Toward improved regional hydrological model performance using a novel soil data-informed calibration method

Chuxuan Li[1], Guo Yu[2], Jiali Wang[3], and Daniel E Horton[1]

[1]Northwestern University
[2]Desert Research Institute
[3]Argonne National Laboratory (DOE)

January 11, 2023

Abstract

Accurate soil moisture and streamflow data are an aspirational need of many hydrologically-relevant fields. Model simulated soil moisture and streamflow hold promise but numerical models require calibration prior to application to ensure sufficient model performance. Manual or automated calibration methods require iterative model runs and hence are computationally expensive. In this study, we leverage the Soil Survey Geographic (SSURGO) database and the probability mapping of SSURGO (POLARIS) to help constrain soil parameter uncertainties in the Weather Research and Forecasting Hydrological modeling system (WRF-Hydro) over a central California domain. After calibration, WRF-Hydro soil moisture exhibits increased correlation coefficients (r), reduced biases, and increased Kling-Gupta Efficiencies (KGEs) across seven in-situ soil moisture observing stations. Compared to four well-established soil moisture datasets including Soil Moisture Active Passive Level 4 data and three Phase 2 North American Land Data Assimilation System land surface models, our POLARIS-calibrated WRF-Hydro produces the highest mean KGE (0.67) across the seven stations. More importantly, WRF-Hydro streamflow fidelity also increases especially in the case where the model domain is set up with an SSURGO-informed total soil thickness. Both the magnitude and timing of peak flow events are better captured, r increases across nine United States Geological Survey stream gages, and the mean Nash-Sutcliffe Efficiency across seven of the nine gages increases from 0.19 in default WRF-Hydro to 0.63 after calibration. Our soil data-informed calibration approach, which is transferable to other spatially-distributed hydrological models, uses open-access data and non-iterative steps to improve model performance and is thus operationally and computationally attractive.

Hosted file

953402_0_art_file_10582675_rnxm1l.docx available at https://authorea.com/users/559535/articles/618234-toward-improved-regional-hydrological-model-performance-using-a-novel-soil-data-informed-calibration-method

Hosted file

953402_0_supp_10582588_rnxzxk.docx available at https://authorea.com/users/559535/articles/618234-toward-improved-regional-hydrological-model-performance-using-a-novel-soil-data-informed-calibration-method

Toward improved regional hydrological model performance using a novel soil data-informed calibration method

Chuxuan Li[1], Guo Yu[2], Jiali Wang[3], and Daniel E. Horton[1]

1. Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL 60208, USA
2. Division of Hydrologic Sciences, Desert Research Institute, Las Vegas, NV, 89119, USA
3. Environmental Science Division, Argonne National Laboratory, Lemont, IL, 60439, USA

Key points:

- Model simulated soil moisture and streamflow fidelity substantially improve by using computationally-efficient soil data-based calibration.
- Calibrated surface soil moisture outperforms four well-established soil moisture products when evaluated against in situ observations.
- After calibration, the model's capability to simulate observed streamflow hydrographs improves, especially peak flow fidelity.

Abstract:

Accurate soil moisture and streamflow data are an aspirational need of many hydrologically-relevant fields. Model simulated soil moisture and streamflow hold promise but numerical models require calibration prior to application to ensure sufficient model performance. Manual or automated calibration methods require iterative model runs and hence are computationally expensive. In this study, we leverage the Soil Survey Geographic (SSURGO) database and the probability mapping of SSURGO (POLARIS) to help constrain soil parameter uncertainties in the Weather Research and Forecasting Hydrological modeling system (WRF-Hydro) over a central California domain. After calibration, WRF-Hydro soil moisture exhibits increased correlation coefficients (*r*), reduced biases, and increased Kling-Gupta Efficiencies (KGEs) across seven in-situ soil moisture observing stations. Compared to four well-established soil moisture datasets including Soil Moisture Active Passive Level 4 data and three Phase 2 North American Land Data Assimilation System land surface models, our POLARIS-calibrated WRF-Hydro produces the highest mean KGE (0.67) across the seven stations. More importantly, WRF-Hydro streamflow fidelity also increases especially in the case where the model domain is set up with an SSURGO-informed total soil thickness. Both the magnitude and timing of peak flow events are better captured, *r* increases across nine United States Geological Survey stream gages, and the mean Nash-

Sutcliffe Efficiency across seven of the nine gages increases from 0.19 in default WRF-Hydro to 0.63 after calibration. Our soil data-informed calibration approach, which is transferable to other spatially-distributed hydrological models, uses open-access data and non-iterative steps to improve model performance and is thus operationally and computationally attractive.

Key words: soil moisture, streamflow, data-informed calibration, hydrological models

Plain language summary

In this study, we develop a method that uses field- and machine learning-derived soil property uncertainties to improve the performance of a hydrological model to simulate observed soil water content and river flows. Specifically, we replace three of the model's default parameters with the corresponding parameters from a probabilistic soil property dataset. After replacement, simulated soil water content more closely resembles observations from seven in-situ observing stations. Compared to four other well-established, satellite-derived and model-simulated products, our soil property-calibrated model performs favorably. For river flows, we find the highest model performance in the case where we modify the total soil thickness according to the soil survey dataset. With modified soil thickness, the timing and magnitude of high flows are much better captured and the similarity between our simulations and the observations substantially increases at almost all observing stations. Compared to calibration methods that require repetitive model runs, our probabilistic soil property calibration method is computationally-efficient and may prove useful in a number of hydrologic modeling contexts.

1. Introduction

Soil moisture and streamflow are two key components of the hydrologic cycle. In the following, we provide examples to show their importance for a plethora of fields including hydrology, geomorphology, natural hazards, ecology, water resource management, and climate science. For natural hazards and geomorphology, both soil moisture and streamflow can influence the likelihood of flooding (Koster et al. 2010; Massari et al. 2014) and debris flows (Coe et al. 2008; Kean et al. 2013; Tang et al. 2019), while soil moisture has also been used to predict drought (Xu et al. 2020) and shallow landslides (Gasmo et al. 2000; Handwerger et al. 2019; Johnson and Sitar 1990; Ray and Jacobs 2007; Sweeney and Robertson 1979). In water supply management, soil moisture influences forest water yield and streamflow controls suspended sediment transport and water quality (Acharya et al. 2022; Colby 1956). Water in soil and river channels also drives the productivity and sustainability of terrestrial ecosystems, especially in arid and semi-arid regions (Legates et al. 2011), influencing crop yields and other aspects of agriculture (Berg and Sheffield 2018; Carrão et al. 2016; Kang et al. 2009). Over climatic timescales, soil moisture affects both short- and long-term climate by modulating the hydro-climate feedback loop (Seneviratne et al. 2010; Seneviratne et al. 2013; Yeh et al. 1984). Because soil moisture and streamflow play important roles in the broad Earth system across various spatiotemporal scales, accurate estimates of them are critical to improve the predictive skills of models in a wide range of fields. For example, initializing models with realistic soil moisture can reduce uncertainties in atmospheric predictions at sub-seasonal to seasonal scales in climate models (Douville and Chauvin 2000; Fennessy and Shukla 1999; Koster 2004) and facilitate accurate landslide predictions in slope stability models (Cai et al. 2019; Di Matteo et al. 2018). On climatic timescales, soil moisture can greatly impact projections of extreme temperature and precipitation in global climate models (Seneviratne et al. 2013). In ecological and agricultural models, soil moisture is needed to simulate carbon cycles (Friend and Kiang 2005; Yuste et al. 2007) and crop growth (Rosenzweig et al. 2002) and is a key variable for predicting agricultural drought (Crow et al. 2012; Narasimhan and Srinivasan 2005). Streamflow is also an indispensable variable used in hydrological hazard mapping and assessment tools, water resource management tools, landscape evolution models, and coupled atmospheric-hydrological models (Davy and Lague 2009; Dottori et al. 2016; Gong et al. 2010; Wagner et al. 2016).

In-situ observations of soil moisture and streamflow are regarded as ground truth. However, they are spatially sparse due to the high costs of large-scale implementation especially in remote and topographically complex regions. This is especially a problem for obtaining in-situ soil moisture observations. Satellites using passive microwave techniques such as Soil Moisture Active Passive (SMAP), on the other hand, provide promising remotely-sensed surface soil moisture with global data coverage (Al-Yaari et al. 2017; Chen et al. 2018; Kumar et al. 2018). However, satellite-derived data is reported to be biased in heavily vegetated areas (Fan et al. 2020; Ma et al. 2019; Reichle et al. 2017) and is subject to data gaps primarily due to satellite orbits (Tavakol et al. 2019; Wang et al. 2012). In addition, remote sensing techniques can only retrieve skin (0–5 cm) or near-surface soil moisture (Mohanty et al. 2017). As such, process-based land surface models (LSMs) are frequently used to fill the data gaps in satellite-derived soil moisture and extend soil moisture estimates to the root zone (~1–2 m below ground) (Koster et al. 2009; Mohanty et al. 2017; Tavakol et al. 2019). However, LSMs at global or regional scales often have rather coarse resolutions (e.g., 1/8 degree in NLDAS-2 LSMs). Due to the high variability in soil moisture across space and time, efforts to produce high-resolution soil moisture are needed for both regional-scale and locally-focused applications.

Physics-based hydrological models that simulate soil moisture and streamflow at high resolutions are critical tools to fill in-situ and remotely-sensed gaps but models need validation prior to application. Improving hydrological models' soil moisture and streamflow performance has been a long-standing research objective. In these models, soil moisture and streamflow are prognostic variables that are often subject to great uncertainties originating from various sources including model physics and structure, meteorological forcing, and parameterizations (Leach et al. 2018; Matgen et al. 2010; Silver et al. 2017). To improve simulation fidelity, a number of different techniques have been employed including data assimilation and manual or automated calibration. So far, data assimilation has been the primary technique to improve soil moisture simulations in hydrological models and it has shown promising results by incorporating remotely-sensed soil moisture data (Crow and Van den Berg 2010; De Santis et al. 2021; Loizu et al. 2018). It is also found that assimilating observational soil moisture can improve the accuracy of both soil moisture and streamflow predictions in various types of models (Aubert et al. 2003; Lee et al. 2011). Assimilating observational streamflow and/or snow data is also applied to improve streamflow simulations (Lahmers et al. 2022) and forecasts (Boucher et al. 2020). Despite its successful applications, data assimilation typically requires a large volume of high-quality observational data which are often not available in data-scarce regions. Other efforts to improve model predictions include model calibration. Typically, hydrological models are calibrated either manually (i.e., via a trial-and-error process (Yucel et al. 2015)) or using an automated algorithm (Becker et al. 2019; Gallagher et al. 2007). Both manual and automated-algorithm-based calibration techniques require iterative model runs to arrive at the optimal combination of parameters. Even though parallelization has saved considerable computing hours (Alvioli et al. 2016; Baum et al. 2008; Wang et al. 2019), this calibration process could still be complicated and resource-demanding, especially when the model domain is large and spatial resolution is high.

