidence angle (Entekhabi et al., 2010). The microwave sensors are categorized into passive and active sensors. The passive sensor, i.e. radiometer, measures the naturally emitted energy from the earth's surface in the form of TB. The active sensor, i.e. radar, sends a focused beam of microwave radiation toward the ground and captures the backscattered signal from the earth's surface (Karthikeyan et al., 2017a). The radar on SMAP failed after 3 months from its launch, but SM estimates based on a single-channel algorithm (SCA) from a single-orbit passive radiometer continue to be produced (Sadri et al., 2018). SMAP Level 1 delivers products from its instrument measurements and Level 2 delivers SMAP's Swath based geographical retrievals. SMAP Level 3 (SPL3SMP) is a composite based on daily passive radiometer estimates of global land SM at the top 5 cm of the soil that is resampled to a global, cylindrical ~36 km Equal-Area Scalable Earth Grid, Version 2.0 (EASE-Grid 2.0) (O'Neill et al., 2016). LSMs assimilating SMAP measurements at the root zone depth (up to 1 m) are delivered as SMAP Level 4 (Reichle, 2017; Colliander et al., 2017). In all release versions of SMAP, regions with permanent snow, ice, frozen, and thawed ground are masked out using a passive freeze-thaw retrieval mask. Additionally, areas with Radio Frequency Interference (RFI), open water, permanent ice, and large Vegetation Water Content (VWC) are also marked out. The SMAP satellite is in a 6 AM/6 PM ascending and descending orbit. The primary objective of SMAP is to retrieve SM from the morning overpasses. For this study, we use version 4 of SPL3SMP retrievals from the morning overpasses that are used (hereafter referred to as SMAP).

The SMOS L-band radiometer was launched in November 2009 (Pan et al., 2010; Kerr et al., 2016) by the Centre Aval de Traitement des Donnã©es SMOS. Its retrievals are based on a Multi Orbit retrieval algorithm. The satellite measures the thermal emission of SM such as the SM EASE grid Level 3 (L3SM) from the Earth at a frequency of 1.4 GHz and for incidence angles from 0° to ~60° sampled at ~25 km resolution, with a revisit time of every 2–3 days (similar to SMAP) and a mission target accuracy of 0.04 (m^3/m^3) (Kerr et al., 2010). Detection of ice and snow is performed and added as flags. For SMOS L3SM both ascending (~6 am local time) and descending (~6 pm local time) orbits are processed separately (Al Bitar et al., 2017; Centre Aval de Traitement des Données SMOS, 2017) on a daily basis. Here, we use only the ascending orbits of SMOS L3SM (hereafter referred to as SMOS), in order to have consistency with the SPL3SMP data as well as smaller uncertainties and bias from the in-situ data (Al Bitar et al., 2017).

2.2. ESA-CCI merged SM

Developed by the Vienna University of Technology (TU Wien), Dutch company VanderSat BV, and the Austrian Earth Observation Data Centre and hosted by ESA, the ESA-CCI merged SM is a quarter degree resolution global scale daily long-term SM record (provided in volumetric units m^3/m^3). Version 4.4 of the CCI product covers a 40 year period from 1978 to 2018 and is based on blending data from 11 active and passive microwave satellite sensors through the method described in (Liu et al., 2011; Wagner et al., 2013). The active data set was generated based on observations from the C-band scatterometers on board of ERS-1-2 SCAT, MetOp-A ASCAT, MetOp-B and for the period of 1991/08/05 to 2018/06/29. The passive data set was generated based on passive microwave observations from Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR), Defense Meteorological Satellite Program (DMSP) F8-F13 Special Sensor Microwave - Imager (SSM/I), the microwave imager on-board Tropical Rainfall Measuring Mission (TRMM) TMI, Aqua's Advanced Microwave Scanning

Table 1
Datasets and their calibration/validation periods used in this study.

Dataset	Source	Depth (cm)	Period	Gridded res. (km)	Temporal. res.	Cosidered period for validation
SMAP (SPL3S MP)	RS	5	2015/03/31-2018/11/12	36	Every 3–4 days	–
SMOS (L3SM)	RS	5	2010/01/15-2018/11/12	25	Every 7–8 days	–
SMOS/SMAP	Integrated	5	2010/01/15-2018/11/12	36	Daily	2010/01/15-2018/06/29
CCI (V04.4)	Integrated	0.5–2	1978/01/01-2018/06/29	25	Daily	2010/01/15-2018/06/29
VIC	LSM	5–10	1978/01/01-2018/11/12	36	Daily	2010/01/15-2018/06/29
SPI-1	MSWEP	–	1978/01/01-2018/11/12	36	Monthly	1979/01/01-2018/06/29

* RS: Remote Sensing
** MSWEP: Multi-Source Weighted-Ensemble Precipitation

Radiometer – Earth Observing System (AMSR-E), GCOM the Advanced Microwave Scanning Radiometer 2 (AMSR2), Coriolis Windsat, and SMOS for the period of 1978/11/01 to 2018/06/29 (Chung et al., 2019). SMOS has been integrated into Version 4.04 but it is a version of SMOS estimated with AMSR-E/Aqua surface soil moisture (LPRM, V5). The way SMOS was integrated was to use a grid per grid change detection approach (scaled by either the ground site or model outputs and then spatially smoothed) to derive ASCAT SM maps on which AMSR LPRM data have been scaled to produce CCI. SMAP, Aquarius, and MetOp-C missions are not considered in the v04.4 and even in the current release (v04.5) of the CCI SM product. The harmonization of CCI datasets aimed to take advantage of both types of microwave techniques but proved difficult owing to sensor degradation, drifts in calibration and algorithmic changes in the processing systems. In fact, one of the challenges in using the ESA's CCI product is guaranteeing consistency between the SM data retrieved from the different active and passive microwave instruments. In addition, it is noteworthy that CCI's product is not a near-real-time product. Albergel et al. (2013) considered in-situ measurements from 2007 to 2010 from 196 stations of five networks in different countries (USA, Spain, France, China, and Australia) to evaluate the CCI product. Results indicate that in general, the product captures the temporal dynamics of observed SM well, with correlations of 0.6 (Â ± 0.06) at the 5% significance level. The merged product is freely available after completion of a simple registration (European Space Agency, 2012). For this study, we used the global gridded CCI version 4.4 (that goes up to June 29, 2018) product (Liu et al., 2011) with a spatial resolution of ~25 km.

2.3. In-situ dataset

The performance of active and passive soil moisture products is assessed by comparing them with station data. The in-situ candidate sites used in this study are from the International Soil Moisture Network (ISMN) (Dorigo et al., 2011a, 2012; Ojo et al., 2015; TUWIEN, 2019). Initiated by the SMOS project, ISMN plays a crucial role in globally assessing the quality of SM estimates from space-borne microwave sensors and LSMs. ISMN is the first, fully operational, and publicly available international initiative to globally host in-situ measurements of SM (Karthikeyan et al., 2017b) from more than 500 stations spanning 18 different networks (Dorigo et al., 2011b; TUWIEN, 2019). The main limitation of the ISMN is that it does not cover the globe homogeneously. Some areas (boreal, or tropical for instance) are almost not covered while others are densely covered (US, Europe) (Karthikeyan et al., 2017b).

2.4. VIC-SM

Large variations in in-situ data and its scarcity have encouraged the use of available global meteorological observations, such as precipitation and temperature, to simulate SM using LSMs such as VIC (Liang et al., 1994, 1996). Sheffield et al. (2004) used VIC to simulate retrospective SM for the contiguous U.S. using subgrid variability in land cover, elevation, soil water storage capacity, and storm coverage. The VIC model divides the soil column into three layers, with a fixed 10 cm top layer depth and a spatially varying depth for the lower layers (Sheffield and Wood, 2008a). The downside of using simulated SM data is that such data are prone to large uncertainties since their forcings and land-atmosphere feedback mechanisms are not well understood for many regions of the world (Sheffield et al., 2004; Dorigo et al., 2015).

2.5. SPI

Since SM responds to precipitation anomalies on a relatively short scale (Sánchez et al., 2016), we compared the SM indices with the SPI at the selected locations to assess whether they provide spatiotemporal consistency. The one-month SPI (SPI-1) is generally related to short-term SM and crop stress, especially during the growing season and therefore, can approximate conditions represented by the SMOS/SMAP drought index. SPI-1 means that at a given month, we compare the total precipitation for that month in a particular year, with total precipitation of that month in all the years of the time record. In this study, the precipitation data are from the Multi-Source Weighted Ensemble Precipitation (MSWEP) Version 2 (Beck et al., 2019) and re-gridded to ~36 km resolution. A drought event occurs any time the SPI-1 is continuously negative, reaches an intensity of -1.0 or less, and ends when the SPI becomes positive.

Table 1 summarizes the type of data and the periods used for calibration and validation in this study.

2.6. Soil moisture climatology on a global scale

Fig. 1 shows the comparison among mean interannual SM from VIC, CCI, SMOS, and SMOS/SMAP. For the CCI SM map, areas of a dense canopy in the mid-latitude have been masked out at the source. The mean SM values of all the groups differ by several volumetric percent. VIC LSM's physical interpretation of the three layers and possession of a smaller dynamic range than remotely sensed data cause a systematic difference against others, showing it to be wetter in general. The spatial patterns of CCI and SMOS maps are highly consistent although SMOS has a higher intensity in dryness. SMOS and SMOS/SMAP agree on dry parts of the globe but disagree on the wetness pattern. The SMOS/SMAP climatology is highly similar to that from SMAP (not shown here). The difference between SMOS/SMAP and SMOS is mostly in the type of configuration algorithms that SMOS and SMAP use for the computation of the Faraday rotation (Al Bitar et al., 2017). In all cases the Sahara Desert, the Arabian Peninsula, central and eastern Iran, and
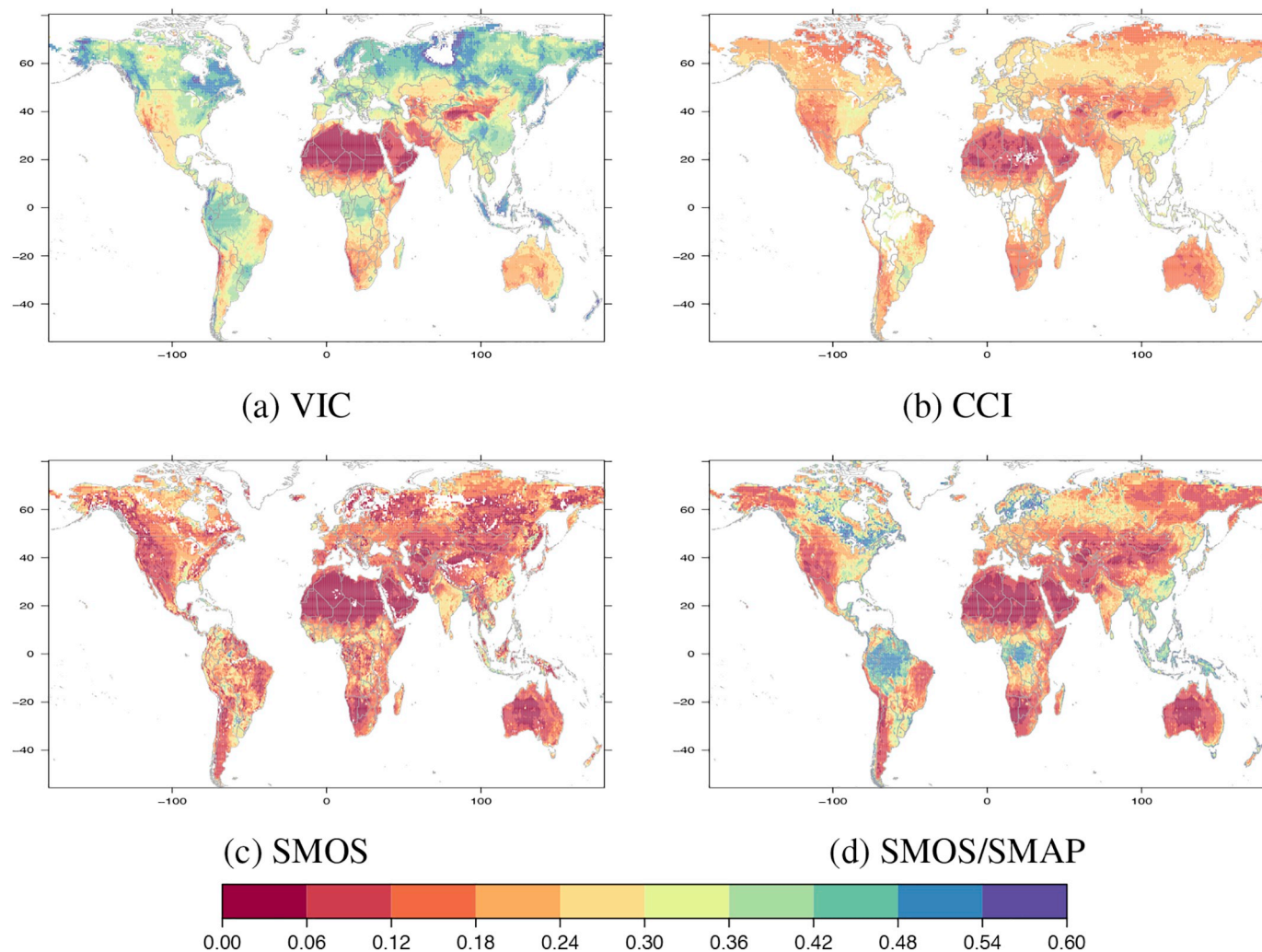
(a) VIC

(b) CCI

(c) SMOS

(d) SMOS/SMAP

0.00   0.06   0.12   0.18   0.24   0.30   0.36   0.42   0.48   0.54   0.60

Fig. 1. Average SM (m^3/m^3) for VIC, CCI, SMOS, and SMOS/SMAP for the period of 2010/01/15 to 2018/06/29.

northwestern China agree in dryness. However, SMOS/SMAP and SMOS show the average dryness in many other parts of the world as being noticeably drier than those from CCI and VIC. These areas include central Australia, Namibia and Botswana, eastern and along the western coast of South America, central/western U.S., and northwest Alaska. The overall wetness patterns among the SM sources disagree more noticeably, especially in the northern latitudes which include Canada and Alaska.

Fig. 2 shows the standard deviation and the areas with the highest SM variability for the four groups of VIC, CCI, SMOS, and SMOS/SMAP. Regional biases are strong, especially in SMOS, with a spatial standard deviation of 0.07 m^3/m^3 for the difference between the time-average fields. In the U.S., VIC shows the strongest deviation in the northwest. Among the remotely sensed data, the differences in standard deviation can be related to the differences in calibration of the sensors and acquisition time (Al Bitar et al., 2017).

3. Methodology

In this section, we present the computation process of the GSMIM. The process has three major steps:

1. For each grid g at ~36 km resolution and each calendar month m, we integrate $SMAP^{g, \, m}$ (2015–2018) and $SMOS^{g, \, m}$ (2010–2018) using two processes: Bayesian conditional and CDFM;
2. For each grid g and each calendar month m, we calculate the

parameters of the Beta distribution of the $SMOS/SMAP^{g, \, m}$ data (therefore every grid has parameters associated with each of the 12 calendar months); and
3. For near-real-time (NRT) application, as the incoming $SMAP^{g, \, d}$ and $SMOS^{g, \, d}$ come in every day d, we integrate them, and then a percentile is assigned for each of the 100,500 grids on the earth from that grid's 4 parameters of the Beta distribution called from the relevant calendar month in the calibration period.

Fig. 3 shows a summary flowchart of the steps taken in order to develop the GSMIM.