Here, we develop a soil property data-informed calibration method to calibrate both soil moisture and streamflow simulations with non-iterative steps in the Weather Research and Forecasting Hydrological modeling system version 5.1.1 (WRF-Hydro; Gochis et al. 2020). WRF-Hydro is a 3-D, fully-distributed, and physics-based open-source community hydrological model. Compared with other traditional hydrological models such as the semi-distributed Variable Infiltration Capacity model (VIC), WRF-Hydro is a fully-distributed model that considers spatially distributed hydrological variables (Yin et al. 2020); compared with the quasi-physically-based Soil and Water Assessment Tool (SWAT) that works at single watershed- to river basin-scales, WRF-Hydro can simulate multi-processes across multiple scales. In operational mode, WRF-Hydro works as the hydrologic core of National Water Model (NWM) to produce streamflow predictions at ~2.7 million river reaches. In research settings, streamflow from WRF-Hydro has been calibrated manually (Yucel et al. 2015) or using automatic algorithms (Lahmers et al. 2020; Lahmers et al. 2019; Wang et al. 2019; Yu et al. 2020). Sofokleous et al. (2022) found streamflow predictions in WRF-Hydro are improved with improved representation of groundwater and transpiration processes, which highlights the importance of replicating the real-world conditions with realistic parameters as opposed to intensively calibrating the parameters. In that respect, in contrast to manual and auto-algorithm-based calibration studies, Silver et al. (2017) outlined a systematic calibration procedure that employs physical soil characteristics derived from remote sensing to calibrate streamflow in WRF-Hydro. However, similar calibration methods have not been applied to soil moisture in WRF-Hydro and the use of soil moisture from WRF-Hydro simulations has thus far been limited. In addition, it is not clear whether the improved soil moisture can improve streamflow simulation as well in WRF-Hydro. Therefore, to simplify the calibration procedure and increase the utility of high-resolution soil moisture simulation in WRF-Hydro, here we develop a calibration approach that relies on two related open-access soil databases i.e., SSURGO and Probability Mapping of Soil Survey Geographic Database (POLARIS; Chaney et al. 2016). Our calibrated experiments show improved simulation-observation fidelity for both soil moisture and streamflow. With the improvement, our approach may increase the utility of WRF-Hydro and potentially other spatially-distributed hydrological models for a number of hydrologically relevant fields, including climate science, natural hazards, agriculture, and ecology. In the following, our study domain and environmental setting are introduced in Section 2, descriptions of the model, data, and the data-informed calibration method are presented in Section 3, Section 4 presents the results, and Section 5 provides discussions and a conclusion.

2. Study area and environmental setting

Our study area is located in the Coast Ranges surrounding Monterey Bay in central California, USA (Fig. 1a). The WRF-Hydro model domain outlined by the black box in Fig. 1a covers several mountainous areas, seven in-situ soil moisture stations, and nine United States Geological Survey (USGS) stream gages. Soil moisture stations Los Gatos (lgs), Gilroy (gry), Soledad (sld), and Lockwood (lwd) are operated by the NOAA Physical Sciences Laboratory (PSL), whereas stations blueoak, norris, and hastings are operated by the Western Regional Climate Center (WRCC). The streamflow measured at the nine USGS stream gages are natural flows (i.e., flows without human regulations). The details regarding the soil moisture and streamflow observational sites are given in Section 3.4.1. California has a Mediterranean climate with distinct wet and dry seasons. About 80% of annual precipitation in California falls within the wet season [defined as November to April in Jong et al. (2016)]. Due to the Mediterranean climate, soil moisture in California also has high seasonal variability, similar to precipitation.

Our model domain features complex topography and heterogeneous vegetation cover (Fig.1b-d). The histograms of elevation and slope are calculated based on USGS National Elevation Dataset (NED) 30-m

Digital Elevation Model (DEM). Both distributions have a bimodal shape i.e., the majority of the model domain has topographic elevations 30-40 m above sea level and minimal slopes, and the secondary peaks in the distributions, however, correspond to topographic elevations of 300 m and slopes of 13°. The interquartile range of topographic slope spans more than 15°, showing the large spatial heterogeneity in topographic gradients. The distribution of the Moderate Resolution Imaging Spectroradiometer (MODIS) normalized difference vegetation index (NDVI) has a median value of ~0.6 and a maximum value approaching 1. According to the MODIS International Geosphere–Biosphere Programme (IGBP) land cover data, evergreen needleleaf forest is the most dominant vegetation cover in our model domain (Supplemental Fig. 1).

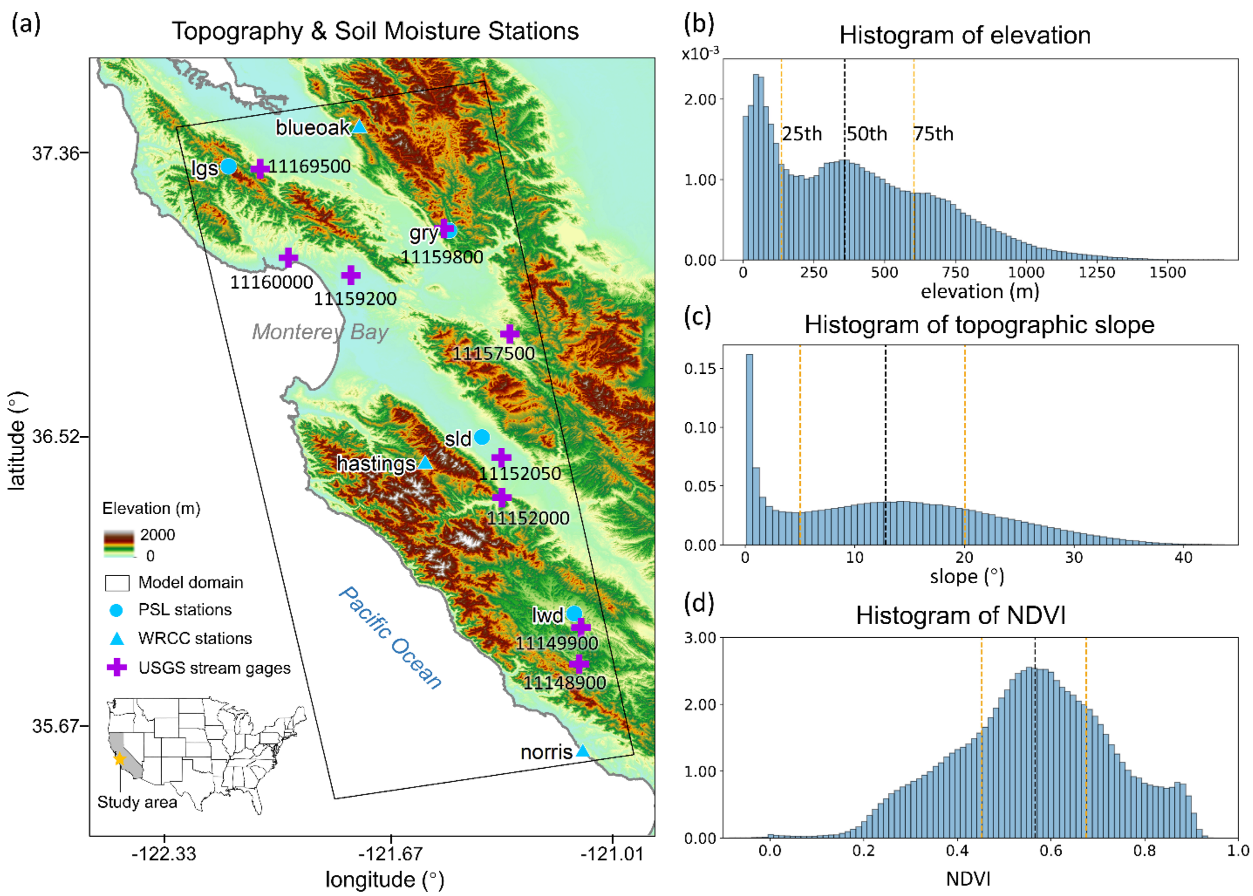

Fig. 1| WRF-Hydro model domain, topography, soil moisture observational sites, USGS stream gages, and statistics of the environmental setting. (a) The model domain covers several mountains in the Coast Ranges of central California (black box). Topography is from the USGS National Elevation Dataset (NED) 30-m DEM (shading). There are seven in-situ soil moisture stations (blue circles for NOAA PSL stations and blue triangles for WRCC stations) and nine USGS stream gages that measure natural flows (purple crosses). The location of the study area in the U.S. is shown in the embedded map with the state of California shaded in grey. Distributions of (b) topographic elevation, (c) topographic slope, and (d) normalized difference vegetation index (NDVI) within the model domain. Median values of the distributions are indicated by the black vertical dashed lines and 25th and 75th percentiles are indicated by the orange vertical dashed lines. The distributions of elevation and slope are calculated using the USGS 30-m DEM, and the distribution of NDVI is calculated based on the Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) Vegetation Indices (MYD13Q1) Version 6.1 data.

3. Data and Methods

3.1 WRF-Hydro model description and configurations

WRF-Hydro is a physics-based, open-source community model that simulates 3-D land surface hydrologic processes (Gochis et al. 2020). WRF-Hydro includes the Noah-MP Land Surface Model (LSM) (Niu et al. 2011), a terrain routing module, a channel and reservoir routing module, and a conceptual baseflow bucket model. The Noah-MP LSM simulates vertical energy fluxes (i.e., sensible and latent heat and net radiation), moisture fluxes (i.e., infiltration, infiltration excess, canopy interception, and evapotranspiration), and soil thermal and moisture state variables. In default configuration, the soil column in Noah-MP LSM has a total depth of 2 m and four soil layers. The thickness of the layers from top to bottom is 10, 30, 60, and 100 cm, respectively. For each of the four soil layers, the simulation of water movement follows the diffusive form of Richard's equation. Users can modify the total depth and thickness of each layer but in the current version of WRF-Hydro the total soil depth and vertical distribution of soil layers can only be the same across the model domain.

Soil moisture and other variables are disaggregated from the relatively coarse grid in Noah-MP LSM (1-km in our study) to the higher resolution grid in the terrain routing module (100-m in our study) which then simulates subsurface and overland flow. The high-resolution terrain routing grid is generated by interpolating the USGS NED 30-m hydrologically-conditioned DEM to our 100-m grid. Once the overland flow and subsurface flow simulated from the terrain routing module flow into the channel grid that is pre-defined in the USGS hydrologically-conditioned DEM, the channel routing module of WRF-Hydro routes the water as channelized streamflow. The channel routing module works at a spatial resolution consistent with the channel bottom width which typically ranges from 1.5 m to 100 m. More details regarding the governing equations and model workflows can be found in Li et al. (2022).

In this study, WRF-Hydro is run in standalone mode, i.e., it is not coupled with an atmospheric model. We use the Multi-Radar/Multi-Sensor System (MRMS) gauge-corrected quantitative precipitation estimation (QPE; Zhang et al., 2011, 2014, 2016) to provide precipitation forcing at hourly, 1-km resolution and the Phase 2 of North American Land Data Assimilation System (NLDAS-2) to provide forcing of other meteorological variables including incoming shortwave and longwave radiation, specific humidity and air temperature at 2 m above the surface, surface pressure, and 10-m wind speed (both u and v components) at hourly, 1/8-degree resolution. The MRMS precipitation and NLDAS-2 forcing data are re-gridded onto the 1-km Noah-MP LSM grid using bilinear interpolation.

WRF-Hydro is initialized with National Centers for Environmental Prediction (NCEP) FNL (Final) Operational Global Analysis data. We spin up the model for one year from October 1, 2015 – September 31, 2016. The one-year spin-up time allows the hydrological variables in the model to reach equilibrium. We run WRF-Hydro in three configurations: one in its default configuration and two calibrated experiments. Details of the calibration experiments are given in Section 3.3.2. Soil moisture is reported hourly on the terrain routing grid (100-m) under three configurations for October 1, 2016 to May 31, 2017.

3.2 Soil hydraulic properties in default WRF-Hydro

Prior to calibration, we performed sensitivity experiments to identify highly-sensitive soil moisture-relevant parameters in WRF-Hydro (Supplemental Fig. 2). Our sensitivity experiments covered numerous soil property and vegetation parameters including *smcmax* (soil porosity), *dksat* (saturated hydraulic conductivity), *bexp* (coefficient *b* in Cosby et al. (1984) that denotes pore size distribution), *smcref* (field capacity), *smcwlt* (wilting point), *slope* (bottom soil layer drainage), *rsurfexp* (surface dryness factor controlling the surface resistance for evaporation), *hvt* (canopy height), and *vcmx25* (maximum carboxylation rate at 25°C). The sensitivity analyses were performed by manually changing the parameter values within a physically-reasonable range based on POLARIS and comparing changes in simulated soil moisture time series. Eventually *smcmax*, *dksat,* and *bexp* were identified as the three most sensitive parameters and their effects on soil moisture simulations are shown in Supplemental Fig. 2.

In the default version of the Noah-MP LSM and WRF-Hydro, *smcmax*, *dksat*, and *bexp* are mapped onto the 16 soil classes defined in the 1-km USDA State Soil Geographic database (STATSGO; Miller and White, 1998) based on the soil analysis from Cosby et al. (1984) (Fig. 2a–c and Supplemental Table 1). Specifically, Holtan et al. (1968) and Rawls et al. (1976) collected 1448 soil samples from 35 locations across 23 states in the U.S. Using these soil samples, Cosby et al. (1984) derived the representative values of soil saturated hydraulic conductivity and porosity for each soil class, whereas the *bexp* (i.e., *b* in the equation below) was calculated via a best fit to the moisture retention data. This soil analysis conducted by Cosby et al. (1984) is used as the default soil hydraulic properties in WRF-Hydro (Supplemental Table 1). The 16-type STATSGO soil map has a relatively coarse spatial resolution and its accuracy was found to be questionable (Dy and Fung, 2016).

3.3 A new soil data-informed calibration method

3.3.1 SSURGO and POLARIS soil databases

To better constrain the uncertainties in the soil parameters of WRF-Hydro, we leverage two related open access soil databases, i.e., the Soil Survey Geographic (SSURGO) database and the probability mapping of SSURGO (POLARIS; Chaney et al. 2016). Here we provide a brief description of both.

SSURGO is a compilation of soil surveys with details gathered over the course of a century for the CONUS (Soil Survey Staff, 2021). It was generated via a combination of observed soil information in the field, lab experiments, expert knowledge, areal images, pedotransfer functions, and extrapolation of observations using soil and/or landscape models. It is managed and updated annually by the National Cooperative Soil Survey. In terms of data format, SSURGO provides a map of polygon features with assigned unique map units and tabular soil texture and property information. Each map unit corresponds to multiple soil components and each component corresponds to multiple soil horizons. Though it has the highest level of details and it is the most up-to-date soil physical property data, it is subject to data gaps and artificial discontinuities between political units that conduct the soil survey (i.e., county or state boundaries).

To fill the data gaps, remove the artificial discontinuities, and spatially disaggregate the multiple components for one map unit in SSURGO, POLARIS probabilistically remaps SSURGO using high-resolution geospatial environmental data such as topography and land cover data with a random forest machine learning algorithm (DSMART-HPC; Chaney et al., 2016). POLARIS provides soil series predictions with uncertainties for six soil layers at 30-m resolution over CONUS. The statistics it provides include the mean, mode, 5th, 50th, and 95th percentiles and the depths of the six soil layers are 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm, and 100–200 cm, respectively.

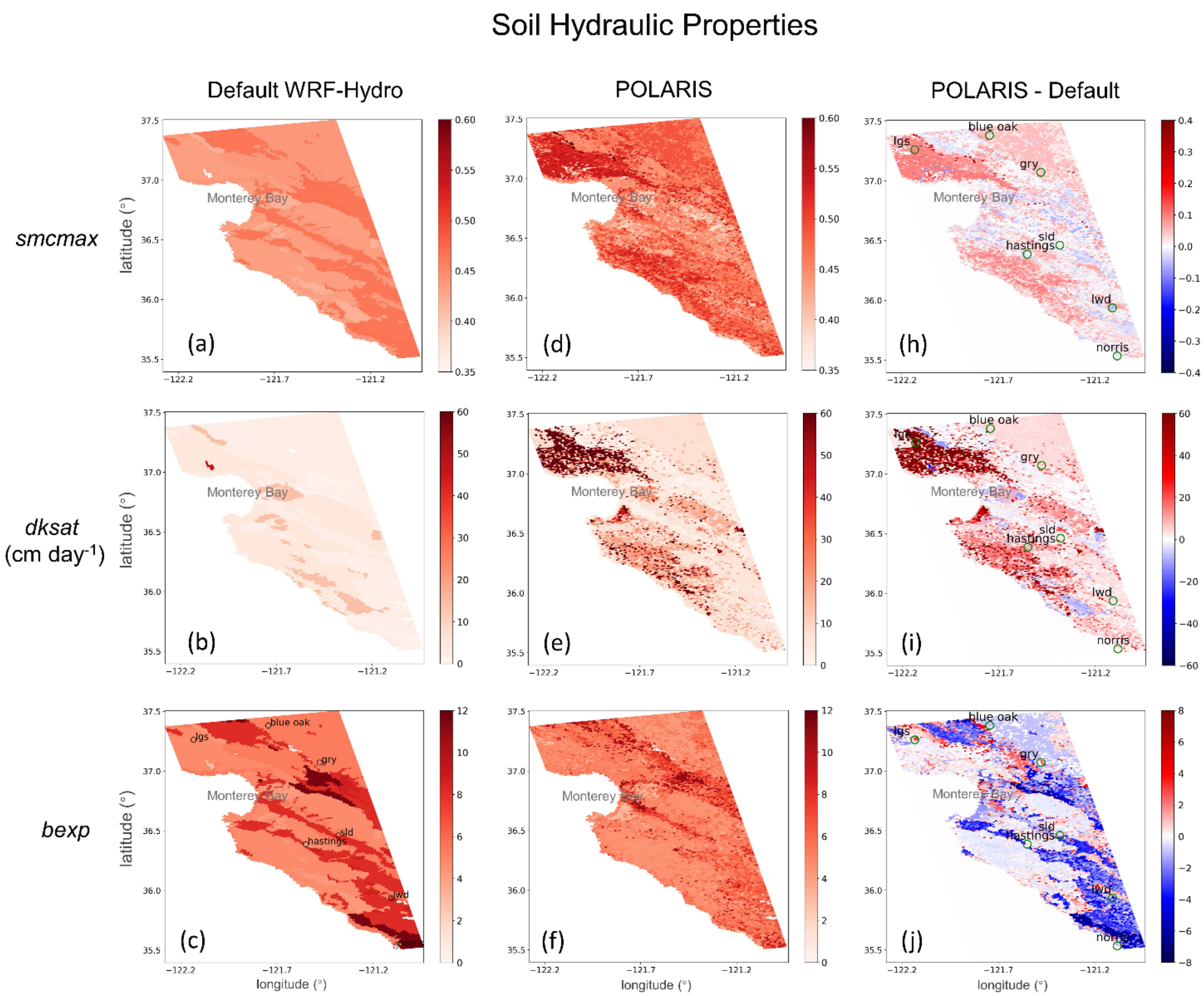

Fig. 2| Maps of soil hydraulic properties including *smcmax* (porosity), *dksat* (saturated hydraulic conductivity; cm day^{-1}), and *bexp* that controls the soil pore size distribution in (a)–(c) default WRF-Hydro, (d)–(f) POLARIS-based parameters, and (h)–(j) difference between the POLARIS-based parameters and default WRF-Hydro. Note that (d) and (e) show the median porosity and saturated hydraulic conductivity from 0–5 cm soil layer in POLARIS and *bexp* in (f) is calculated using the POLARIS 0–5 cm median clay fraction based on the linear regression model in Cosby et al. (1984). The green circles in (h) – (j) show the seven in-situ soil moisture stations.