3.1. Integrating SMAP and SMOS

Satellite remote sensing data are often spatially and temporally discontinuous (Nijssen et al., 2001). SMAP provides data in near-real time for operational applications but its 5-year scattered data makes it technically difficult to meaningfully define an index. This raises the challenge of creating a longer-term high temporal resolution global dataset with remote sensing data provided by sensors on different satellites (Coccia et al., 2015). We use the global data from SMOS, upscaled from 25 km to 36 km resolution, from the period of 2010/01/15 to 2018/06/29 but since the characteristics of SMAP and SMOS sensors and their viewing geometry used to observe SM are slightly different, their retrievals also generate specific differences. To make a consistent dataset of these two satellites, we use the Bayesian conditional process
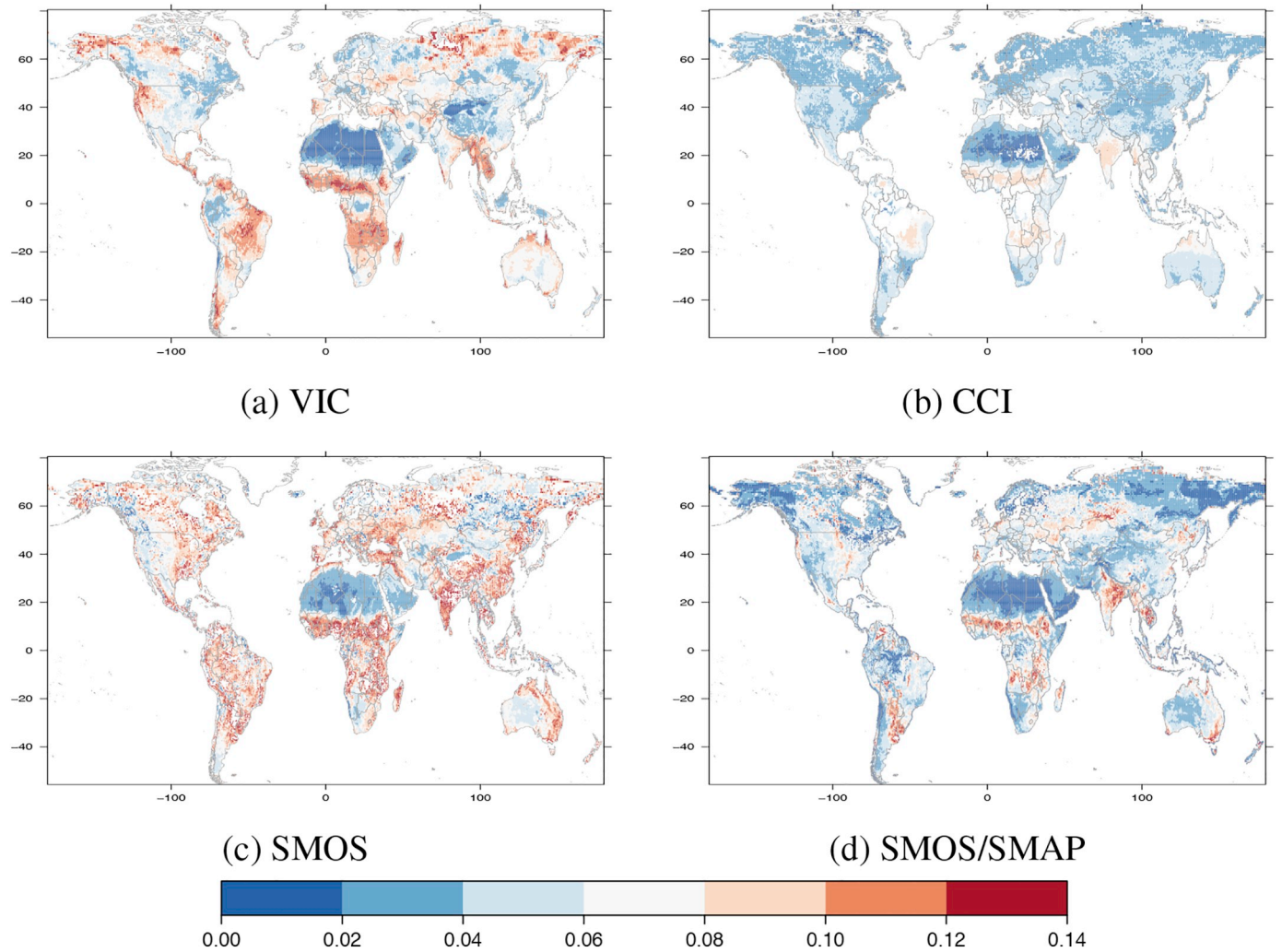
Fig. 2. Standard deviation of SM (m^3/m^3) for VIC, CCI, SMOS, and SMOS/SMAP for the period of 2010/01/15 to 2018/06/29. Areas of dense canopy in the mid-latitude have been masked out from CCI data.
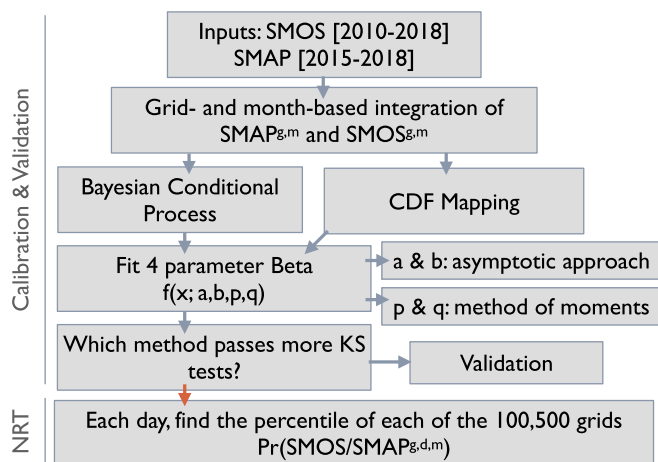

Fig. 3. Flowchart describing steps involved in developing a global NRT drought index monitor.

and CDFM. In both approaches, we aim to create a continuous SMOS/SMAP dataset consistent with SMAP, using their 2015–2018 overlapping period. In other words, we treat the SMAP data as the predictand and the SMOS data as the predictor. The Bayesian conditional process approach has been extensively explained by Coccia et al. (2015)

and Sadri et al. (2018) and will result in an extended-dataset consistent with sparse (short record length) remotely sensed data as well as provide in-filling of data due to the orbit gaps. CDFM has been a strong and standard way of rescaling and matching two or more datasets on the basis of their common periods (Liu et al., 2012). Both of these approaches are transferable to other areas of hydrology.

3.1.1. Bayesian conditional process overview

We follow the Bayesian conditional process from Coccia et al. (2015); Sadri et al. (2018), which operates as follows. It is based off the assumption that, if we normalize both SMOS and SMAP in an order-preserving way to be unit normal distributions, then their joint distribution is a bivariate normal, determined uniquely by the normalized variables' correlation. This principle is applied separately to each grid cell and each calendar month. Then, it gives us a way to randomly sample a missing value from one sequence, given the corresponding value from the other sequence.

The grid g- and month m-based process begins with calculating the empirical distribution of the values of $SMOS^{g,\,m}$ and $SMAP^{g,\,m}$, for the days that they both had an observation. We apply a quantile-preserving map of these values to unit normal functions, naming the new values $\eta_{SMOS}{}^{g,\,m}$ and $\eta_{SMAP}{}^{g,\,m}$.

For a number of random grids, we plotted the CDF of each month for each grid to find the minimum number of common days required to balance between having a reasonably-spread CDF while letting most

grids from most months participate in the analysis. For the CPE, for example, the maximum number of common days was 60. The CDF showed that if we choose 10 days or more the non-exceedance probability would be at least 30% meaning that we can use 70% of all the data. Therefore, we deemed at least 10 days of common observations required to proceed with that grid for that month.

We compute the correlation $\rho^{g, m}$ of the normalized variables $\eta_{SMOS}^{g, m}$ and $\eta_{SMAP}^{g, m}$. For any day d such that $SMAP^{g, d}$ is missing, but $SMOS^{g, d}$ is not, we find the realization of $SMOS^{g, d}$ transformed to the standard normal space. Our assumption is that the normalized variables form a Gaussian Bivariate Distribution with unit normal marginals and correlation $\rho^{g, m}$, so the distribution $f(\eta_{SMAP}^{g, d}|\eta_{SMOS}^{g, d})$ of $\eta_{SMAP}^{g, d}$ conditioned on this $\eta_{SMOS}^{g, d}$ is a normal distribution with mean and variance

$$\mu = \rho^{g,m}.(\eta_{SMOS}^{g,d}) \tag{1}$$

$$\sigma^2 = 1 - (\rho^{g,m})^2 \tag{2}$$

Here $\eta_{SMOS}^{g, d}$ is the transformed value of $SMOS^{g, d}$ in the standard normal space. Next, in the stochastic step, we draw a number from this distribution, which will be our missing $\eta_{SMAP}^{g, d}$. We map this value back to real space using the inverse quantile-preserving map for g and m. This new number replaces the missing $SMAP^{g, d}$.

3.1.2. CDFM

The CDFM approach is simpler and deterministic. It also can be viewed as a variant of the Bayesian approach where we assume perfect correlation (i.e. $\rho = 1$) for all grids and months. For days with a $SMOS^{g, d}$ record but no $SMAP^{g, d}$ observation, the closest value to $SMOS^{g, d}$ from the $SMOS^{g, m}$ calibration period is selected. We compute its exceedance probability and then apply the inverse CDF for $SMAP^{g, m}$; we define this $SMAP^{g, d}$ as the missing SMAP for that day.

3.2. Fitting a four-parameter Beta distribution to the integrated datasets

Application of the Beta distribution for capturing SM variability was suggested by (Famiglietti et al., 1999) and first tested by (Ryu and Famiglietti, 2005; Famiglietti et al., 2008). The Beta distribution is a family of continuous probability distributions, defined on the interval of [0,1], and parameterized by two positive shape parameters p and q. The probability density function (PDF) of the standard Beta distribution, a random variable x between [0,1] with positive shape parameters p and q, is as follows:

$$f(x; p, q) = \frac{(x)^{(p-1)}(1 - x)^{(q-1)}}{B(p,q)} \quad 0 \le x \le 1; p, q > 0 \tag{3}$$

$B(p,q)$ is a normalizing constant known as the Beta function, related to the Gamma function:

$$B(p,q) = \int_0^1 u^{(p-1)}(1 - u)^{(q-1)}du \tag{4}$$

Here u is a parameter introduced for the sake of integration.

The generalized form of Beta distribution is transformed linearly to have lower and upper bounds a, b instead of 0 and 1, with probability density function:

$$f(x; a, b, p, q) = f\left(\frac{x - a}{b - a}; p, q\right)/(b - a)$$
$$= \frac{(x - a)^{(p-1)}(b - x)^{(q-1)}}{B(p,q)(b - a)^{(p+q-1)}} \quad a \le x \le b; p, q > 0 \tag{5}$$

The Beta distribution has the capability of mapping the behavior of SM as a random variable with limited variability range (Sheffield et al., 2014; Sadri et al., 2018), however, the fact that the SMOS/SMAP calibration data spans only 8 years can be still a statistical challenge for meaningful assessment of its indices. Even with long enough datasets, we still want to extend the tails of the distribution beyond the observed values for those extreme drought and flood events. To the authors'

knowledge, no existing methodology works for this case, although there have been efforts to establish such procedures in the past. For example, Sadri et al. (2018) extended the lower and upper bounds of the integrated SMAP/VIC data proportional to the difference of long-term (1979–2017) with short-term (2015–2017) of VIC data for each grid. Sheffield et al. (2004) used the first and last 10% of the sorted soil moisture values to linearly extrapolate the cumulative distribution functions. This approach does not consistently extend the tails of the data and sometimes it even gives bounding parameters less than the already observed values.

3.2.1. Finding a and b using an asymptotic model endpoints

Since there has been a lack of scientific exploration on extending the tails of the Beta distribution, and to better support our needs, we introduce a new and robust approach that uses asymptotic model endpoints for finding a and b parameters. We calculate the limit of the CDF as x gets closer to a,

$$\frac{d}{dx}(\lim_{x \to a}(F(x, a, b, p, q))) = \lim_{x \to a} f(x, a, b, p, q) \tag{6}$$

and by definition $\lim_{x \to a} f(x, a, b, p, q) = (x - a)^{p-1}. k_0$ where $k_0 = \frac{(b-a)^{q-1}}{B(p,q)(b-a)^{(p+q-1)}} = \frac{1}{B(p,q)(b-a)^p}$. So integrating, the CDF of the Beta distribution in the limit as x approahces a becomes:

$$\lim_{x \to a} F(x, a, b, p, q) = \int_a^b (u - a)^{(p-1)}. k_0 du = (x - a)^p. k \tag{7}$$

where $k = k_0/p_0$. The problem then becomes to fit the limiting CDF $(x - a)^p. k$ to the empirical CDF.

Let n be the total length of SMOS/SMAP values in grid g in month m, and $\hat{x}_i$ denote the observed data point of rank i. To mimic the behavior of the limit, we consider only the lowest decile of the observed CDF. So the empirical CDF is $(\hat{x}_i, i/n)$ for $i : [1,2,3,\ldots N/10]$. Taking logs, we equivalently want to fit:

$$\log k + p. \log(\hat{x}_i - a) \approx \log(i/n) \tag{8}$$

We formalize the objective as finding a, such that the resulting line of best fit minimizes the sum of squared vertical residuals to

$$\begin{pmatrix} \text{X axis: } \log(\hat{x}_i - a) \\ \text{Y axis: } \log(i/n) \end{pmatrix} \tag{9}$$

Similarly for finding b, the objective is to choose b yielding the line of best fit to:

$$\begin{pmatrix} \text{X axis: } \log(b - \hat{x}_{(n-j)}) \\ \text{Y axis: } \log((n - j)/n) \end{pmatrix} \tag{10}$$

where j is defined as $j : [1,2,3,\ldots N/10]$, the highest decile of observed data points.

This can be implemented by trying all possible values of a and b in a reasonable range by brute force (SMAP data are $\in[0.02, 0.99]$), then using a standard linear regression computation for each one. The benefits of this method are:

- it guarantees that the value of a will be smaller than the smallest soil moisture value, and similarly for b
- it removes the complexity to past approaches—we do not try to solve for more than one Beta parameter at a time.

We discard any other information from the optimal solutions other than the values of a, b (i.e. we don't try to infer approximate information about p, q; See the next section). We also tried alternative definitions of the line of best fit:

- We calculated the line of best fit in terms of squared horizontal residuals but the values of a (and b) tend to be unrealistically very close to $a = 0$ and $b = 1$, which were not close enough to the actual

(a) CDF   (b) PDF
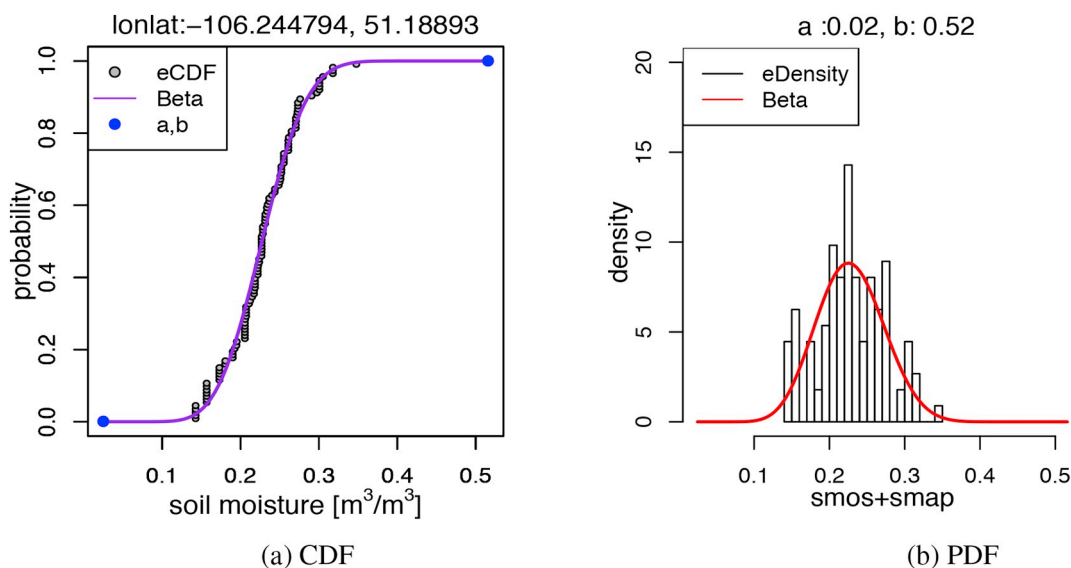
Fig. 4. Example from fitting 4-parameter Beta to months of April between the period of 2010–2018 at a grid point in SK, Canada (lon:-106.2448, lat:51.18893). (left) The a and b values are shown on the CDF by blue dots are the lower and upper bounds outside of the observed values. (right) The density plot is consistent with the CDF plot on extended beyond the observed values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

observed soil moisture range for most of the candidate grids we tested.

- We also tried picking a, b to maximize the correlation. The results were very close to the line of best fit in the first scenario.

It is important to note that here, we want a model that describes the two vectors in the best linear relationship it can get, and therefore, tests such as NSE, KGE, Bias or RMSE are not as useful as their purpose is to measure exact correspondence.

3.2.2. Finding p, q using method of moments

We applied the method of moments (1st and 2nd) to estimate p and q after we transformed: $X' = (X - a)/(b - a)$. Note that when estimating a and b, due to the asymptotic analysis we looked at only the first and last 10 percentiles of the distribution. But for estimating p and q we look at the entire data.

Fig. 4 shows an example of CDF and PDF plots before and after fitting a 4-parameter Beta distribution at a grid point in Saskatchewan for all Aprils from 2010-to 2018. The observed data (black circles in the left plot) are in the range of 0.13 to 0.36 m^3/m^3. However, using the asymptotic model, the endpoints of a and b (blue dots) extend the tails of the distribution to a range of 0.02 to 0.52 m^3/m^3. The longer tailed distribution can take care of those unseen drought and flood events. This process was done for every calendar month of every grid point.

We used the Kolmogorov-Smirnov (KS) test to examine the goodness-of-fit of Beta distribution at the 95% significance. The null hypothesis of the KS test is that the data follow the specified distribution. The null hypothesis regarding the distributional form was rejected if the test statistic, D, was greater than the critical value, $D_{crit.}$, obtained from the KS table. Otherwise, there was no evidence to support the rejection of the null hypothesis, and therefore we accepted that the data follow the Beta distribution.

3.3. Finding the percentile as the new integrated SMAP and SMOS come in every 24 h

For NRT application, as the $SMAP^{g, d}$ and $SMOS^{g, d}$ come in every day d, we integrate them as $SMOS/SMAP^{g, d, m}$, and then we find the corresponding percentile of the integrated $SMOS/SMAP^{g, d, m}$ using the grid's Beta distribution parameters repository from the month m of that

day d belongs to. The global grids' percentiles are then smoothed up (each grid is the average of neighbor grids) and transformed into a colored map similar to the U.S. Drought Monitor (USDM), which uses levels D0-D4 to indicate drought severity. In addition, we extend our indices to pluvial conditions similar to the maps from the Gravity Recovery and Climate Experiment (GRACE). We have developed an online platform that overlays these maps over Google Maps and stores the archive for download (i.e. GSMIM). We programmed the GSMIM to run on a daily schedule.

4. Validation strategies

Ensuring the accuracy of integrated data originating from space is a crucial aspect of any satellite data processing effort (Sánchez et al., 2016). One way of achieving this is to compare the in-situ measurements on the ground with the integrated remote sensing data (Rodríguez-Fernández et al., 2017). Even though ground-based measurements are sparse and not necessarily representative of large-scale satellite soil moisture (Molero et al., 2018), such evaluation is important to understand the variations and diversity of conditions between the point measurements and the SMOS/SMAP data.

4.1. Canadian Prairies Ecozone (CPE)

The Canadian Prairies (or grassland) are the three provinces of Alberta, Saskatchewan, and Manitoba. The southern parts of these provinces are called the CP Ecozone (CPE), an area of 520,000 km^2, shown in Fig. 5, an area with relatively homogeneous soil texture. CP has about 50% of Canada's number of farms and more than 80% of the farmland (in terms of hectares) that are mostly concentrated in the CPE. The CPE is the core of the agricultural economy of Canada with some of the world's highest climate and weather variability. Such variabilities have significant effects on CPE's agriculture, environment, economy, and culture on a year to year basis since most of the agriculture is still rain-fed. For example, the drought of 2001–2002 cost approximately $3.6 billion in agricultural production losses (Wheaton et al., 2005). Between 2008 and 2012, federal-provincial disaster relief payouts for climate-related events totaled more than $785 million, and more than $16.7 billion in crop insurance was paid out. The 100-year record-breaking drought in 2017 caused massive wildfires (also in British
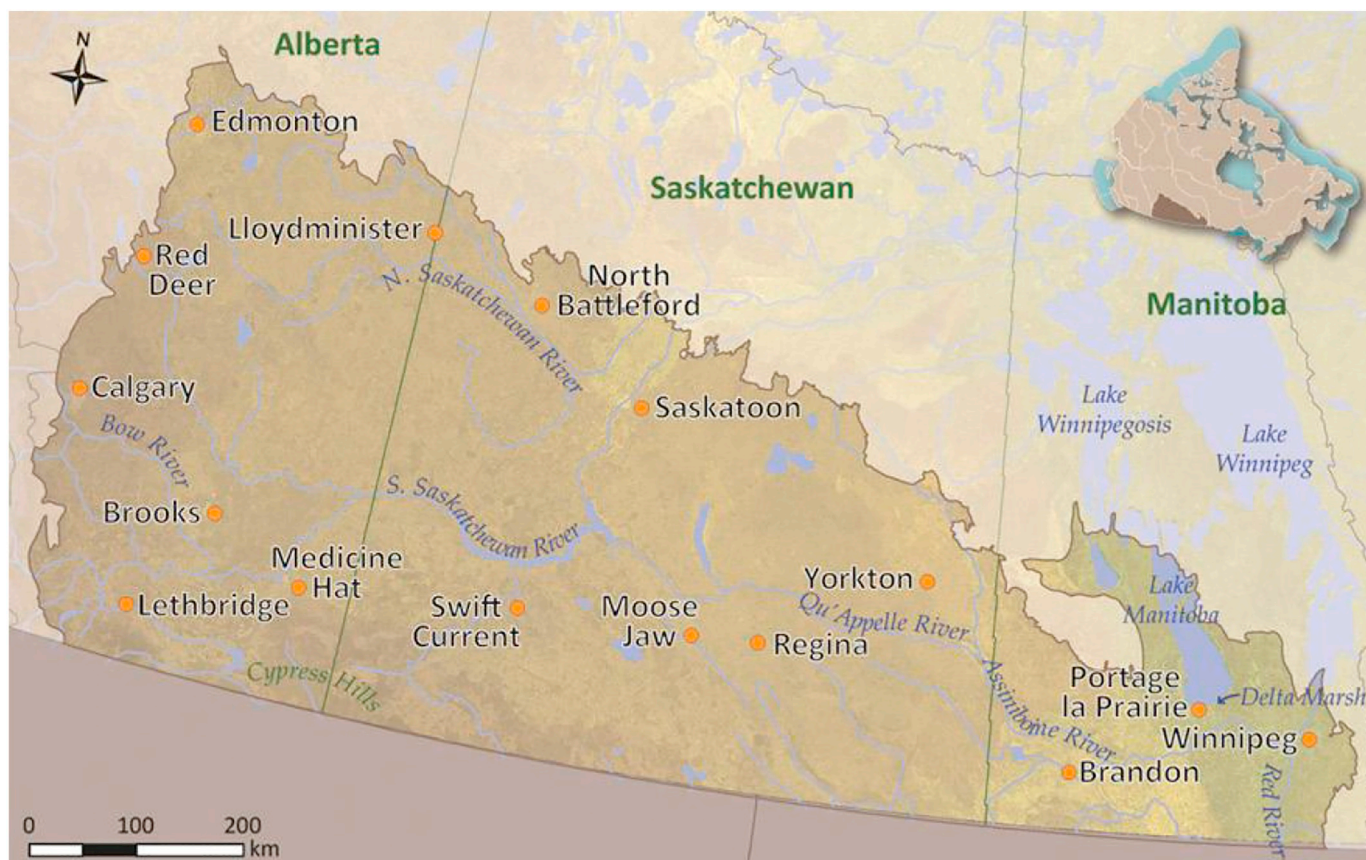
Fig. 5. CPE position relative to Canada and the CP region (three provinces of Manitoba, Saskatchewan, and Alberta). Source: (ESTR Secretariat, 2014).

Columbia), reduced yields (particularly canola), heat stress, poor grain fill, livestock feed shortages, and relocation of nearly 3000 cattle in Saskatchewan and Alberta (Cherneski, 2018). Given the vulnerability of the CPE to drought and agricultural production risks and the future scenarios of climate, which show more severe and frequent droughts with declining trends of precipitation and surface water resources during summer and fall, there is a need to develop drought outlook and forecasting for the CPE to better inform decisions by producers, especially at the start of the growing season (Sauchyn and Kulshreshtha, 2008; Cherneski, 2018).

High seasonal and interannual variability in rainfall are shown in Figs. 6 and 7. The mean annual precipitation in the CPE from 1979 to 2018 ranges from 130 mm to 860 mm. The average spring and fall rainfall, the seasons of the great demand for agriculture, is low. Summer has the highest average precipitation but with great interannual variability (Fig. 7) and a declining trend on some future projections scenarios. In general, the eastern and western margins of the area receive slightly higher precipitation but the major part of the CPE is suffering from the narrow window of favorable agricultural weather for much of the year.

4.2. Regional comparison of SMOS/SMAP soil moisture with in-situ data

In the CPE, the in-situ data of ISMN are collected from SMAP core validation sites and RISMA, a Canadian monitoring network of in-situ SM established by Agriculture and Agri-Food Canada (AAFC) and Environment and Climate Change Canada (ECCC). RISMA has three in-situ monitoring networks near Kenaston (Saskatchewan) (Tetlock et al., 2019), Carman (Manitoba) (Bhuiyan et al., 2018), and Casselman (Ontario) as part of the Sustainable Agriculture Environmental Systems (SAGES) project entitled Earth Observation Information on Crops and Soils for Agri-Environmental Monitoring in Canada (Government of

Canada, 2017). A total of 111 in-situ sites were gathered in combination which many of them were geographically very close to each other or closely overlapped. Both SMAP core validation sites and RISMA are chosen based on a requirement of having continuous soil moisture measurements at 5 cm depth with replication within our grid cells with a spatial scale of ~36 km. Fig. 8 shows the locations of these 111 sensors in the CPE region with overlapping transparent orange circles. The centers of the ~36 km grids containing these sites are numbered. Four of these grids are in Saskatchewan (SK) and 4 are in Manitoba (MB). Grids 1 to 6 have a higher number of sensors and therefore, higher confidence in the representativeness of data. Grids 7 and 8 can be used in algorithm testing but the confidence of representativeness is not high enough due to the low number of sensors within the grids. Furthermore, land use land cover studies show that both of these grids are located in urban and developed areas which might contribute to their discrepancy between the in-situ and remotely sensed data. Therefore, the in-situ data from these two grids are eliminated from the study.

Our study period for point validation spans from January 2012 to December 2017. According to the Canadian Drought Monitor (CDM), there were at least two periods of drought in the prairies during this period. One severe drought occurred in 2012 and the other occurred in 2017. In contrast, the year 2016 was a very wet year with above-normal precipitation during summer and fall. In the next section, we look for these events in our results.

5. Results and discussions

5.1. CDFM vs. Bayesian conditional process

Fig. 9 compares correlation and RMSE from the two processes of integrating SMOS/SMAP SM data. The SMOS/SMAP dataset obtained
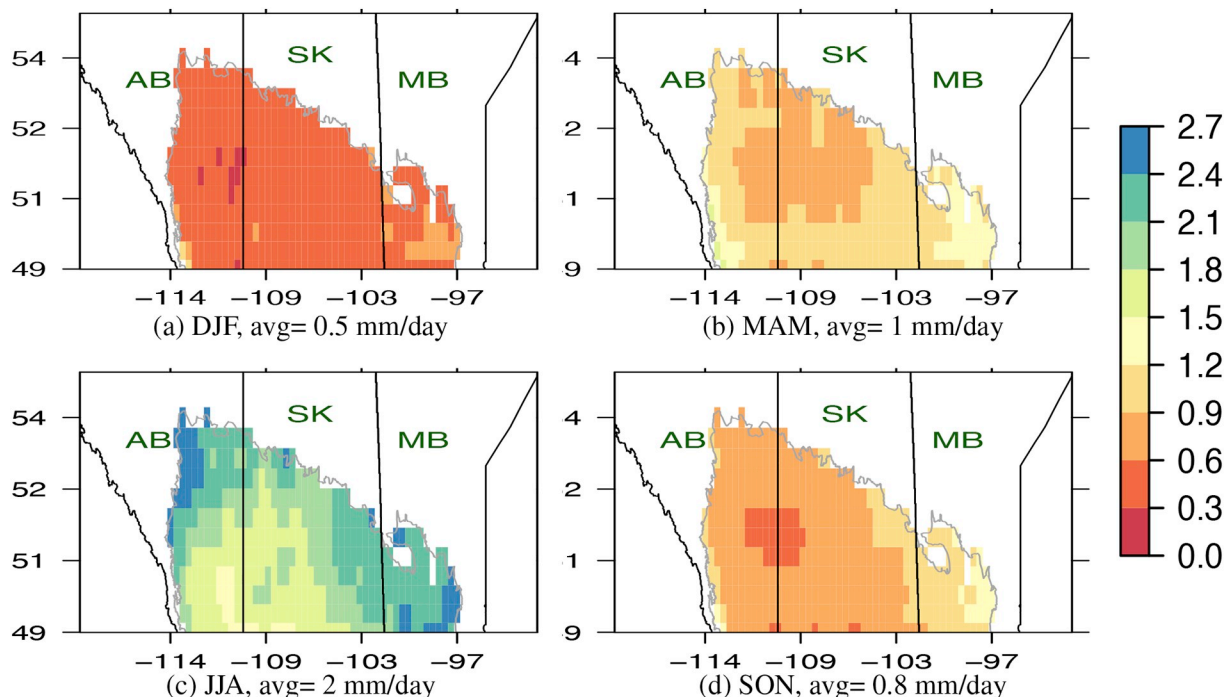
Fig. 6. Seasonal mean precipitation for 1979 to 2018 for (a) Dec-Feb, (b) Mar-May, (c) Jun-Aug, and (d) Sep-Nov. AB: Alberta, SK: Saskatchewan, MB: Manitoba. Legend unit is mm/d.

from CDFM in column 2 revealed to have lower RMSE (and bias) and higher correlation with SMAP over the globe compared with the dataset obtained from the Bayesian process in column 1. It can be, therefore, concluded that there is less uncertainty associated globally with the CDFM dataset. However, this step was only preliminary in understanding which dataset could be used to proceed with developing a global SM index.