3.3.2 Soil data-informed calibration experiments

In this study, except for the experiment using default WRF-Hydro, we perform two WRF-Hydro calibration experiments by incorporating the information from SSURGO and POLARIS. Fig. 3 shows a flowchart summarizing the information and methods used in the two experiments, i.e., 1) POLARIS-calibrated WRF-Hydro with a 2 m soil column (hereafter referred to as "POLARIS-calibrated WRF-Hydro") and 2) the POLARIS-calibrated WRF-Hydro with a modified total soil thickness of 40 cm (hereafter referred to as "POLARIS-40 cm soil"). It is worth mentioning that reducing total soil thickness does not influence surface soil moisture simulations so POLARIS-40 cm soil experiment simulates the same surface soil moisture as POLARIS-calibrated WRF-Hydro. We perform both calibration experiments starting October 1, 2016 and both calibrated and default WRF-Hydro run for eight months from October 1, 2016 – May 31, 2017.

In the POLARIS-calibrated WRF-Hydro, we use the median values in the top soil layer (0–5 cm) of the following parameters: porosity (in m^3 m^{-3}), saturated hydraulic conductivity on $\log_{10}$ scale (cm hr^{-1}), and clay fraction (in %). We use median values because they are more representative for the entire distribution. Specifically, the median POLARIS porosity and saturated hydraulic conductivity are re-gridded onto the Noah-MP LSM grid using a nearest-neighbor interpolation and are used to replace the parameters *smacmax* and *dksat* in all four soil layers in default WRF-Hydro (Fig. 2d&e). To derive *bexp* which denotes the soil pore size distribution (Fig. 2f), we re-grid the median clay fraction onto the LSM grid and apply a linear regression model adopted from Cosby et al. (1984):

$$bexp = 0.159 \times c + 2.91 \quad (1)$$

where c is clay fraction in %. Differences between the POLARIS-based and default soil parameters are displayed in Fig. 2h,i&j.

In the second calibration experiment, we set up the model domain of POLARIS-calibrated WRF-Hydro with a reduced total soil thickness of 40 cm, and each of the four soil layers has a thickness of 10 cm. 40 cm is derived via calculating the domain average of the depth-to-bedrock data from SSURGO. Based on SSURGO, the soil depth in our model domain ranges from 0–173 cm with a mean and a standard deviation of 40 and 39 cm, respectively (Supplemental Fig. 3). Therefore, the 2-m soil in the default setting of WRF-Hydro is likely overestimating the actual soil conditions in central California.

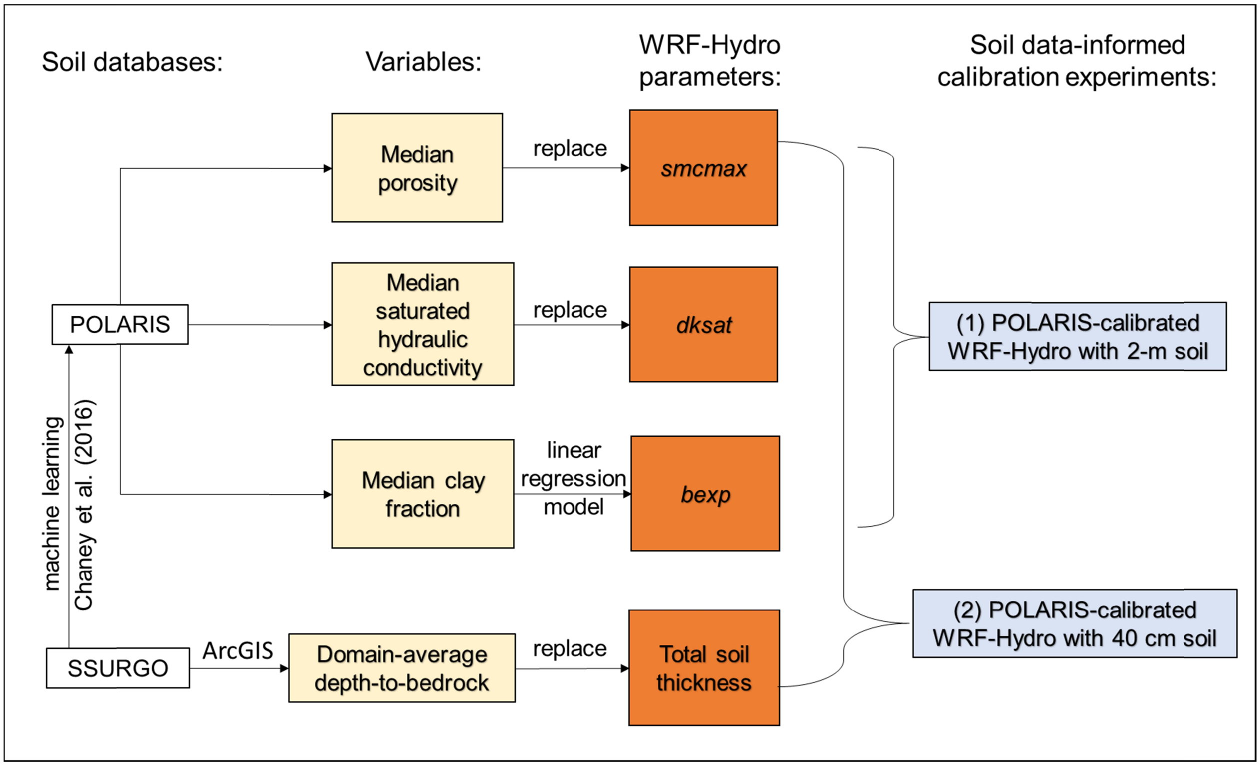

Fig. 3| A schematic flowchart summarizing the datasets, variables, model parameters, and methods used in our soil data-informed calibration experiments: (1) the POLARIS-calibrated WRF-Hydro, and (2) the POLARIS-40 cm soil according to the domain average of the SSURGO depth-to-bedrock data.

3.4 Model performance evaluation

3.4.1 Soil moisture in-situ stations and USGS streamflow gages

We then evaluate both the default and the calibrated WRF-Hydro against soil moisture and streamflow observations over October 1, 2016 – May 30, 2017 at hourly time steps. It is noteworthy that most stream gages in our study area experience no-to-low flows outside the period of December to March. Therefore, although our model evaluations span the whole 8-month period (October 2016 – May 2017), we find a similar conclusion when evaluating the streamflow just for December to March.

The soil moisture observations we use in this study is volumetric soil moisture indirectly measured at seven in-situ stations. The water content reflectometers CS616 and CS625 at the four PSL stations (sld, lgs, gry, and lwd) measure the soil temperature and output period at 10 cm below ground at 2-minute interval. We use these two variables to compute the volumetric soil moisture. We firstly perform a data correction to the output period with the measured soil temperature following the equation in the reflectometer instruction manual (Campbell Scientific INC, retrieved 2021):

$$\tau_c(T) = \tau_o + (20 - T) \times (0.526 - 0.052 * \tau_o + 0.00136 \times {\tau_o}^2) \quad (2)$$

where τ_o and τ_c are output periods in microseconds before and after the correction, respectively, and T is soil temperature in °C. The corrected output period is then converted to volumetric soil water content (m^3 m^{-3}) using a quadratic calibration equation documented in the instruction manual (Campbell Scientific INC, retrieved 2021):

$$VWC = -0.0663 - 0.0063 \times \tau_c + 0.0007 \times {\tau_c}^2 \quad (3)$$

At the three WRCC stations (blueoak, hastings, and norris), reflectometer CS615 is used to measure the soil moisture at 2 inches (~5 cm) below ground at 10-minute resolution.

To compare WRF-Hydro soil moisture simulations with the in-situ observations, we first compute the hourly mean for the soil moisture observations during October 1, 2016 – May 31, 2017. Next, time series of soil moisture simulations are collected from the WRF-Hydro high-resolution routing grid cells (100-m) that are closest to the seven observational stations. We also use the in-situ precipitation recorded by the soil moisture observational stations to investigate the uncertainties in the precipitation forcing (i.e., the MRMS).

Nine USGS stream gages with natural flows (i.e., no human regulation) are available in our model domain, as shown in in Figure 1a. They are Saratoga Creek at Saratoga (ID 11169500), Soquel Creek at Soquel (ID 11160000), WB Soquel C NR Soquel (ID 11159800), Corralitos Creek at Freedom (ID 11159200), Tres Pinos Creek near Tres Pinos (ID 11157500), Arroyo Seco NR Soledad (ID 11152000), Arroyo Seco BL Reliz C NR Soledad, CA (ID 11152050), San Antonio River near Lockwood (ID 11149900), and Nacimiento River below Sapaque Creek near Bryson (ID 11149800). The streamflow observations are at 15-minute resolution and we calculate the hourly mean of the observations to compare with our model simulations.

3.4.2 Other remotely-sensed and LSM-simulated soil moisture products

For further evaluation of the performance of WRF-Hydro simulated soil moisture, we also compare POLARIS-calibrated soil moisture with four other widely-used soil moisture products: SMAP L4, NLDAS-2 Noah, VIC, and Mosaic LSMs.

SMAP L4 is a merged soil moisture product that assimilates SMAP satellite L-band brightness temperature observations into the NASA's GEOS-5 Catchment LSM using a spatially-explicit ensemble Kalman filter (Reichle et al., 2017). Catchment LSM has a temporal resolution of 3 hours, a spatial resolution of 9 km, and provides soil moisture estimates at the surface (0–5 cm) and root zone (0–1 m). SMAP L4 is chosen because SMAP satellite-derived soil moisture has been reported to be superior to other remotely-sensed soil moisture products by various studies (Al-Yaari et al., 2017; Zhang et al., 2017; Chen et al., 2018; Kumar et al., 2018; Tavakol & Rahmani, 2018) (Ford and Quiring 2019). Compared to SMAP L1–3 products, SMAP L4 is continuous over space and time and combines both observation and simulation components.