5.2. Global and regional goodness-of-fit of the Beta distribution

For the SMOS/SMAP datasets, we performed the KS test on both CDFM and the Bayesian conditional process integrated datasets at a 95% confidence margin to filter the grids whose time series for each calendar month fit a Beta distribution. The SMOS/SMAP SM dataset from CDFM showed less flexibility to fit the Beta distribution regionally and globally. The number of grids that passed the KS test varied in each month. We summed the total number of passed grids in every month for
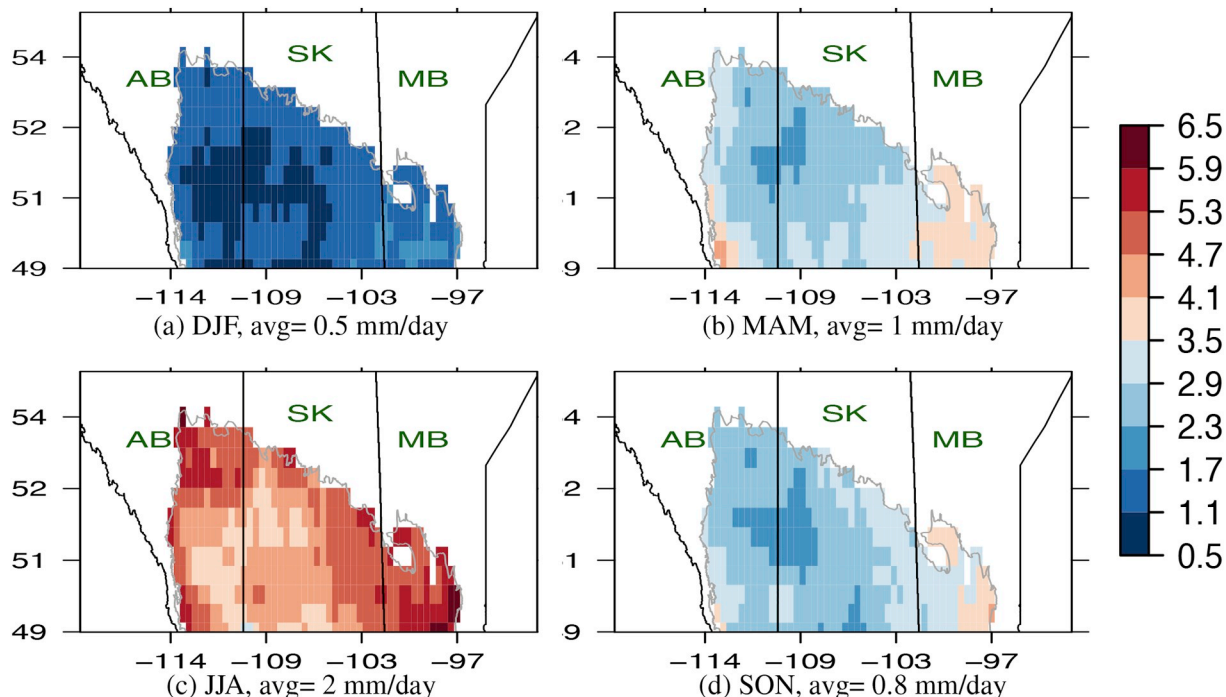

Fig. 7. Same as Fig. 6 but showing seasonal standard deviation of precipitation for 1979 to 2018 for (a) Dec-Feb, (b) Mar-May, (c) Jun-Aug, and (d) Sep-Nov.
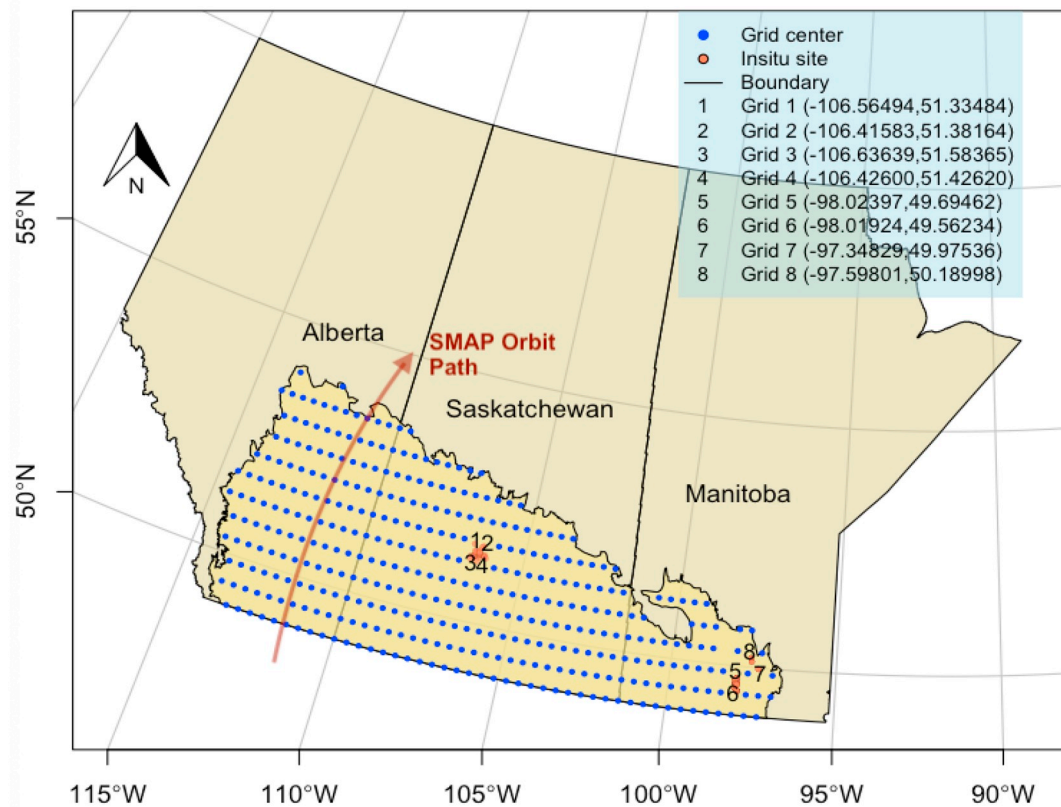
Fig. 8. The CPE is covered by 382 grids at ~36 km resolution and the blue dots are the centers of SMAP grids points. The ascending SMAP orbit path is drawn with an arrow. The locations of 111 in-situ sites are shown with orange dots with some transparency. The legend lists the number of the candidate grids from the global dataset and their respective longitude and latitude. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

12 months and divided them by the total number of grids that were overpassed for 12 months to get the percentage of grids that fit the Beta distribution both globally and for the CPE region. On the global scale, we eliminated 18,872 grids (from the total of 100,500 grids) that were permanent bare soil/deserts from the percentages of KS passing calculation, whereas for the CPE, only 4 grids were identified as bare soil and so we did not eliminate them. The summary of results is presented in Table 2. For comparison, we did a similar analysis of VIC and CCI datasets. Although there is a chance that the results in Table 2 are affected by spatial autocorrelation which may affect the total passing percentages, considerably more grids passed the KS test using the Bayesian conditional process than the CDFM approach both globally and in the CPE. While the CDFM process is widely practiced for integrating data, the assumption of perfect correlation between SMOS and SMAP, the process might not lead to better parametric fit.

Since we used monthly batches to extract indices, we performed a "breakpoint" analysis to check for temporal biases in month-to-month changes. For every grid, we made two frequency distributions: one by calculating the differences between the indices of every two consecutive dates that fall inside of the same calendar month; the second by calculating the difference between the indices of every two consecutive dates that fall on the last day of the old and first day of the new month. If breakpoints are not an issue, the expectation is that the two distributions are the same. A KS test with a p-value of 95% found that out of 382 grids in CPE, only 36 grids did not have the same distributions. So we concluded that the breakpoints could be an issue for less than 10% of grids.

We also investigated whether the 3-year SMOS and SMAP common period introduced any stability issue of the parameters. To do this, we performed cross-validation on the data by removing 20% of the days from common days of SMOS and SMAP in the 3 years overlapping

period and recalculated the SMOS/SMAP data, fitted the 4-parameter Beta distribution, and obtained new indices for the period of 2010–2018 for all the grids in the CPE. The index bias for each grid was very small, between $[-0.9$ to $0.9]$, and scattered like random noise across the CPE with mean and median of close to zero (~ -0.01). The cross-validation test showed that using only 3-year common period between SMOS and SMAP is stable enough to carry the index identification. However, the 3 years affected the choice between the Bayesian process and CDFM.

5.3. In-situ comparison

For the 6 validation grids in the CPE, we calculated BIAS, RMSE, correlation coefficient, and the Nash-Sutcliffe Efficiency (NSE) coefficient between the average in-situ SM data and the SM from SMOS, SMAP, SMOS/SMAP (from both CDFM and the Bayesian process). We tried two combinations of SMOS/SMAP SM time series. One approach was to fully simulate the SMOS/SMAP series even for the days that SMAP satellite had a retrieval. The other approach was to partially simulate the SMOS/SMAP series meaning that we replace the SMAP retrievals back in the SMOS/SMAP datasets for the days that they existed.

Fig. 10 shows the boxplots of summary statistics between the average in-situ SM values and SMOS, SMAP, fully simulated SMOS/SMAP using CDFM, partially simulated SMOS/SMAP using CDFM, and partially simulated SMOS/SMAP using Bayesian process. Comparing fully simulated SMOS/SMAP using CDFM (orange bars) and partially simulated SMOS/SMAP using CDFM (dark brown bars), we see that they have similar biases but in the partially simulated dataset, we get a better correlation and NSE with the in-situ data. Therefore, we decided to proceed with the partially simulated integrated time series in general. We then compare the partially simulated SMOS/SMAP using
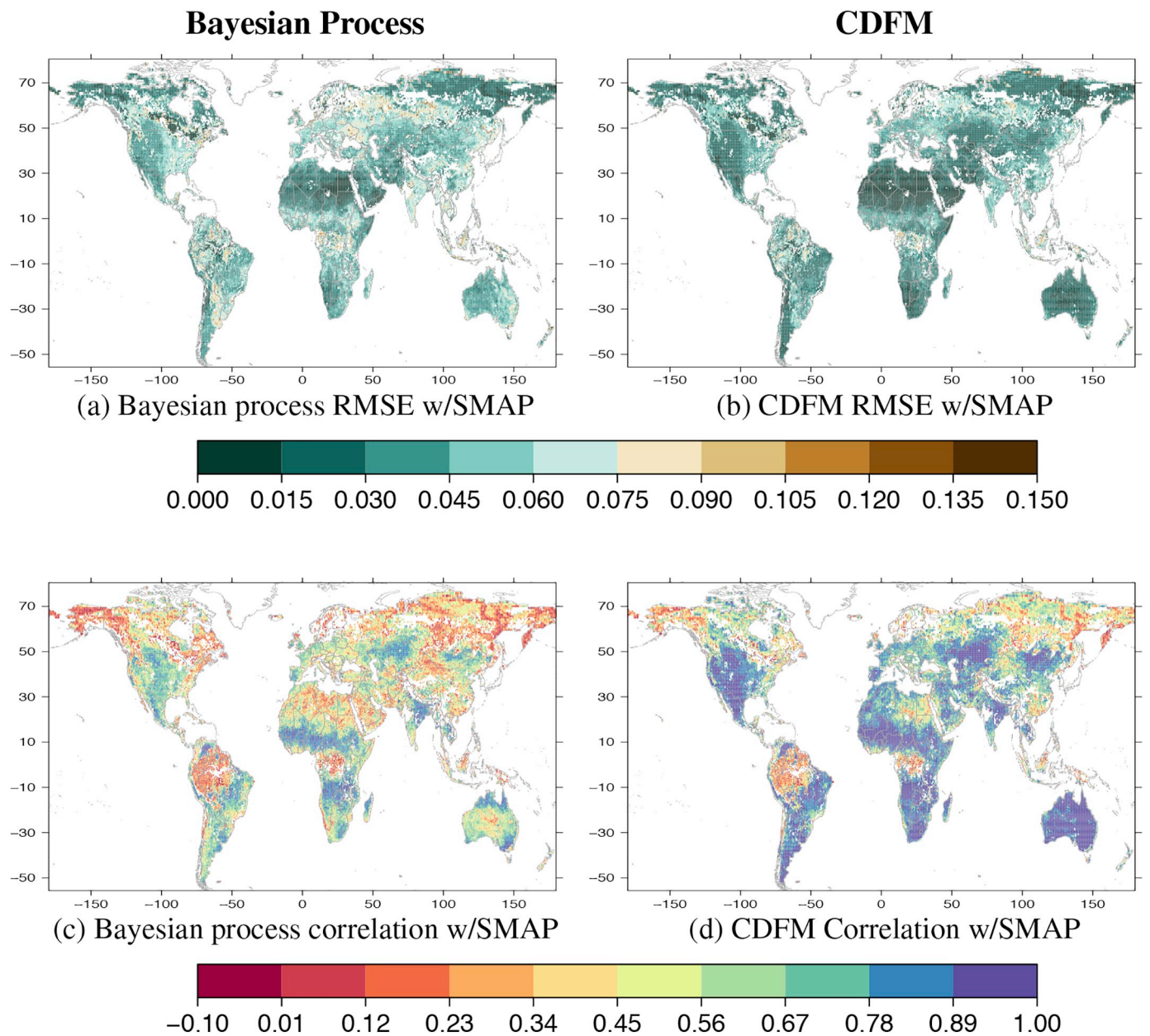
Bayesian Process

CDFM

(a) Bayesian process RMSE w/SMAP

(b) CDFM RMSE w/SMAP

(c) Bayesian process correlation w/SMAP

(d) CDFM Correlation w/SMAP

Fig. 9. Comparing RMSE (a) and (b) and correlation (c) and (d) between SMOS/SMAP and SMAP from the two methods of Bayesian conditional process and CDFM. The period of comparison is from 2015/03/31 to 2018/06/29.

Table 2
The percentage of grids that fit the Beta distribution with 95% confidence in the globe and CPE.

Dataset	Pass KS Test 81,628 global grids (Bayesian)	Pass KS Test 81,628 global grids (CDFM)	Pass KS Test 382 CPE grids (Bayesian)	Pass KS Test 382 CPE grids (CDFM)
VIC	58%	58%	64%	64%
SMOS/SMAP	76%	51%	81%	71%
CCI	86%	86%	98%	98%

CDFM (dark brown) and the partially simulated SMOS/SMAP using the Bayesian conditional (light brown) series. In this latter case, the two, again, have very similar bias and RMSE statistics. Some of the bias values from both the CDFM and the Bayesian process are slightly in the negative range for both datasets. But the average bias values are less than 0.04 m^3/m^3 and close to zero. Correlation (Corr in Fig. 10) and NSE for the SMOS/SMAP from partially simulated CDFM is slightly better than in the partially simulated SMOS/SMAP from the Bayesian

conditional process, specifically in terms of median and upper extreme values. However, such statistical improvements are at the cost of fewer grids fitting the Beta distribution in CPE and globally (Table 2), since the Bayesian process showed a significant improvement in terms of the number of grids that passed the KS test.

The decision for choosing an approach for integrating satellite data is ultimately about our final goal to develop an SM-based index. We decided to carry the rest of the analysis using the partially simulated
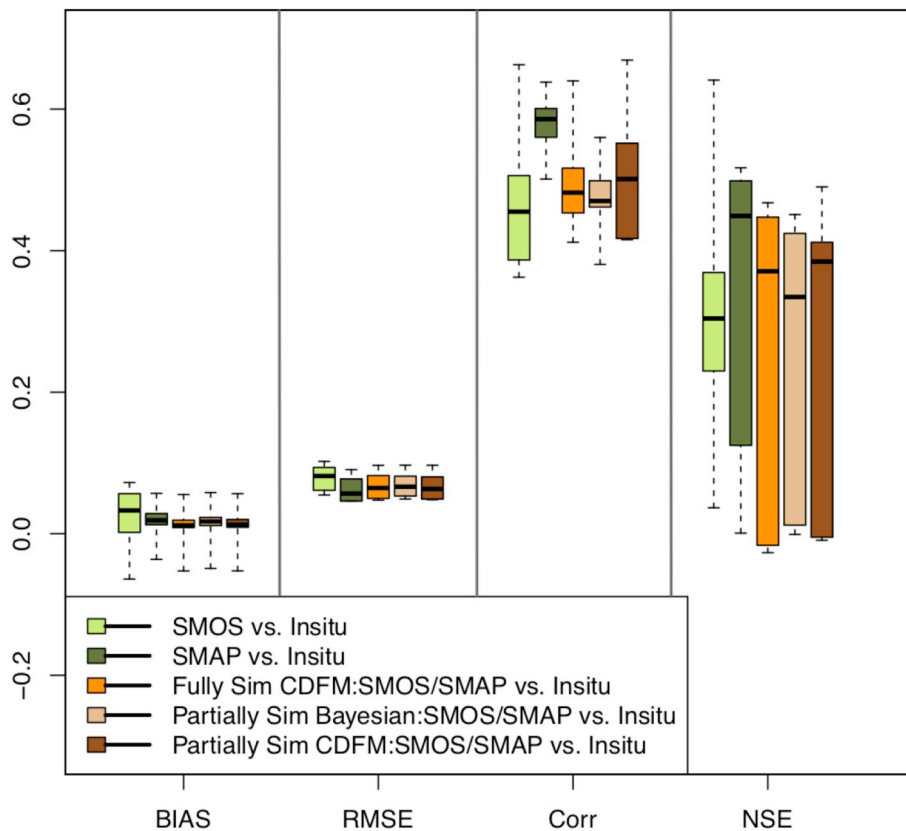
Fig. 10. Boxplot comparison between in-situ data and SMOS, SMAP, partially and fully simulated SMOS/SMAP using two types of CDFM approach, and partially simulated SMOS/SMAP SM using the Bayesian conditional approach. The period of comparison is from 2010 to 2018.

SMOS/SMAP dataset using the Bayesian conditional approach (with SMAP values replacing the simulated SM on any day and grid that has a SMAP SM observation). The best reasons for choosing the Bayesian process over the CDFM are firstly a similar accuracy to partially simulated SMOS/SMAP using CDFM and secondly, the higher number of grid coverage. The latter was especially important since for prediction purposes we need to have more SM data reliability fit a parametric distribution.

5.4. SM time series comparison

Fig. 11 shows a more in-depth analysis of the SM time series of the 4 validation sites (grids 1 to 4) in Saskatchewan and 2 validation sites in Manitoba (grids 5 and 6). Here we only show the period from 2014 to 2016. We are using partially simulated SMOS/SMAP and that is why the series covers over the observed SMAP data (blue circles). Overall, the SMOS/SMAP time series has captured the behavior of the on-site soil moisture dynamics over all the 6 sites.

In almost all plots of Fig. 11, the in-situ sites show a pronounced low level of soil moisture from November to early-April, expected from the CPE due to the frozen winter season. Except for grids 1 and 2, SMOS tend to show a lower amount of soil moisture for higher in-situ SM values. While SMOS series can be slightly more correlated to the on-site data, the SMOS/SMAP series showed a lower bias with the in-situ values (Fig. 10). This means SMOS/SMAP on average has less sum of differences with the in-situ data while it can capture the in-situ seasonality reasonably well. The SMOS data also capture seasonality but show a bigger bias. Fig. 11 (as well the P–P plots of the SMAP vs. the SMOS/SMAP) confirm that the SMOS/SMAP dataset has captured the SMAP pattern while using the information that SMOS provided (having more revisit frequency and a longer period of record).

5.5. Comparison between drought indices

A comparison of the SPI-1 with the average monthly SMOS/SMAP index allows assessing time of co-occurrence of historic events as well as their intensity and duration. We calculated the monthly average precipitation, SPI-1, as well as monthly average of drought indices from SMOS/SMAP, VIC, and CCI for each of the candidate grids in the CPE. In Fig. 12, we present the time series of such analyses for grids 2, 3, and 6 and for the period from 2012 to 2017.

In the first panels, average monthly precipitation and SPI-1 are plotted. In the second panels, we can compare the pattern and overall temporal variability of SMOS/SMAP, CCI, and VIC average monthly indices. The SMOS/SMAP and CCI indices follow each other quite consistently and closely, while VIC shows different behavior from the other two on some months. The CCI and SMOS/SMAP also corresponds more with the SPI-1. All grids and indices captured the extreme droughts of 2012 and 2017 quite well, although they show different levels of severity. The year 2013 was abnormally dry (D0), dominant in southern Alberta and Saskatchewan, for most of August since rainfall in the region was 40–60% of normal. This is reflected in the plots of grids 2 and 3 showing Saskatchewan. The wet summer and fall of 2016 is also captured in the plots.

In Fig. 13, the spatial extent and severity of the monthly average of the SMOS/SMAP drought indices for August 2012, 2013, 2016, and 2017 are compared with the monthly spatial maps released by AAFC's CDM. The monthly maps from CDM are produced through data sources from consultations with federal, provincial, regional, and academic scientists to establish a single drought index based on the five categories of intensity of drought (Agriculture Agrifood Canada, 2019). Therefore, CDM is more comparable to USDM than our SM-based inex maps, but the purpose here is to compare the overall patterns of dryness. The CDM maps are generated Canada-wide but here, we cropped the maps to
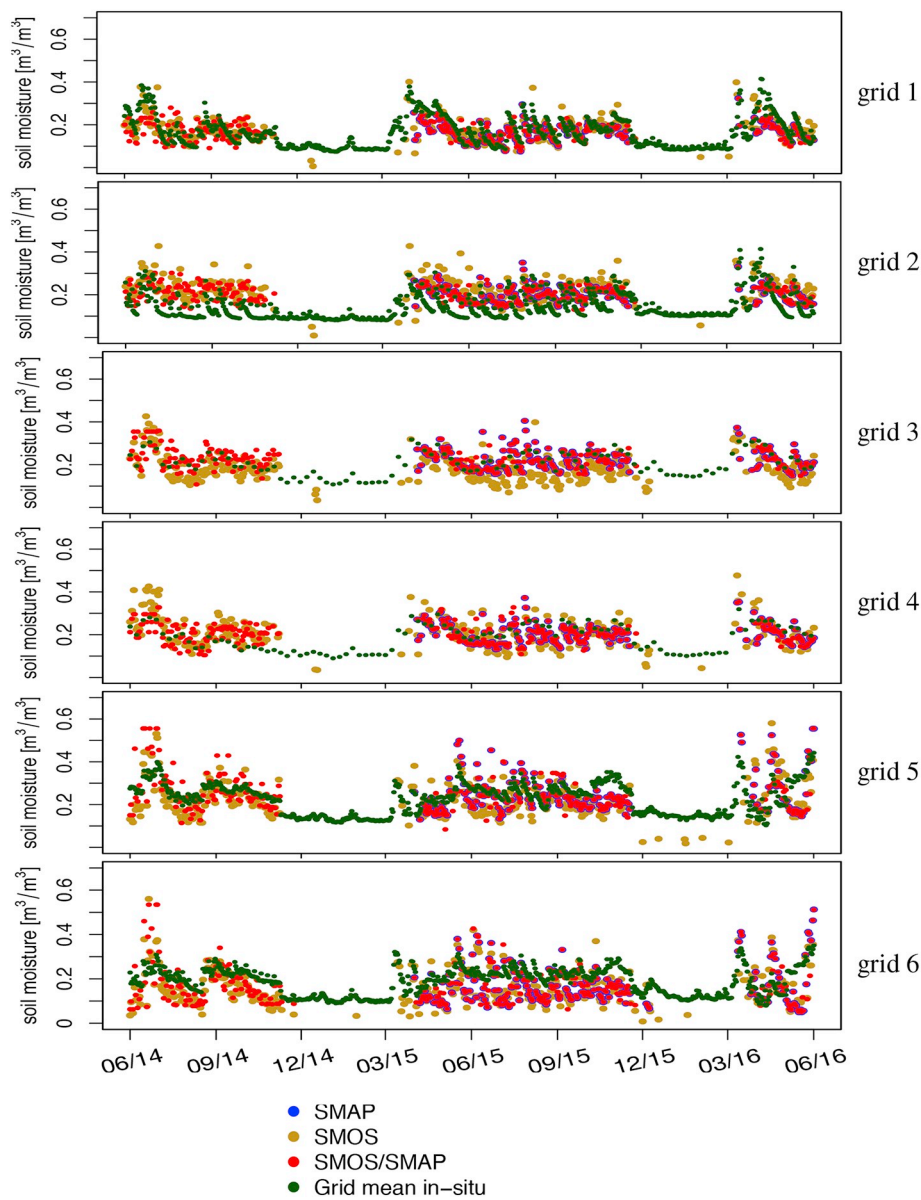
Fig. 11. Time series for the validation sites from 2014 to 2016. Note that the partially simulated SMOS/SMAP (red dots) overlap with the SMAP (blue dots) for days that SMAP has a retrieval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

show only the CPE. Visual examination of SMOS/SMAP with CDM maps shows that SMOS/SMAP captured extreme events in terms of spatial extent and severity of drought conditions quite well. For example, years 2012 and 2013 were moderately dry (D1) years since February, but particularly from August to October in southern Manitoba and the southeast part of Saskatchewan, less than 70% of the normal amount of rain fell. Very relevant pattern of dryness is captured in the left panels of Fig. 13. The year 2016 was a very wet year with an overabundance of precipitation in most of the CPE, leading to the deletion of both abnormally dry (D0) and moderate drought (D1) conditions in the area from the past years. Only a small packet in southeastern Saskatchewan began to show short-term dryness with D0-D1. This is, again, well captured in the SM-index map of Fig. 13. Finally, in 2017, severe to exceptional drought (D2 to D4) in Alberta and Saskatchewan negatively impacted most crops and pasture yields and quality. Pockets of extreme drought (D3) caused wildfires that burned more than 100,000 ha. This is also well shown in Fig. 13. The SM-based index map showed to be capable of capturing the sptail and temporal patterns and severity of drought conditions in most cases in the CPE.

### 5.6. Joint distribution heat maps

Although studying the monthly pattern of indices show the overall agreement and disagreement of the indices, to understand the similarity of the daily index maps, we generated two-by-two comparison heat maps from the actual daily indices of SMOP/SMAP, VIC, and CCI. Fig. 14 shows heat maps from such comparisons for selected grids (2,3, and 6) for the entire period of 2010 to 2018. Percentiles from 1 to 100 are partitioned into 5-percentile increments for each of the indices and the number of times that two indices at the time fall within the same bin is counted. For each grid, a heat map denser diagonally would be more ideal because it shows that the two indices identify similar severity of drought or pluvial conditions numerically.

There are warmer colors in higher percentiles between VIC and SMOS/SMAP and between VIC and CCI percentiles meaning that, in the CPE region, the daily maps of indices from VIC datasets most likely agree more with CCI or SMOS/SMAP in terms of severity of the pluvial conditions but agrees less in terms of drought conditions. CCI and SMOS/SMAP show noticeable agreement in recognition of drought
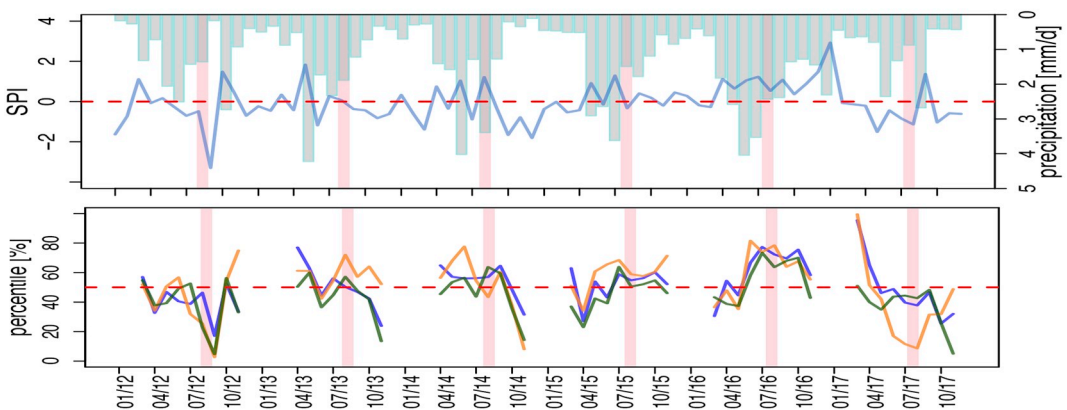
Grid 2:
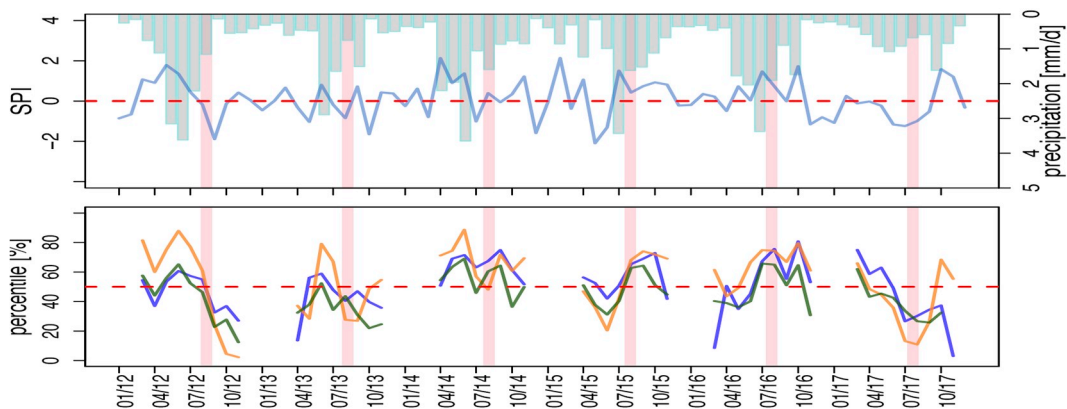
Grid 3:
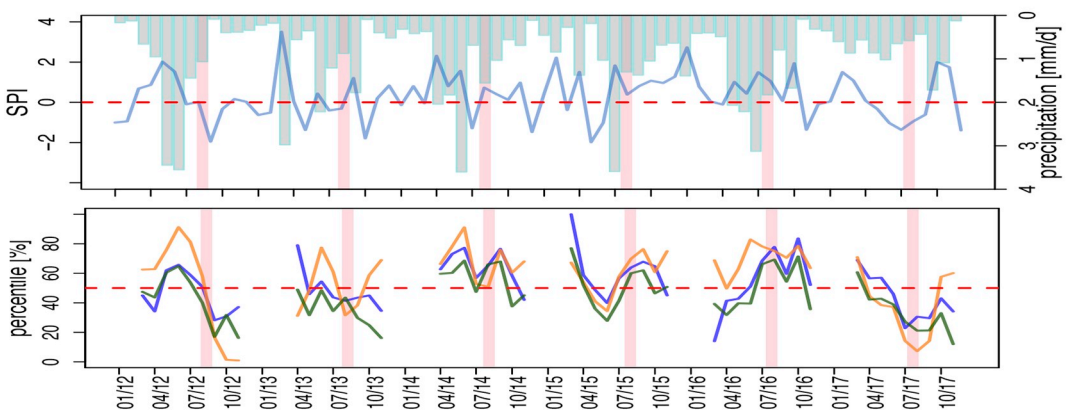
Grid 6:

—— SMOS/SMAP   —— VIC–SM   —— CCI

Fig. 12. Comparative evaluation between mean monthly precipitation, SPI-1 (top) and mean monthly drought indices from SMOS/SMAP, VIC, and CCI (bottom) for test grids in the CP for the period from 2012 to 2017. The month of August of each year is shown by a pink shadow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
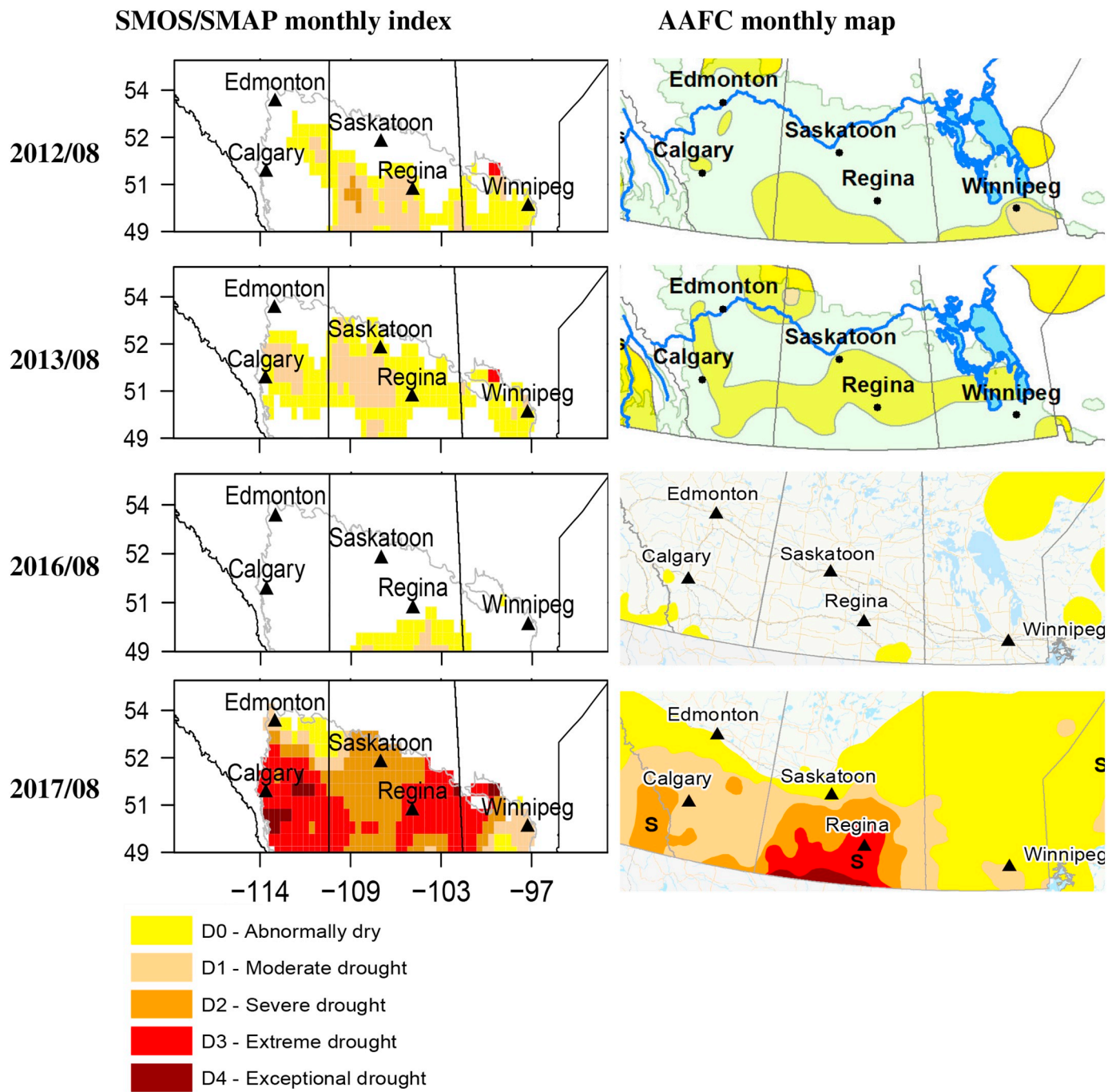
SMOS/SMAP monthly index   AAFC monthly map

Fig. 13. Comparison of indices from SMOS/SMAP with those from CDM for the months of August and the years of 2012, 2013, 2016, and 2017. Note that projection of the right and left column are different, and even for AAFC's CDM (right column), the projection of year 2012–2013 is different from that in years 2016–2017. Drought conditions outside of the CPE is not plotted in the left column.

conditions in terms of percentiles especially for the grids in Manitoba. CCI and SMOS/SMAP indices also show a clearer diagonal relation, which is less observed between VIC index and these two indices.

5.7. All indices comparison

In comparing all indices, we want to answer two questions. Firstly, which of these indices are more reliable, and secondly, which of these indices are more suitable and has the potential of assessing agricultural drought impacts. Reliability is a matter of matching ground truth; as far as the CPE region is concerned, SMAP datasets showed to be the most reliable, with the partially simulated SMOS/SMAP next in reliability.

Suitability of use can be assessed by more statistical tests to find which index is more harmonized with major drought and pluvial conditions and can be used as a monitoring and prediction tool. We compared all 382 grids in the CPE in terms of their statistics and plotted the CDFs of CCI, SMOS/SMAP, and VIC indices in Fig. 15 knowing that, in estimating the indices, every grid on any given day should have an equal chance of having all the percentiles. In the long run, this means that the CDF of all the indices should be very close to the 1:1 line. Although that is the expectation, none of the 3 indices in Fig. 15 are completely matching the 1:1 line, meaning that they all have some sort of bias. Fig. 15 also explain the tails or extreme indices. For further investigation, in Fig. 16 the 10th, 50th, and 90th percentiles of VIC, SMOS/SMAP,
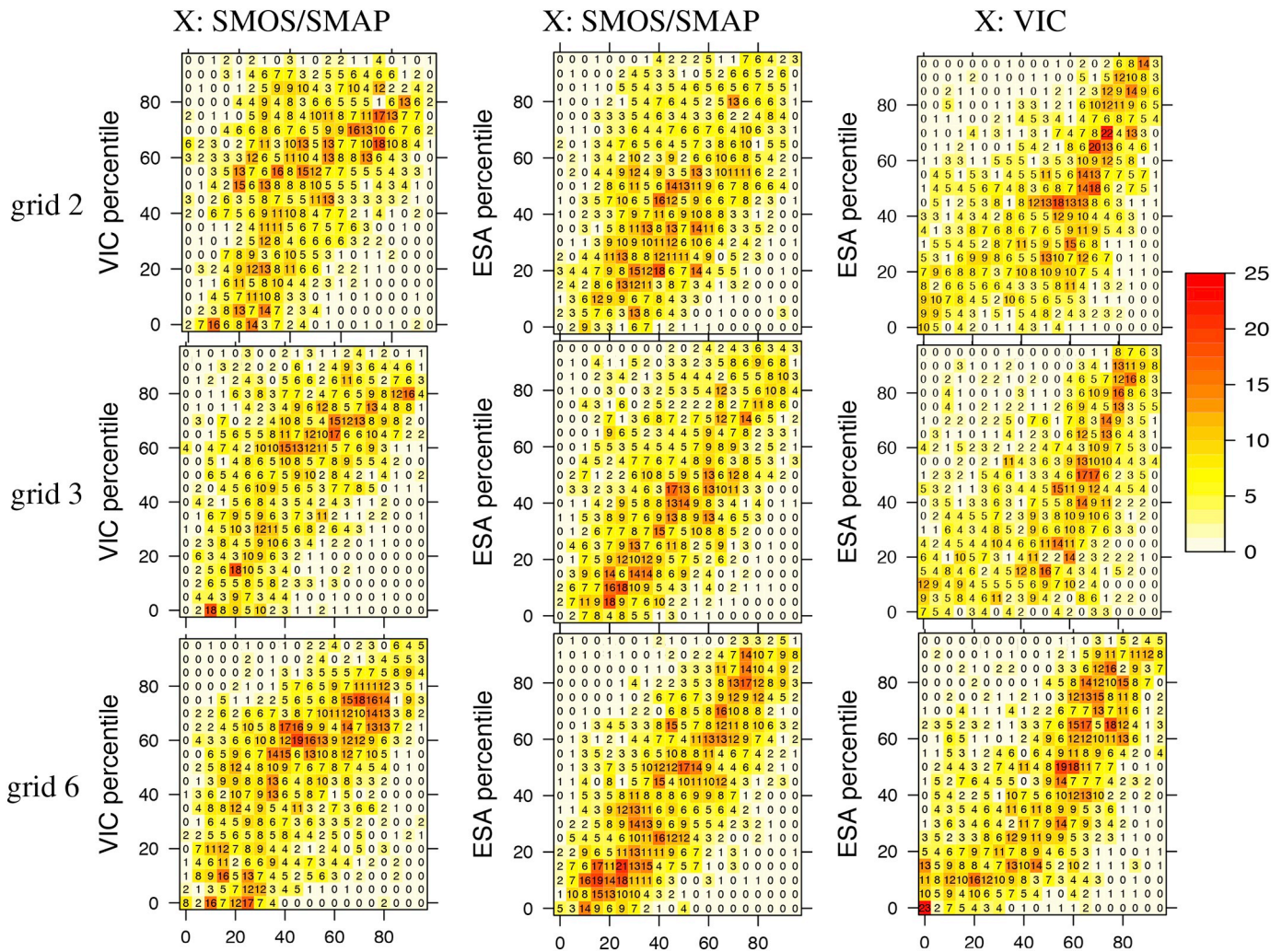
Fig. 14. Two-by-two comparison of VIC, SMOS/SMAP, and CCI using joint distribution heat maps. The period represented by this comparison is 2010–2018.
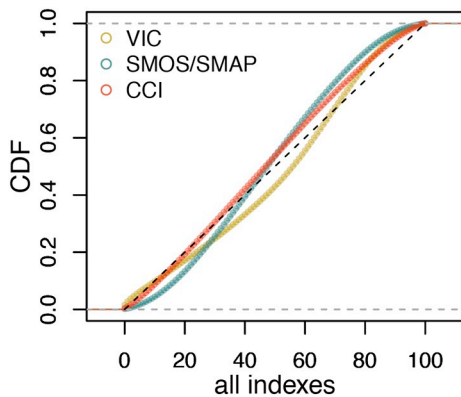

Fig. 15. Summary CDF of each index for the 382 grids of the CPE region. The period represented is 2010–2018.

and CCI are drawn.

Studying 15 together with Fig. 16 reveals that all of the three indices can show extreme pluvial conditions slightly less wet than what might be happening on the ground. The SMOS/SMAP index can show extremely dry conditions slightly less dry than what might be happening on the ground as well. The reason behind this behavior is that SMOS/SMAP is still having shorter tails due to having a shorter record, even after we extended its tails using the asymptotic model. However, this

might not cause a major problem when translating the SMOS/SMAP indices into the colour map according to D0-D4 categories given that each category includes a range of percentiles. For example, D1 and D0 categories allow indices to be anywhere in the range of 11% to 20% and 21% to 30%, respectively. Therefore, SMOS/SMAP index is most likely to categorize a D4 drought as a D3 and a D3 drought as a D2 but much less likely to categorize a D2 drought as a D1. Around the 50th percentiles, CCI and SMOS/SMAP indices show a very similar assessment which probably corresponds realistically with what is happening on the ground. This is evident also in Fig. 15 where the two CDFs cross each other around the middle percentiles. VIC indices, in contrast, show a great negative bias for indices between 30% to 70%. An event with the non-exceedance probability of 30 percentiles can be presented with an index of 40% in the VIC index and therefore, in mid percentiles, VIC tends to show the D0-type droughts as no dryness at all. This is further confirmed in Fig. 16. VIC also shows extremely severe droughts (D4 to D2) that fall in the 1% to 10% index range more severe than it might be on the ground, for example it can show a D3 drought as a D4.

Overall, the CCI indices are almost on the 1:1 line in lower indices (less than 30%) so we can say that CCI indices in lower percentiles tend to show the ground as dry as it is in reality. CCI and SMOS/SMAP performed closest to each other in capturing seasonality (due to close correlation) and their spatial patterns. Their CDFs show similar patterns and we can say that SMOS/SMAP CDF is an exaggerated version of CCI, which means that SMOS/SMAP sometimes tones the dry and wet situations slightly down. SMOS/SMAP indices bias can be sensitive to one
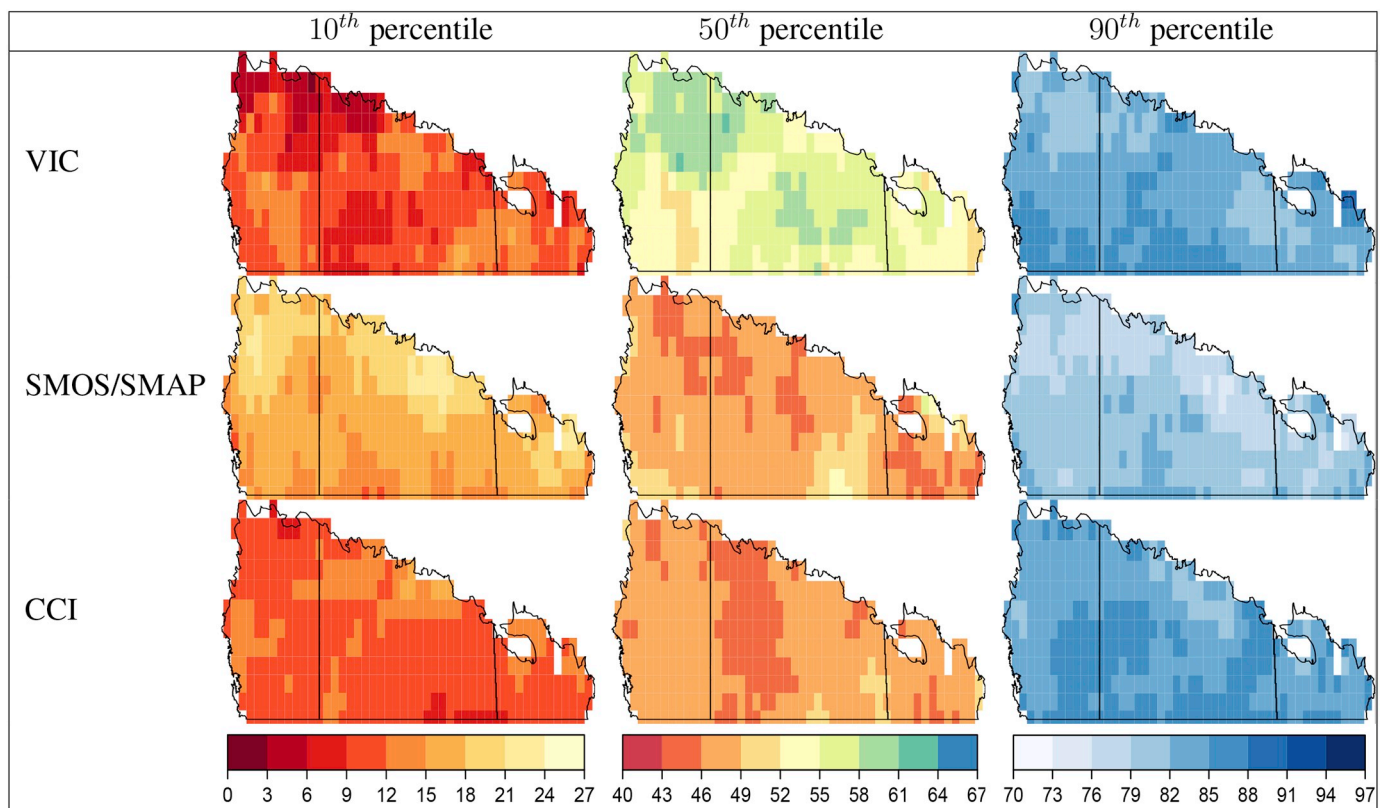
Fig. 16. Comparing VIC, SMOS/SMAP, and CCI spatial pattern of tails distributions and medians.

or more of factors such as the way that the two types of data got integrated, the record length, or the regional characteristics of indices. Understanding the behaviors of indices are important in interpreting the daily index maps, however, such analysis also brings out the importance of investigating the source and removal of biases, although, this is not in the scope of this study.

5.8. Global drought monitor maps

Fig. 17d shows an example of the SMOS/SMAP-based index mapped for the GSMIM for August 31, 2017. Areas without vegetation and deserts are eliminated using FAO IIASA mask (FAO, 2019). The GSMIM is a 3-day composite map for more continuous coverage, e.g., August 31 is the average of August 29, 30, and 31. For comparison, other drought maps from other organizations and agencies are presented although not all of them represent merely soil moisture nor are they only based on the day of August 31, 2017. CDM (Fig. 17b), for example, is an average over the entire month of August. USDM (Fig. 17a) presents weekly maps and, similar to CDM, it integrates several drought indices and experts' inputs (Svoboda et al., 2002). European Drought Monitor (EUDM in Fig. 17c) is also a combined drought indicator and its maps here represent the last 10 days of August. For our GSMIM, we include wetness percentiles similar to AFDM (Fig. 17e) except that AFDM is colour-coded as shades of green instead of blue.