NLDAS-2 applies state-of-the-art observational and simulated data as forcing to drive physically-based, uncoupled, distributed LSMs to simulate land surface conditions at hourly and 1/8-degree resolutions over the U.S. NLDAS-2 uses three physics-based LSMs, i.e., Noah (Betts et al., 1997; Chen et al., 1997), VIC (Liang et al., 1994), and Mosaic (Koster and Suarez, 1994, 1996). All NLDAS-2 LSMs share the same atmospheric forcing, soil classification, and land cover, but they yield different results due to different model physics, configuration, and parameter choices. Noah-LSM was developed as the land component of the mesoscale Eta model by NOAA and NCEP (Betts et al., 1997; Chen et al., 1997). It also works as the LSM of WRF atmospheric model and NOAA/NCEP Global Forecast System (GFS) and Climate Forecast System (CFS). Noah has four soil layers with depths of 10, 30, 60, and 100 cm from top to bottom. VIC-LSM is a macroscale, semi-distributed hydrologic model designed by University of Washington and Princeton University (Liang et al., 1994; Wood et al., 1997). It has three soil layers with spatially-varying layer depths depending on the vegetation type and root distribution. The Mosaic LSM was developed for use in NASA's global climate models by Koster and Suarez (1994 and 1996). It uses a tile approach to represent vegetation variability at sub-grid scales. Each vegetation tile simulates its own soil moisture and consists of three soil layers with depths of 10, 30, and 160 cm from top to bottom. The soil moisture from NLDAS-2 LSMs has been widely evaluated (Xia et al. 2015; Xia et al. 2014; Zhuo et al. 2015). In a nation-wide soil moisture product evaluation study, soil moisture simulations from NLDAS-2 LSMs are found to have the best performance among various modeled and remotely sensed soil moisture products (Ford and Quiring 2019).

We average soil moisture observations and WRF-Hydro simulations to 3-hourly to compare with the surface-layer, 3-hourly SMAP L4 soil moisture. For the three NLDAS-2 LSMs, we use the hourly soil moisture product in the surface layer. We perform both point-scale and spatial comparisons between calibrated WRF-Hydro and the four soil moisture products. For comparisons at observational site level, soil moisture simulated at the LSM grid points that are closest to the seven in-situ stations are evaluated against the seven observational stations and the evaluation metrics are compared with those in the calibrated WRF-Hydro (Section 4.3). For spatial comparisons, soil moisture in WRF-Hydro is interpolated to the grids of the Catchment, Noah, VIC, and Mosaic LSMs using bilinear interpolation, and the evaluation metrics are calculated for each products' time series at each grid (Supplemental Text 2).

3.4.3 Evaluation metrics

To evaluate model performance, we use five metrics including the Pearson correlation coefficient (*r*), mean bias, root mean square error (RMSE), mean absolute error (MAE), Kling-Gupta Efficiency (KGE) (Gupta et al. 2009; Kling et al. 2012), and Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe 1970).

KGE is a comprehensive metric used to evaluate the performance of hydrologic models, and has been applied in other studies to evaluate soil moisture simulations (Lahmers et al. 2019; Vergopolan et al. 2020). It is calculated as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2}\,, \tag{4}$$

where *r* is the correlation coefficient between the simulation and observation, α is the ratio of the standard deviation of simulation to that of the observation, and β is the ratio of the mean of simulation to that of the observation. A model has higher fidelity to observations if KGE and *r* are closer to 1 and if MAE and RMSE approach 0.

NSE is commonly used to evaluate the performance of streamflow simulations in hydrologic models (Wang et al. 2019; Xia et al. 2012). It is calculated as follows:

$$NSE = 1 - \frac{\sum_{t=1}^{t=T}(Q_{sim}(t) - Q_{obs}(t))^2}{\sum_{t=1}^{t=T}(Q_{obs}(t) - \overline{Q_{obs}})^2}\,, \tag{5}$$

where T is the length of the time series, $Q_{obs}(t)$ and $Q_{sim}(t)$ are observed and simulated discharge at time t, respectively, and $\overline{Q_{obs}}$ denotes the mean observed discharge over time. Generally, NSEs of 1 stand for a perfect model-observation match and NSE values approaching 1 indicate excellent model performance in simulating streamflow. Typically, NSEs that are between 0.5 and 0.65 are suggested to be an indication of sufficient model performance (Wang et al. 2019) and negative NSEs represent poor model performance (Schaefli and Gupta 2007).

4. Results

4.1 Simulations of soil moisture with POLARIS-calibrated WRF-Hydro

In this study, we use the fully-distributed WRF-Hydro to simulate soil moisture at high spatial (100-m) and temporal (hourly) resolutions in a central California domain and we leverage observation-based soil databases to inform model calibration (see details regarding the POLARIS dataset and calibration method in Section 3.3). In POLARIS-calibrated WRF-Hydro, default parameters are replaced with POLARIS-based soil parameters. Differences are evident between the POLARIS-based and default soil parameters (Fig. 2h–j). We find that the 16 soil parameters in the default version of WRF-Hydro underestimate the spatial heterogeneity of soil characteristics in the field, while the more spatially refined POLARIS-based soil parameters display greater spatial variation (Fig. 2).

To visualize the simulation of soil moisture in WRF-Hydro over space and time, Fig. 4 shows the simulated evolution of surface soil moisture before, during, and after a storm event in the POLARIS-calibrated WRF-Hydro. The three time slices shown in Fig. 4 are marked by vertical dashed lines in Fig. 5a. The chosen storm spanned 5 days from January 16, 2017 to January 20, 2017 with a maximum precipitation intensity of ~240 mm day^{-1} according to MRMS. During this storm event, the soil moisture at many stations reached their maximum value over our study period. WRF-Hydro simulates the wetting and drainage processes related to the passing of the storm. The high-resolution terrain routing module of WRF-Hydro is able to simulate the interactions between hydrology and the microtopography at finer scales. In addition, the channel routing module of WRF-Hydro simulates channelized streamflow at scales comparable to the channel widths ranging from 1.5 to 100 m, such that WRF-Hydro can simulate greater level of details including the elevated surface soil moisture within channel networks (zoomed-in maps in Fig. 4).

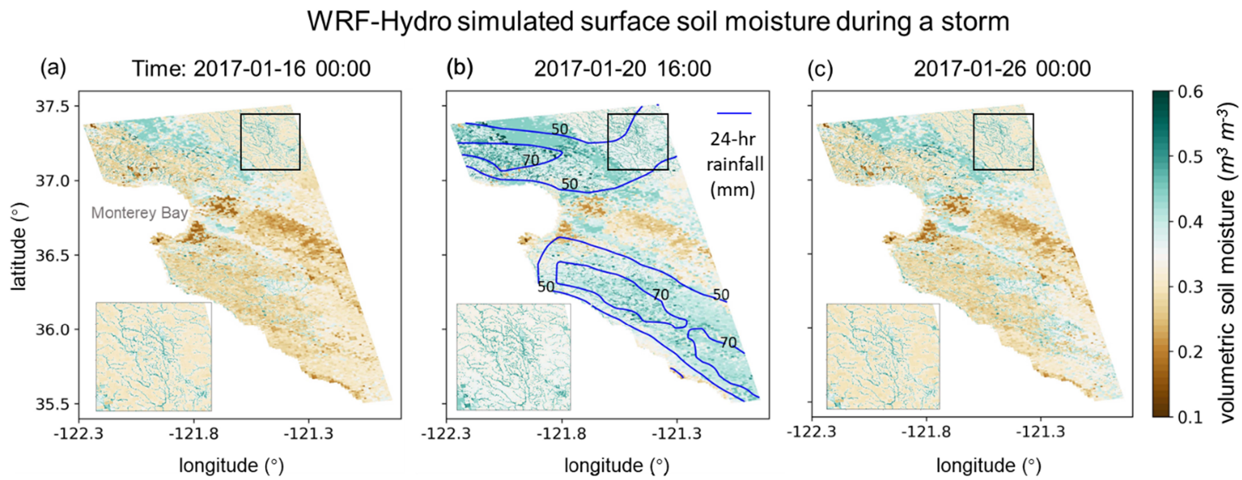

Fig. 4| Evolution of surface soil moisture simulated by the POLARIS-calibrated WRF-Hydro (a) before a storm (2017 January 16 00:00), (b) during the soil moisture peak (2017 January 20 16:00), and (c) after a storm event (2017 January 26 00:00). The three time slices are marked by vertical dashed lines (i), (ii), and (iii) in Fig. 5a, respectively. Embedded maps in the bottom left show zoomed-in details of soil moisture in and near channel networks within the black boxes. 24-hr accumulated precipitation (mm; blue contours) from 00:00 to 23:59 on January 20, 2017 is shown in (b) and contours of 50, and 70 mm are labeled. 24-hr accumulated precipitation for January 16 and 26, 2017 have zeros everywhere.

4.2 Evaluation of WRF-Hydro soil moisture against in-situ soil moisture

To assess the performance of our POLARIS-calibrated WRF-Hydro, soil moisture time series before and after calibration are compared with the observations at seven in-situ stations (Figs. 5&6). Generally, both default and POLARIS-calibrated WRF-Hydro capture the magnitude and variability of the observations, and the ±1 standard deviation simulation envelope for POLARIS-calibrated WRF-Hydro encapsulates the observed soil moisture during the majority of the study period for most stations (Figs. 5&6). After the POLARIS-informed calibration, *r* increases at six of the seven stations, while RMSE and MAE decrease and KGE increases across all in-situ stations (Supplemental Table 2). The average correlation coefficient across the seven stations increases from 0.84 to 0.89, mean RMSE decreases from 0.0916 m^3 m^{-3} to 0.0754 m^3 m^{-3}, and mean KGE increases from 0.57 to 0.67. Four of seven stations have above-average KGEs. Stations gry, hastings, and norris have KGE values above 0.85. Compared to the default WRF-Hydro simulation, the average of the percent change in correlation coefficients across seven stations increases by ~6%, the average RMSE percent change decreases ~18%, and average KGEs increase ~25%, indicating skill improvements in POLARIS-calibrated WRF-Hydro to simulate surface soil moisture.