The extension of abnormal dryness (D0) to exceptional (D4) drought conditions from the CPE in CDM shows in the USDM map in the high plains of Montana, northern plains of Nebraska, and the Dakotas. Pasture and crop conditions were poor to very poor in these areas during August. The northern Rockies and Pacific Northwest also experienced moderate (D1) to severe (D2) drought. Some of the D0 and D1 situations lingered in parts of southern California and Arizona, which was not captured by GSMIM. One reason for this can be human interference and the use of groundwater wells for irrigation during the warm season. This was further confirmed by NOAA's monthly report for

August 2017 which indicated that the southern California and Arizona drought were mostly based on groundwater and GRACE observations (NOAA, 2017). Abnormal dryness and moderate hydrological drought expanded in the Northeast of the U.S. as well. The GSMIM map also captured the drought conditions in the Northwest and interior Southeast of Alaska, as well as the pluvial conditions everywhere else in Alaska. GSMIM is mostly in agreement with AFDM but disagrees in terms of severity of drought in some regions, particularly in South Africa and Madagascar. The areas where we see the lack of correspondence with GSMIM are exactly the areas of poor or no rain gage availability and hence, we speculate that this affects AFDM is showing similar conditions as those in GSMIM. Summer 2017 was exceptionally dry in Europe, particularly Southwest Europe which was hit by heatwave "Lucifer". By August 31, most of the Southwest recovered from the heatwave, but the drought continued to linger in some of the Scandinavian countries, East and Central France, West and South of Spain, as well as in Ukraine, Serbia, and Croatia (C3S, 2017). We compared more maps of GSMIM from various dates with the regional maps and overall, there is a high correspondence among the drought maps from different regions and GSMIM. Future developments can, however, focus on the quality of performance of the GSMIM in various areas of the globe.

5.9. Data and maps availability

For accessibility and as an outreach tool, we have developed an online platform that overlays these maps over Google Maps and stores the archive for download. The website can be reached at http://hydrology.princeton.edu/ sadri/nasasmap_globe/index.html. The CDM monthly maps are available at https://open.canada.ca/data/en/dataset/292646cd-619f-4200-afb1-8b2c52f984a2. USDM weekly maps of the U.S. are available at https://droughtmonitor.unl.edu/Maps/MapArchive.aspx. GRACE-based assimilated weekly maps are available at https://

(a) USDM

(b) CDM

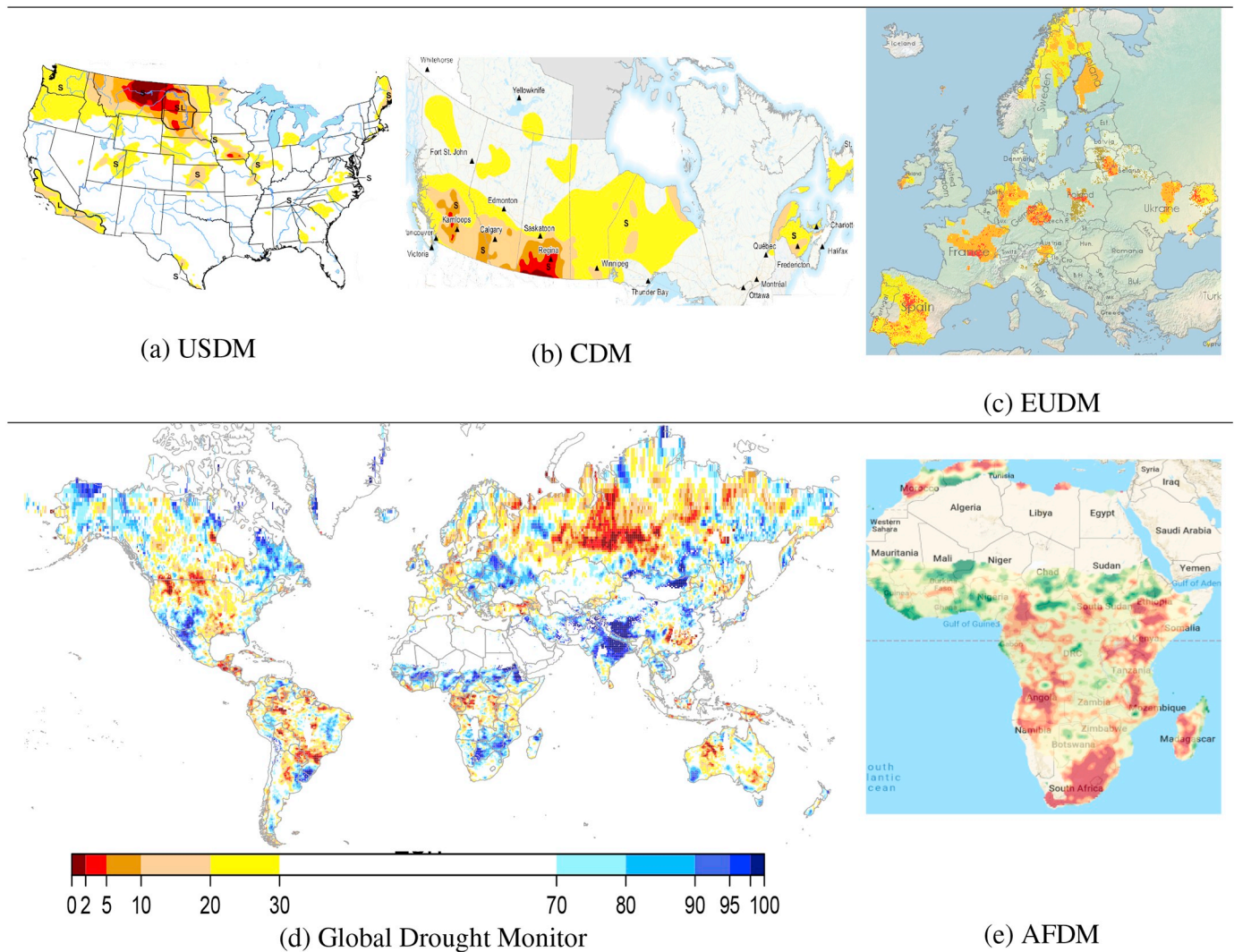(c) EUDM

(d) Global Drought Monitor

(e) AFDM

Fig. 17. Comparison between GSMIM and other agencies' drought monitor maps around August 31, 2017. Areas of the globe that have bare soil (i.e. dominated by deserts) have been removed in GSMIM.

nasagrace.unl.edu/Archive.aspx. The 10-day European Drought Monitor maps are from https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1111. AFDM is generated by Princeton Climate Analytics (PCA) at https://platform.princetonclimate.com/AFDM/interface.php?locale=en.

6. Conlusions

In this study, we developed GSMIM, the first operational global SM-based index product using merged near-real-time remote sensing data. The constructed dataset used SMAP observations as predictand while taking advantage of the spatial information and temporal coverage that SMOS provides since 2010. We make the index maps available online every 24 h. In its current form, the system only monitors the SM index state of the globe for studying drought and flood evolutions and does not address impacts.

The GSMIM is distinguished from other operating systems in that it uses 4-parameter Beta distributions with asymptotics to describe its limiting behavior and extend its endpoints beyond the observed data. Furthermore, the value of this study is enhanced by the use of remotely sensed data which characterizes the drought condition spatially and temporally in a near-real-time capacity. The algorithm developed in this study is implemented operationally for the use of SMAP/SMOS dataset as the basis for GSMIM, however, the methodology for integrating two or more datasets, fitting a non-standard Beta distribution whose extreme tails go beyond the observed data, and translating the incoming integrated soil moisture into index percentile is essentially applicable to any datasets in any regional or temporal scales. Since GSMIM uses a parametric distribution, it also has great potential for being developed further into a prediction tool. In this current stage, the 7–8 day temporal latency of GSMIM comes from the latency of the SMOS SM and if that is addressed, the GSMIM can easily be synced with SMAP's latency (3-day). Compared with the C3S SM, GSMIM's 7–8 day lag is lower which may be critical for some applications.

The differences in VIC-based indices and other indices can be due to the way that these data are produced. Soil moisture through satellite data is calculated from the C-band passive microwave signal, whereas LSM models such as VIC integrate soil moisture from antecedent meteorological forcing mostly precipitation. It can be argued that the ground measurements could be the most accurate but they are also prone to errors and uncertainties and can be altered or not representative. Moreover, different sites can have different instruments and protocols for onsite measurement, be far more sparsely distributed in space and in time, and not necessarily representative of large-scale soil moisture. The synergy between precipitation index SPI-1 and SMOS/SMAP index makes this index more powerful for future use in depicting drought conditions in the CPE than the use of observations separately.

SMOS/SMAP indices might not show the extreme drought condition as severe as they are on the ground albeit capturing their presence and spatial extent. This can be due to the fact that both CCI and VIC SM datasets have longer record periods than SMOS/SMAP SM datasets. The CCI still has biases to be recognized and removed, but it performed closest to an index. The CCI, however, is not currently available in near-real-time. SMOS/SMAP's indices performed better than VIC indices. Because SMOS and SMAP monitor soil moisture directly and provide critical information for drought early warning, GSMIM can be a reasonable global index with operational capacity and competitive quality. It is important, however, to acknowledge that such an operational tool relies on continuous efforts to improve the system, e.g. recalibration and other updates, which can lead to changes in the product.

Although this study overcame several obstacles, it also shed light on other future developments and research. It is important that future developments focus on merging CCI with SMOS/SMAP SM data for building a version of GSMIM based on a longer-term record while maintaining the near-real-time capacity of the product. Downscaling of the SMOS/SMAP indices and correcting biases (such as the bias detected from the SMOS/SMAP indices in the CPE shown in Fig. 15) are other possible future directions. Next steps can be also on understanding the relationship and feedback between surface SM, root zone soil moisture, crop development stages, and land surface temperature (LST). However, evaluation of drying and wetting trends of SM and associated feedback with LST would still remain difficult in many regions worldwide due to a lack of availability of global long-term SM observations. It would be crucial to run validation in other regions, especially in underrepresented parts of the world. The relationship between point measurements in a region, SM, and crop yields is another interesting direction that this work can take. Developing GSMIM has been a step forward toward building a global SM index prediction tool and bridging the gap between the satellite data information and on-the-ground improvements in water and food security.

Funding

Authorship Contribution Statement

Sara Sadri: Conceptualization, Methodology, Software, Formal analysis, Writing original and final draft. Ming Pan: Resources, Methodology, Editing, Consulting. Yoshihide Wada: Editing, Consulting. Noemi Vergopolan: Resources, Editing. Justin Sheffield: Resources. James S. Famiglietti: Resources, Consulting. Yann Kerr: Resources, Conceptualization, Consulting, Research. Eric F. Wood: Supervision, Conceptualization, Funding acquisition.

*The authors would like to sincerely thank the reviewers of this manuscript, whose comments enhanced the science and fluency of the work.

Declaration of Competing Interest

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

References

Adler, R., Huffman, G., Chang, A., Ferraro, R., Xie, P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., Nelkin, E., 2003. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979 present). J. Hydrometeorol. 4, 1147.

AghaKouchak, A., Nakhjiri, N., 2012. A near-real-time satellite-based global drought climate data record. Environ. Res. Lett. 7.

Agriculture Agrifood Canada, 2019. About the Canadian Drought Monitor. URL: http://www.agr.gc.ca/eng/agriculture-and-climate/drought-watch/canadian-drought-monitor/about-the-canadian-drought-monitor/?id=1463576995558. [Last access: Feb 2020].

Al Bitar, A., Mialon, A., Kerr, Y.H., Cabot, F., Richaume, P., Jacquette, E., Quesney, A., Mahmoodi, A., Tarot, S., Parrens, M., Al-Yaari, A., Pellarin, T., Rodriguez-Fernandez, N., Wigneron, J.P., 2017. The global SMOS Level 3 daily soil moisture and brightness temperature maps. Earth Syst. Sci. Data 9, 293–315.

Albergel, C., Dorigo, W., Reichle, R.H., Balsamo, G., de Rosnay, P., Muoz-Sabater, J., Isaksen, L., de Jeu, R., Wagner, W., 2013. Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing. AMS 1259–1277.

Alvarez-Garreton, C.A.W., Western, D.R., Crow, W., Su, C., Robertson, D., 2016. Dual assimilation of satellite soil moisture to improve streamflow prediction in data scarce catchments. Water Resour. Res. 52, 5357–5375.

Beck, H., Wood, E., Pan, M., Fisher, C., van Dijk, D.M.A., Adler, T.M.R., 2019. Mswep v2 global 3-hourly 0.1 precipitation: methodology and quantitative assessment. Bull. Am. Meteorol. Soc. 100, 473–500.

Berg, A., Sheffield, J., Milly, P.C.D., 2017. Divergent surface and total soil moisture projections under global warming. Geophys. Res. Lett. 44, 236–244.

Bhuiyan, H.H.M., Powers, J., Friesen, M., Pacheco, A., Jackson, T., Cosh, M., Colliander, A., Berg, A., Rowlandson, T., Bullock, P., Magagi, R., 2018. Assessing SMAP soil moisture scaling and retrieval in the Carman (Canada) study site. Vadose Zone J. 17.

Bindlish, R., Jackson, T.J., Chan, S., Colliander, A., Kerr, Y., 2018. Integration of SMAP and SMOS L-BAND observations. Technical report GSFC-E-DAA-TN42820. NASA Goddard Space Flight Center, Greenbelt, MD, USA.

Bolten, J., Crow, W., Zhan, X., Jackson, T., Reynolds, C., 2010. Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. IEEE J. Selected Top. Appl. Earth Observ. Remote Sens. 3, 57–66.

C3S, 2017. Climate in 2017 - Focus Region: Southwest Europe. URL: https://climate.copernicus.eu/climate-2017-focus-region-southwest-europe.

C3S, 2020. Soil moisture gridded data from 1978 to present. url: https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-soil-moisture?tab=overview.

Carrão, H., Russo, S., Sepulcre-Canto, G., Barbosa, P., 2016. An empirical standardized soil moisture index for agricultural drought assessment from remotely sensed data. J. Appl. Earth Observ. Geoinform. 48, 74–87.

Centre Aval de Traitement des Données SMOS, 2017. CATDS data access on ftp site.

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

Cherneski, P., 2018. The impacts and costs of drought to the Canadian agriculture sector. URL. https://www.drought.gov/nadm/sites/drought.gov.nadm/files/activities/2018Workshop/8_3_CHERNESKI-Agricultural_Drought_Impacts_Canada.pdf.

Chung, D., Dorigo, W., Reimer, C., Hahn, S., Melzer, T., Paulik, C., Vreugdenhil, M., Wagner, W., Kidd, R., 2019. Algorithm theoretical baseline document (ATBD) d2.1 version 04.4, active soil moisture retrievals, ESA CCI soil moisture. URL. http://dap.ceda.ac.uk/thredds/fileServer/neodc/esacci/soil_moisture/docs/v04.4/CCI2_Soil_Moisture_DL2.1_ATBD_v4.4_02_active.pdf.

Climate Change Service, 2018. Soil moisture gridded data from 1978 to present. URL: https://cds.climate.copernicus.eu/portfolio/dataset/satellite-soil-moisture. [Laste access: Apr 2020].

Coccia, G., Siemann, A.L., Pan, M., Wood, E.F., 2015. Creating consistent datasets by combining remotely-sensed data and land surface model estimates through Bayesian uncertainty post-processing: the case of land surface temperature from HIRS. Remote Sens. Environ. 170, 290–305.

Colliander, A., Jackson, T.J., Bindlish, R., Chan, S., Das, N., Kim, S.B., Cosh, M.H., Dunbar, R.S., Danga, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., Yueh, S., 2017. Validation of SMAP surface soil moisture products with core validation sites. Remote Sens. Environ. 191, 215–231.

de Jeu, R., van der Schalie, R., Dorigo, W., Paulik, C., Scanlon, T., Kidd, R., Reimer, C., 2018. Algorithm Theoretical Basis Document Soil Moisture Products from Active and Passive Microwave Sensors. ECMWF Copernicus Report. Copernicus Climate Change Service, URL. http://datastore.copernicus-climate.eu/c3s/published-forms/c3sprod/satellite-soil-moisture/algorithm-theoretical-baseline-v2.2.pdf.

Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., Jackson, T., 2011a. The international soil moisture network: a data hosting facility for global in situ soil moisture measurements. Hydrol. Earth Syst. Sci. 15, 1675–1698.

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

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A., Zamojski, D., Cordes, C., Wagner, W., Drusch, M., 2012. Global automated quality control of in situ soil moisture data from the international soil moisture network. Vadose Zone J. 12.

Dorigo, W., Gruber, A., Jeu, R.D., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R., Kidd, R., 2015. Evaluation of the ESA CCI soil moisture product using ground-based observations. Remote Sens. Environ. 162, 380–395.

Drusch, M., Scipal, K., Rosnay, P.D., Balsamo, G., Andersson, E., Bougeault, P., Viterbo, P., 2009. Towards a Kalman filter based soil moisture analysis system for the operational ECMWF integrated forecast system. Geophys. Res. Lett. 36.

EM-DAT, 2020. International Disaster Database. Université Catholique de Louvain, Louvain, Brussels, Belgium URL. www.emdat.net.

Entekhabi, D., Rodriguez-Iturbe, I., Castelli, F., 1996. Mutual interaction of soil moisture state and atmospheric processes. J. Hydrol. 184, 3–17.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., 2010. The soil moisture active passive (SMAP) Mission Proc. IEEE 98, 704–716. https://doi.org/10.1109/JPROC.2010.2043918.

ESTR Secretariat, 2014. Prairies Ecozone: evidence for key findings summary. Canadian Biodiversity: Ecosystem Status and Trends 2010, Evidence for Key Findings Summary. Technical Report 4. Canadian Councils of Resource Ministers, Ottawa, ON Ix + 115 p.

European Space Agency, 2012. Ecv sm data access. https://esa-soilmoisture-cci.org/index.php?q=node/134 [Last access: Feb 2020].

Famiglietti, J.S., Devereaux, J.A., Laymon, C.A., Tsegaye, T., Houser, P.R., Jackson, T.J., Graham, S.T., Rodell, M., van Oevelen, P.J., 1999. Ground-based investigation of soil moisture variability within remote sensing footprints during the southern great plains 1997 (sgp97) hydrology experiment. Water Resour. Res. 35, 1839–1851.

Famiglietti, J.S., Ryu, D., Berg, A.A., Rodell, M., Jackson, T.J., 2008. Field observations of soil moisture variability across scales. Water Resour. Res. 44, 1–16.

FAO, 2009. How to Feed the World in 2050. URL. http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf.

FAO, 2019. Fao soils portal. Online. URL. http://www.fao.org/soils-portal/soil-degradation-restoration/global-soil-health-indicators-and-assessment/global-soil-health/en/.

Government of Canada, 2017. Soil Monitoring Stations. URL. https://agriculture.canada.ca/SoilMonitoringStations/ [Last access: Apr 2020].

Haile, M., 2005. Weather patterns, food security and humanitarian response in sub-Saharan Africa. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 360, 2169–2182.

Hao, Z., AghaKouchak, A., Nakhjiri, N., Farahmand, A., 2014. Global integrated drought monitoring and prediction system. Sci. Data 1, 140001.

Heim, R., Brewer, M., 2012. The global drought monitor portal: the foundation for a global drought information system. Earth Interact. 16.

Hill, H., Pittman, J., 2013. Agriculture and Disaster Risk Reduction. In: Background Paper prepared for the Global Assessment Report on Disaster Risk Reduction 2013. UNISDR. Geneva, Switzerland, URL. https://www.preventionweb.net/english/hyogo/gar/2013/en/bgdocs/Hill%20and%20Pittman,%202012.pdf.

Huffman, G., Bolvin, D., Nelkin, E., Wolff, D., 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8, 38–55.

Hunt, E., Hubbard, K., Wilhite, D., Arkebauer, T., Dutcher, A., 2009. The development and evaluation of a soil moisture index. Int. J. Climatol. 29, 747–759.

JPL, 2020. Why it matters. URL. https://smap.jpl.nasa.gov/mission/why-it-matters/ [Last access: Mar 2020].

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D.N., Wood, E., 2017a. Four decades of microwave satellite soil moisture observations: part 1. A review of retrieval algorithms. Adv. Water Resour. 109, 106–120.

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D.N., Wood, E., 2017b. Four Decades of Microwave Satellite Soil Moisture Observations: Part 2. (Product validation and inter-satellite comparisons. Advances in Water Resources).

Kerr, Y.H., Njoku, E.G., 1990. A semiempirical model for interpreting microwave emission from semiarid land surfaces as seen from space. IEEE Trans. Geosci. Remote Sens. 28, 384–393.

Kerr, Y., Waldteufel, P., Wigneron, J.P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M., Font, J., Reul, N., Gruhier, C., 2010. The SMOS MISSION: new tool for monitoring key elements of the Global Water Cycle. IEEE 98, 666–687.

Kerr, Y.H., Al-Yaari, A., Rodriguez-Fernandez, N., Parrens, M., Molero, B., Leroux, D., Bircher, S., Mahmoodi, A., Mialon, A., Richaume, P., Delwart, S., Pellarin, A.A.B.T., Bindlish, R., Rudiger, T.J.C., Waldteufel, P., Mecklenburg, S., Wigneron, J., 2016. Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation. Remote Sens. Environ. 180, 40–63. https://doi.org/10.1016/j.rse.2016.02.042.

Knutson, M.J.H.O.V.W.C.L, 2004. Reducing drought risk: bridging theory and practice. Nat. Haz. Rev. 106–113.

Kogan, F., Sullivan, J., 1993. Development of global droughtwatch system using NOAA/AVHRR data. Adv. Space Res. 13, 219–222.

Koster, R., Guo, Z., Yang, R., Dirmeyer, P., Mitchell, K., Pum, M., 2009. On the nature of soil moisture in land surface models. J. Clim. 22, 4322–4335.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res. 99, 415–427.

Liang, X., Wood, E.F., Lettenmaier, D.P., 1996. Surface soil moisture parameterization of the Vic-2l model: evaluation and modifications. Glob. Planet. Chang. 13, 195–206.

Lintner, A.B.B., Findell, K., Malyshev, S., Loikith, P., Gentine, P., 2014. Impact of soil moisture–atmosphere interactions on surface temperature distribution. J. Clim. 27, 7976–7993.

Liu, Q., Reichle, R.H., Bindlish, R., Cosh, M.H., Crow, W.T., de Jeu, R., Lannoy, G.J.M.D., Huffman, G.J., Jackson, T.J., 2011. The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. J. Hydrometeorol. 12, 750–765.

Liu, Y., Dorigo, W., Parinussa, R., de Jeu, R., Wagner, W., McCabe, M., Evans, J., van Dijk, A., 2012. Trend-preserving blending of passive and active microwave soil moisture retrievals. Remote Sens. Environ. 123, 280–297.

Lloyd-Hughes, B., Saunders, M.A., 2007. University college London global drought monitor. URL. http://web.archive.org/web/20070210122816/http://drought.mssl.ucl.ac.uk/intro.html.

Manzoni, S., Vico, G., Palmroth, S., Porporato, A., Katul, G., 2013. Optimization of stomatal conductance for maximum carbon gain under dynamic soil moisture. Adv. Water Resour. 62, 90–105.

McKee, T., Doesken, N., Kleist, J., 1993. The relationship of drought frequency and duration of time scales. In: Proceedings of the Eighth Conference on Applied Climatolog, Anaheim, CA, USA, pp. 179–186 17–23 January.

Molero, B., Leroux, D., Richaume, P., Kerr, Y.H., Merlin, O., Cosh, M., Bindlish, R., 2018. Multi-timescale analysis of the spatial representativeness of in situ soil moisture data within satellite footprints. J. Geophys. Res.-Atmos. 123, 3–21.

Narasimhan, B., Srinivasan, R., 2005. Development and evaluation of soil moisture deficit index (SMDI) and evapotranspiration deficit index (ETDI) for agricultural drought monitoring. Agric. For. Meteorol. 133, 69–88.

Nijssen, B., O'Donnell, G., Lettenmaier, D., Lohmann, D., Wood, E., 2001. Predicting the discharge of global rivers. J. Clim. 14, 3307–3323.

Nijssen, B., Shukla, S., Lin, C., Gao, H., Zhou, T., Sheffield, J., Wood, E.F., Lettenmaier, D.P., 2013. A prototype global drought information system based on multiple land surface models. J. Hydrometeorol. 15, 1661–1676. https://doi.org/10.1175/JHM-D-13-090.1.

NOAA, 2017. Drought- August 2017. URL. https://www.ncdc.noaa.gov/sotc/drought/201708.

Ojo, E., Bullock, P., L'Heureux, J., Powers, J., McNairn, H., Pacheco, A., 2015. Calibration and evaluation of a frequency domain reflectometry sensor for real-time soil moisture monitoring. Vadose Zone J. 14.

O'Neill, P., Chan, S., Njoku, E.G., Jackson, T., Bindlish, R., 2016. SMAP L2 Radiometer Half-Orbit 36 km EASE-Grid Soil Moisture. In: Distributed Active Archive Center Version 4. Boulder, Colo, NASA National Snow and Ice Data Center. https://doi.org/10.5067/XPJTJT812XFY.

Palmer, W., 1965. Meteorological drought. In: NOAA library and information services division. Bureau, U.S.W, Washington, DC, USA.

Pan, M., Li, H., Wood, E., 2010. Assessing the skill of satellite-based precipitation estimates in hydrologic applications. Water Resour. Res 46. https://doi.org/10.1029/2009WR008290.

Pozzi, W., 2013. Towards global drought early warning capability: expanding international cooperation for the development of a framework for global drought monitoring and forecasting. Bull. Amer. Meteorol. Soc. 94, 776–785.

Reichle, R.H., 2017. Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements. J. Hydrometeorol. 18, 2621–2645.

Rodríguez-Fernández, N.J., Sabater, J.M., Richaume, P., de Rosnay, P., Kerr, Y.H., Albergel, C., Drusch, M., Mecklenburg, S., 2017. SMOS near-real-time soil moisture product: processor overview and first validation results. Hydrol. Earth Syst. Sci. 21, 5201–5216.

Ryu, D., Famiglietti, J.S., 2005. Characterization of footprint-scale surface soil moisture variabilityusing gaussian and beta distribution functions during thesouthern great plains 1997 (sgp97) hydrology experiment. Water Resour. Res. 41, 1–13.

Sadri, S., Wood, E.F., Pan, M., 2018. Developing a drought-monitoring index for the contiguous US using SMAP. Hydrol. Earth Syst. Sci. 22, 6611–6626.

Sánchez, N., González-Zamora, Ángel, Piles, M., Martínez-Fernández, J., 2016. A new soil moisture agricultural drought index (SMADI) integrating MODIS and SMOS products: a case of study over the Iberian Peninsula. Remote Sens. 8.

Sauchyn, D., Kulshreshtha, S., 2008. Prairies: From Impacts to Adaptation: Canada in a Changing Climate 2007. Technical Report. Government of Canada, Ottawa, ON (275-32).

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., Rudolf, B., 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theor. Appl. Climatol. 115, 15–40.

Seneviratne, S.I., Nicholls, N., 2018. Changes in Climate Extremes and their Impacts on the Natural Physical Environment. Report 3. Intergovernmental Panel on Climate Change (IPCC) URL. https://www.ipcc.ch/site/assets/uploads/2018/03/SREX-Chap3_FINAL-1.pdf.

Seneviratne, S., Corti, T., Jaeger, E.D.M.H.E., Orlowsky, I.L.B., Teuling, A., 2010. Investigating soil moisture-climate interactions in a changing climate: a review. Earth-Sci. Rev. 99, 125–161.

Sheffield, J., Wood, E.F., 2007. Characteristics of global and regional drought, 1950-2000: analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. J. Geophys. Res. 112, 1984–2012.

Sheffield, J., Wood, E.F., 2008a. Global trends and variability in soil moisture and drought characteristics, 1950–2000, from observation-driven simulations of the terrestrial hydrologic cycle. J. Clim. 21, 432–458.

Sheffield, J., Wood, E.F., 2008b. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. Clim. Dyn. 31, 79–105.

Sheffield, J., Wood, E.F., 2011. Drought: Past Problems and Future Scenarios. EarthScan, New York.

Sheffield, J., Goteti, G., Wen, F., Wood, E.F., 2004. A simulated soil moisture based drought analysis for the United States. J. Geophys. Res. 109, D24108. https://doi.org/10.1029/2004JD005182.

Sheffield, J., Wood, E.F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., 2014. A drought monitoring and forecasting system for sub-Sahara African water resources and food security. Bull. Am. Met. Soc. 95, 861–882.

Sorooshian, S., Hsu, K.L., Gao, X., Gupta, H., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. Bull. Am. Meteorol. Soc. 81, 2035–2046.

Svoboda, M., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., Stephens, S., 2002. The drought monitor. Bull. Am.

Meteorol. Soc. 83, 1181–1192.

Tetlock, E., Toth, B., Berg, A., Rowlandson, T., Ambadan, J., 2019. An 11-year (2007–2017) soil moisture and precipitation dataset from the Kenaston network in the Brightwater Creek basin, Saskatchewan, Canada. Earth Syst. Sci. Data 11, 787–796.

TUWIEN, 2019. International Soil Moisture Network. URL. https://ismn.geo.tuwien.ac.at/en/.

UN/ISDR, 2007. Drought Risk Reduction Framework and Practices: Contributing to the Implementation of the Hyogo Framework for Action. United Nations Secretariat of the International Strategy for Disaster Reduction (UN/ISDR). Technical Report 98 + vi pp. United Nations Secretariat of the International Strategy for Disaster Reduction (UN/ISDR), Geneva, Switzerland.

Wagner, W., Scipal, K., Pathe, C., Lucht, D.G.W., Rudolf, B., 2003. Evaluation of the agreement between the first global remotely sensed soil moisture data with model and precipitation data. J. Geophys. Res. Atmos. 108.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., Steinnocher, K., 2013. The ASCAT soil moisture product: a review of its specifications, validation results, and emerging applications. Meteorol. Z. 22, 5–33. https://

doi.org/10.1127/0941-2948/2013/0399.

Wanders, N., Karssenberg, D., Roo, A., Jong, S.D., Bierkens, M., 2014. The suitability of remotely sensed soil moisture for improving operational flood forecasting. Hydrol. Earth Syst. Sci. 18, 2343–2357.

Wanders, N., Pan, M., Wood, E., 2015. Correction of real-time satellite precipitation with multi-sensor satellite observations of land surface variables. Remote Sens. Environ. 160, 206–221. https://doi.org/10.1016/j.rse.2015.01.016.

Wheaton, E., Wittrock, V., Kulshreshtha, S., Koshida, G., Chipanshi, A., Bonsal, B., 2005. Lessons Learned from the Canadian Drought Years of 2001 and 2002: Synthesis Report for Agriculture and Agri-Food Canada. Technical Report SRC Publication No. 11602-46E03. Saskatchewan Research Council. Saskatoon, Saskatoon.

Wilhite, D.A., 2000. Drought as a Natural Hazard: Concepts and Definitions, in Drought: A Global Assessment. Routledge, London, pp. 3–18.

Wilhite, D.A., Svoboda, M.D., Hayes, M.J., 2007. Understanding the complex impacts of drought: a key to enhancing drought mitigation and preparedness. Water Resour. Manag. 21, 763–774.

Zhan, W., Pan, M., Wanders, N., Wood, E., 2015. Correction of real-time satellite precipitation with satellite soil moisture observations. Hydrol. Earth Syst. Sci. 19, 4275–4291. https://doi.org/10.5194/hess-19-4275-2015.