Our results show that the POLARIS-calibrated WRF-Hydro performs reasonably well during the wet season (October 2016 – February 2017 in our case). Mean KGE during the wet season across seven stations reaches 0.74. However, its performance over the entire study period is negatively affected by the performance during the dry season (starting March 2017 in our case). Dry-season mean KGE across seven stations drops to 0.45. Model performance varies the greatest between wet and dry periods at stations sld, lwd, and norris. KGE values for stations sld, lwd, and norris during 2016 October 1 – 2017 February 28 (wet season) are 0.52, 0.77, and 0.97, respectively, and 0.13, 0.09, and 0.52 during 2017 March 1 – May 31 (dry season). In these stations, moisture in the surface soil layer decreases more slowly during dry weather conditions than observations indicate (Fig. 5a&d & Fig. 6c). Based on the sensitivity experiments we performed prior to our calibration, both *smcmax* and *bexp* can greatly impact the dry-period water drainage rate (from April 15 onwards in Supplemental Fig. 2). Accordingly, we hypothesize that wet biases simulated during the dry period are related to the uncertainties of these two parameters.

In addition to uncertainties in model parameters, another important source of uncertainty that leads to differences between the observations and simulations, especially during the wet season, is the uncertainties of MRMS precipitation. By comparing MRMS precipitation with observational precipitation measured at the seven sites, we found that model biases at stations sld and lgs, which are the two stations with the lowest KGE scores, can be largely explained by the discrepancies in precipitation (Fig. 5a&b; Supplemental Table 2). At station sld, the accumulated precipitation total during the 8-month study period is more than double of that found in MRMS, whereas at station lgs the in-situ precipitation is ~35% higher than the MRMS, leading to the positive bias at sld and negative bias at lgs (Fig. 5a&b). For station gry, MRMS precipitation underestimates the in-situ precipitation by ~25%, which also agrees with the negative mean bias in our WRF-Hydro simulations (Fig. 5c; Supplemental Table 2). The discrepancy in accumulated precipitation amount at station norris is consistent with the dry bias in modeled soil moisture before March 2017 (Fig. 6c). During March – May 2017, however, the parameter uncertainties associated with *smcmax* and *bexp* are likely causing the positive model bias (Supplemental Fig. 2). The differences in accumulated precipitation at the other three sites are relatively small (Figs. 5d & 6a&b).

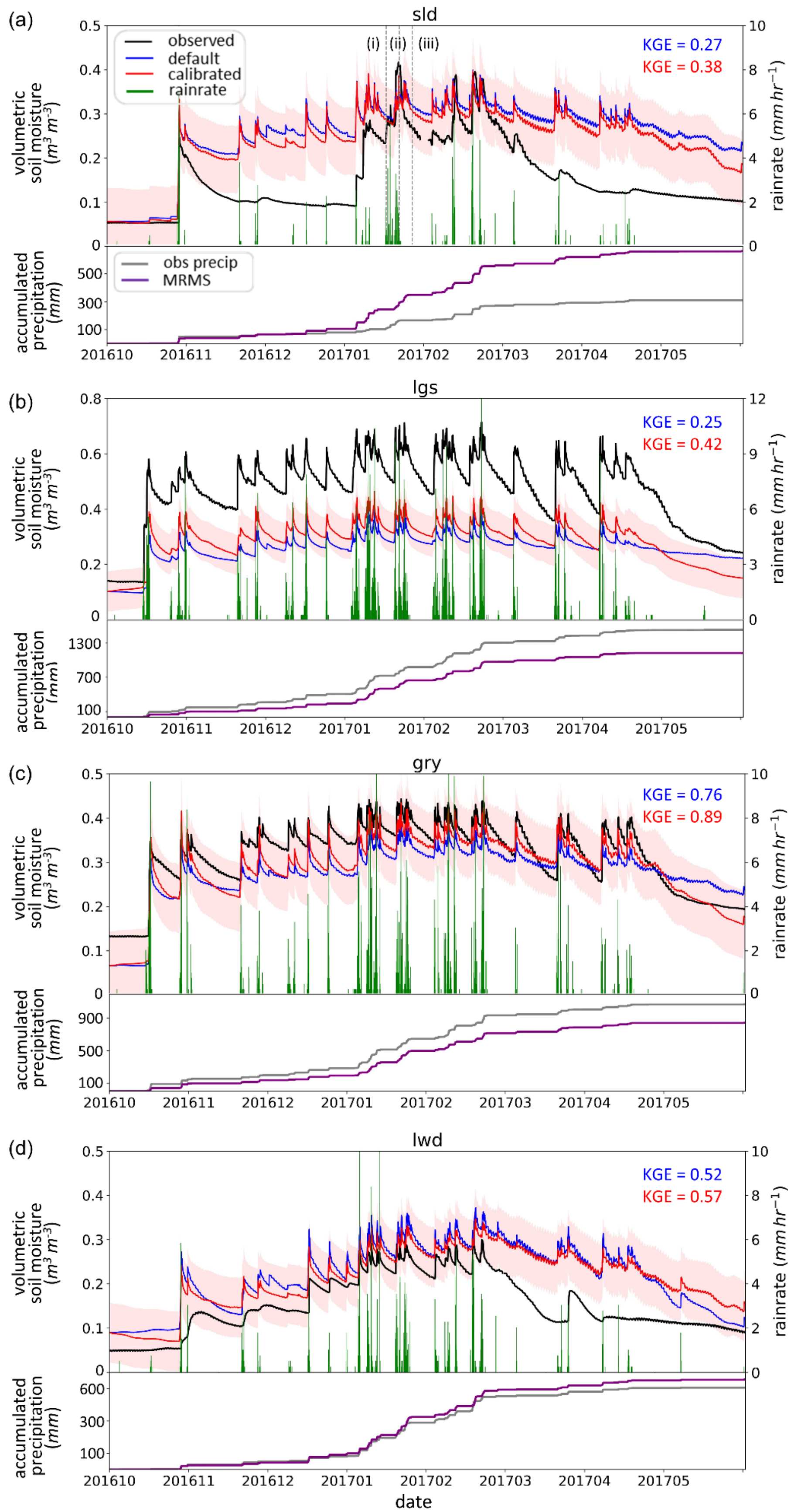
(a)
sld
observed
default
calibrated
rainrate
(i) (ii) (iii)
KGE = 0.27
KGE = 0.38
volumetric soil moisture ($m^3\,m^{-3}$)
rainrate ($mm\,hr^{-1}$)
obs precip
MRMS
accumulated precipitation (mm)
(b)
lgs
KGE = 0.25
KGE = 0.42
(c)
gry
KGE = 0.76
KGE = 0.89
(d)
lwd
KGE = 0.52
KGE = 0.57
201610 201611 201612 201701 201702 201703 201704 201705
date

Fig. 5| Volumetric soil moisture time series and accumulated precipitation amount from 2016 October 1 to 2017 May 31 at NOAA PSL in-situ soil moisture stations (a) sld, (b) lgs, (c) gry, and (d) lwd. Top panels in (a)–(d) show volumetric soil moisture in the observations (in m^3 m^{-3}; black line), default WRF-Hydro simulation (in m^3 m^{-3}; purple line), and POLARIS-calibrated WRF-Hydro simulation (in m^3 m^{-3}; red line). The pink color shading shows the ±1 standard deviations around the POLARIS-calibrated simulation. Hourly precipitation rate in MRMS is shown in green bars (mm hr^{-1}). Grey vertical dashed lines marked with (i), (ii), and (iii) in (a) indicate the three time slices shown in Fig. 4a-c, respectively. Bottom panels in (a)–(d) show the accumulated precipitation measured at the in-situ soil moisture stations (in mm; grey line) and in the MRMS gauge-corrected quantitative precipitation estimation (QPE; in mm; purple line). The accumulated MRMS precipitation is calculated by summing up the precipitation falling on the grid points that are closest to the stations. KGE values are shown in the top right for default simulations (blue) and POLARIS-calibrated simulations (red).

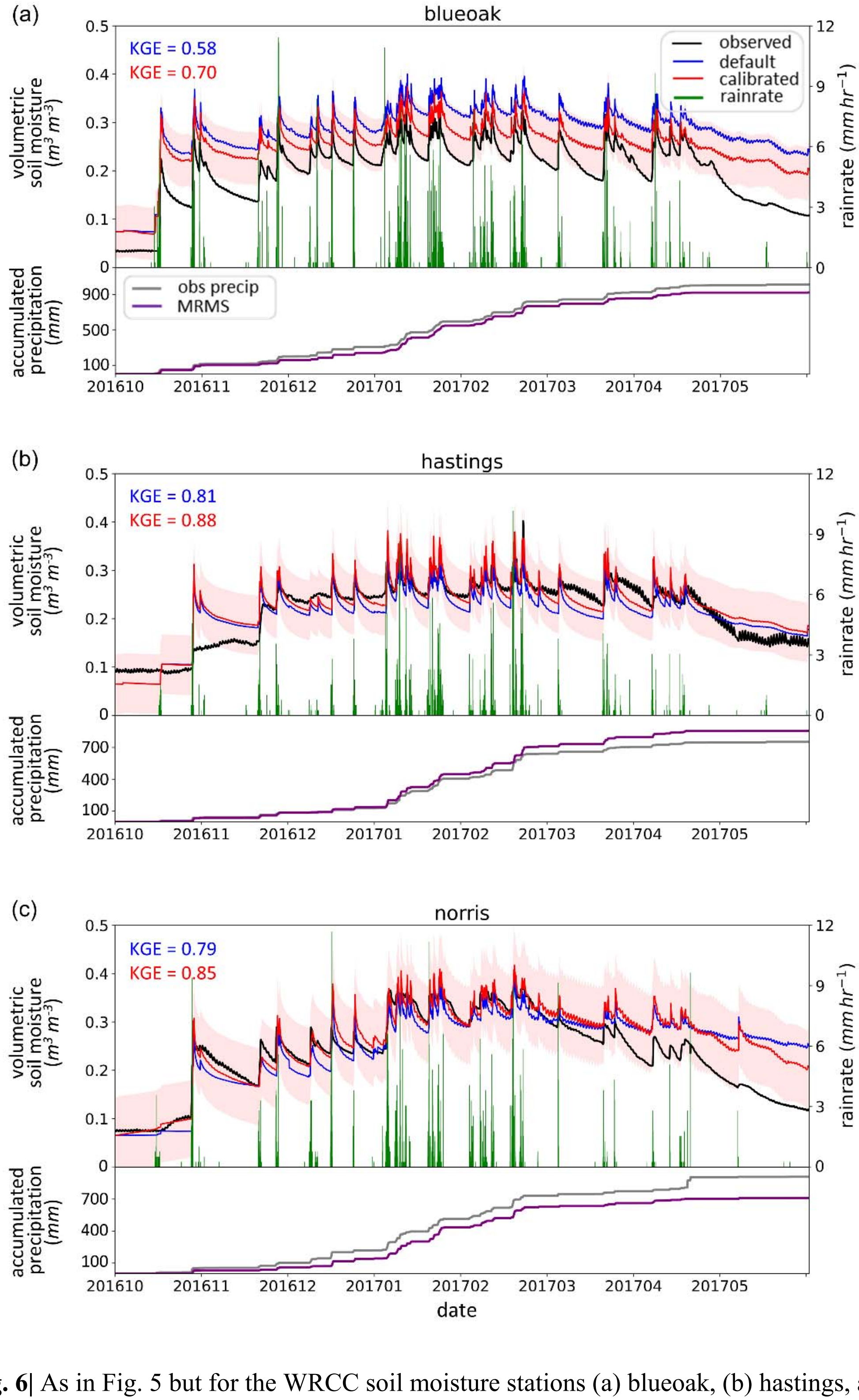

Fig. 6| As in Fig. 5 but for the WRCC soil moisture stations (a) blueoak, (b) hastings, and (c) norris. KGE values are shown in the top left for default (blue) and POLARIS-calibrated simulations (red).

4.3 Comparisons of POLARIS-calibrated WRF-Hydro with other soil moisture products

Next, we compare four other soil moisture products (i.e., SMAP L4, NLDAS-2 Noah, VIC, and Mosaic LSMs) against POLARIS-calibrated WRF-Hydro at the seven in-situ soil moisture observing stations. In general, all five soil moisture products capture the broad variabilities in the observations at sub-daily to sub-seasonal scales (Supplemental Fig. 4). Stations lgs, gry, and blueoak have relatively smaller inter-model variations, whereas the inter-model difference is the largest at stations lwd, hastings, and norris (Supplemental Fig. 4). For stations hastings and norris, the inter-model range roughly encapsulates the observation, whereas at other stations, there are systematic positive or negative biases in all products throughout the study period.

On average, we find that POLARIS-calibrated WRF-Hydro has the best performance in that it has the highest mean KGE across seven stations ($\overline{KGE}$=0.67). Mean KGEs for SMAP L4, Noah, VIC, and Mosaic LSMs are 0.53, 0.55, 0.31, and 0.61, respectively. The KGE scores are the highest in POLARIS-calibrated WRF-Hydro at three of seven stations (i.e., lgs, gry, and lwd) and the second highest at stations blueoak, hastings, and norris. At station sld, POLARIS-calibrated WRF-Hydro has the lowest KGE, which however, are likely explained by the uncertainties in MRMS precipitation at this station (Fig. 5a and Section 4.2). KGE scores are highest in Mosaic LSM at stations hastings and norris, in SMAP L4 at station sld, and in Noah LSM at station blueoak. In VIC LSM, there is a substantial wet bias at stations lwd and hastings and dry bias at gry and norris (Supplemental Fig. 4 and Supplemental Table 3). Indeed, the RMSE in VIC exceeds 0.2 m^3 m^{-3} at stations lgs and lwd and is over 0.1 m^3 m^{-3} at stations gry and hastings, and the KGE score in VIC is the lowest at five of seven stations among the four soil moisture products and POLARIS-calibrated WRF-Hydro (Supplemental Table 3). In addition, because of the limitation of L-band frequency that the radar and radiometer on SMAP spacecraft use to measure soil moisture, SMAP L4 soil moisture may be biased in highly vegetated and topographically complex regions like California (Supplemental Text 1).

We summarize our soil moisture product temporal comparison in Supplemental Tables 2&3 and in scatter plot format (Fig. 7). In Fig. 7, evaluation metrics including *r*, RMSE, and KGE of default WRF-Hydro, SMAP L4, Noah, VIC, and Mosaic LSMs are plotted against the evaluation metrics of POLARIS-calibrated WRF-Hydro. Each point represents an evaluation metric of a soil moisture product at a soil moisture station and how it compares with the metric of the POLARIS-calibrated WRF-Hydro at the same station. In figures of *r* and KGE, points below the one-to-one line indicate higher performance in the

POLARIS-calibrated WRF-Hydro (Fig. 7a&c), whereas points above the one-to-one line in the figure of RMSEs represent reduced bias in the POLARIS-calibrated WRF-Hydro. In most cases, the POLARIS-calibrated WRF-Hydro has increased *r* (25 of 35 points), reduced errors (28 of 35 points), and increased KGEs (29 of 35 points), indicating its higher soil moisture fidelity compared to other soil moisture products and default WRF-Hydro (Fig. 7). The POLARIS-calibrated WRF-Hydro has either the highest or the second highest KGE at stations other than sld (Supplemental Table 3). The lowest KGE at station sld in POLARIS-calibrated WRF-Hydro can largely be explained by the large uncertainties of MRMS precipitation as we show in Fig. 5a and discussed in Section 4.2. To fully evaluate the performance of POLARIS-calibrated WRF-Hydro over our model domain, we also provide a spatial comparison between the soil moisture products in Supplemental Figs. 5–8 and a description on the comparison between POLARIS-calibrated WRF-Hydro and NLDAS-2 Mosaic LSM can be found in Supplemental Text 2.

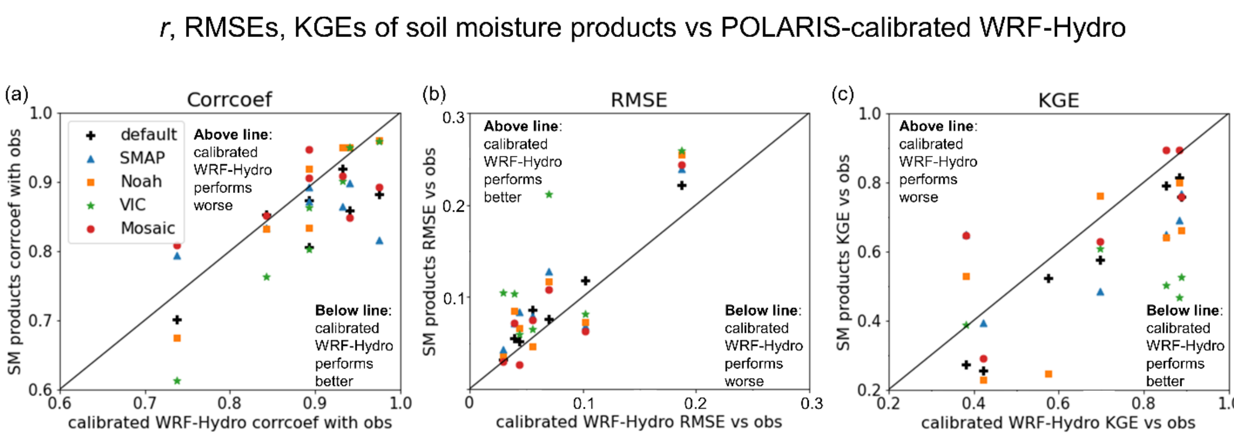

Fig. 7| Evaluation metrics of default WRF-Hydro and four soil moisture products (i.e., SMAP L4, NLDAS-2 Noah, VIC, and Mosaic LSMs) compared with the evaluation metrics of POLARIS-calibrated WRF-Hydro against in-situ observations. (a) Correlation coefficients, (b) RMSEs (m^3 m^{-3}), and (c) KGEs of default WRF-Hydro (black crosses) and other soil moisture products (SMAP L4: blue triangles, Noah: orange squares, VIC: green stars, Mosaic: red circles) versus that of POLARIS-calibrated WRF-Hydro. One-to-one line is indicated by the black solid line.

4.4 Improved streamflow fidelity in the soil data-informed calibrated simulations

Given the key role soil moisture plays in overland flow and subsurface flow production processes, improved soil moisture simulation has been found to improve streamflow simulations spontaneously (Aubert et al. 2003; Lee et al. 2011). In this section, we evaluate the streamflow simulations in the POLARIS-calibrated WRF-Hydro to investigate the linkages between improved surface soil moisture simulations and streamflow simulations in WRF-Hydro. We show that improved surface soil moisture accuracy only moderately improves streamflow simulation. However, when we account for SSURGO depth-to-bedrock data total soil thickness (i.e., the POLARIS-40 cm soil experiment), the effects of soil moisture on streamflow fidelity are enhanced. Meanwhile, the soil moisture model fidelity in POLARIS-40 cm soil experiment is not diminished.

We show that compared to default WRF-Hydro, streamflow fidelity improves in both calibrated experiments with the greatest improvement in the POLARIS-40 cm soil experiment (Fig. 8 and Supplemental Table 4). In the POLARIS-calibrated WRF-Hydro with 2 m soil experiment, r increases, error decreases, and NSE increases across most stations (8/9 stations see an improvement). However, the improvement is quite moderate – none of the stations have NSEs above 0.5 after the calibration (Fig. 8c), which suggests that improving surface soil moisture solely is not sufficient to significantly improve streamflow simulation in WRF-Hydro and the total soil column needs to be considered. Indeed, model performance improves by a large fraction in the POLARIS-40 cm soil experiment (Fig. 8). In the POLARIS-40 cm soil experiment, r increases across all nine gages with a mean of 0.84 (p value$<<0.0001$). Six gages have NSEs exceeding 0.5, indicating sufficient-to-good model performance in these basins (Fig. 8c). RMSE decreases and NSE increases significantly at seven of the nine stations and the mean NSE score across the seven improved gages reaches 0.63. At the other two gages 11152050 and 11152000 (Fig. 1a), however, r increases but NSE decreases to negative values in POLARIS-40 cm soil experiment because the model overestimates the discharge magnitude (Fig. 8). We hypothesize that the positive model bias at these two gages can be partially attributed to the overestimation in the MRMS precipitation in that area. Though in-situ precipitation data is not available at USGS stream gages, we make this assumption based on the fact that the two gages are located in proximity to the soil moisture station sld (Fig. 1a) that has in-situ precipitation measurements. We have discussed in Section 4.2 that the surface soil moisture at sld is overestimated in the POLARIS-calibrated WRF-Hydro due to the positive bias in the MRMS precipitation so it is likely that streamflow simulations in POLARIS-40 cm soil experiemnt are also biased high at gages 11152050 and 11152000 due to overestimated precipitation.

To closely examine the improvement and remaining biases after calibration, we compare the modeled and observed hydrographs at three selected stations in Fig. 9. Corresponding results for the rest of the stations are shown in Supplemental Figs. 9&10. The three representative gages we choose to show here are Gage 11169800 which has the highest NSE among nine stations after calibration (Fig. 9a&b), Gage 11157500 which has the largest improvement in the POLARIS-40 cm soil experiment (Fig. 9c&d), and Gage 11152000 which has a diverging change in the two calibration experiments (i.e., an increased performance in POLARIS-2 m soil experiment but decreased performance in POLARIS-40 cm soil experiment) (Fig. 9e&f). Generally, POLARIS-40 cm performs the best at capturing the magnitude and timing of peak flow events among the three sets of simulations (Fig. 9). It is likely because the default 2-m soil column overestimates the realistic soil thickness in Coast Ranges of central California. With thinner soils, less water can be stored in the soil column and the hydrologic response is much faster. Indeed, in default and POLARIS-2 m soil experiments, the model underestimates the discharge magnitude especially for the storm events prior to January 15, 2017, while the POLARIS-40 cm experiment is able to capture the first few storms in January for all nine gages (Fig. 9a, c&e and Supplemental Figs. 9&10). In addition, the default WRF-Hydro does not capture any streamflow at gage 11157500 during the entire time period (blue line in Fig. 9c), whereas in the POLARIS-40 cm case streamflow is simulated with a small mean bias of -0.3 $m^3\ s^{-1}$ and an NSE of 0.56.

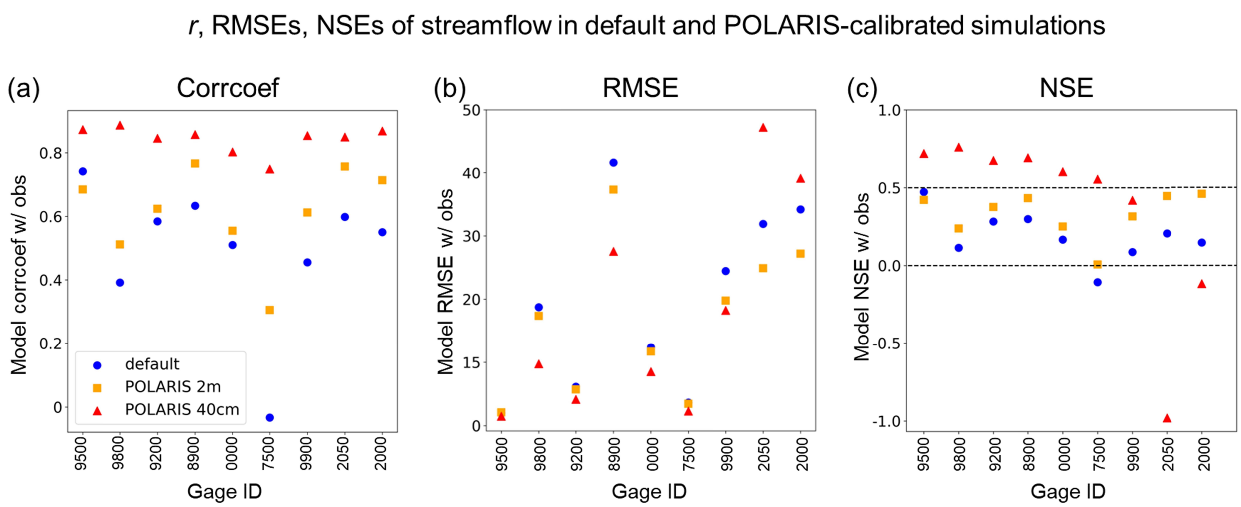

Fig. 8| Scatter plots of performance metrics including (a) correlation coefficients, (b) RMSEs, and (c) NSEs in the default (blue circles), POLARIS-calibrated WRF-Hydro (orange squares), and POLARIS-calibrated WRF-Hydro with 40 cm soil (red triangles). The labels on the x axes show the last 4 digits of the USGS stream gage IDs. The horizontal dashed lines in (c) indicate NSE of 0.5 and 0.0, respectively. NSE of 0.5 is suggested to be the threshold of sufficient model performance whereas NSE below 0 indicates poor model performance.

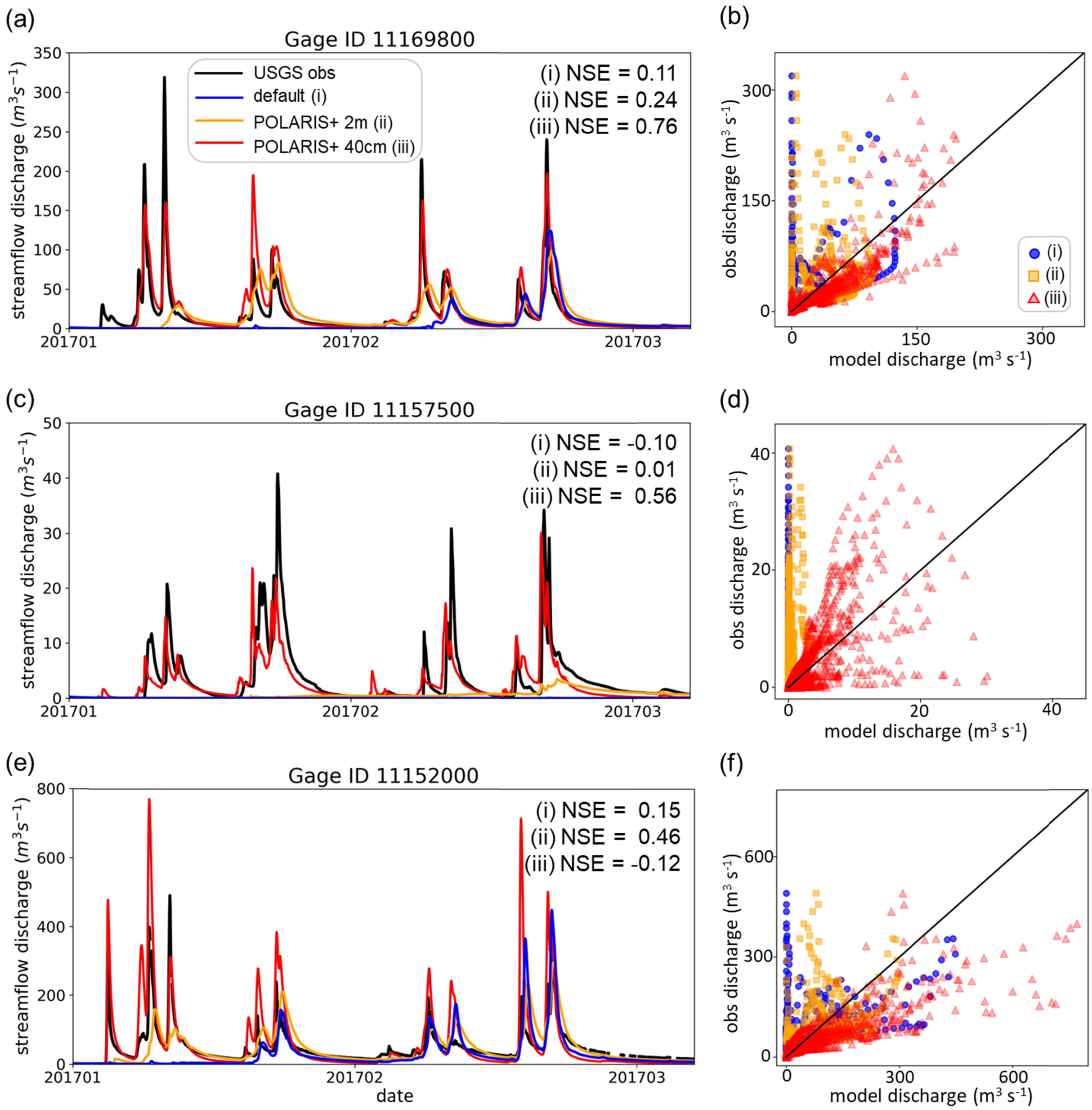

Fig. 9| (a), (c), and (e) Streamflow hydrographs of USGS stream observations (black line), default WRF-Hydro (blue line; labeled with (i)), POLARIS-calibrated WRF-Hydro with 2 m soil (orange line; labeled with (ii)), and POLARIS-calibrated WRF-Hydro with 40 cm soil (red line; labeled with (iii)) at three selected stations (Gages 11169800, 11157500, and 11152000). NSEs of the simulations are shown. (b), (d), and (f) show the scatter plots of experiment (i) (blue circles), (ii) (orange squares), and (iii) (red triangles) along the x-axis versus the observations along the y-axis. The 1:1 line is shown as black solid line in (b), (d) and (f).

5. Conclusions and discussions

In this study, we use open access soil databases to inform the parameters in two calibration experiments in WRF-Hydro. We not only create a simulated soil moisture product that outperforms four well-established soil moisture products but we also significantly improve model streamflow fidelity. In our first experiment (i.e., POLARIS-calibrated WRF-Hydro), we replace the soil hydraulic parameters in the default version of WRF-Hydro with the POLARIS-based soil parameters to calibrate surface soil moisture simulations. We evaluate the POLARIS-calibrated WRF-Hydro simulated soil moisture over an 8-month period against seven in-situ soil moisture stations and see an improvement across all seven stations after the calibration. On average, KGE increases ~25% after calibration. Compared to other spatially-distributed soil moisture simulations in SMAP L4, NLDAS-2 Noah, VIC, and Mosaic LSMs, the POLARIS-calibrated WRF-Hydro has the best average performance across seven sites and produces the highest correlation, lowest error, and highest KGE in most cases. Despite the improved surface soil moisture fidelity in POLARIS-calibrated WRF-Hydro, streamflow simulation is only moderately improved. As such, we reduce the soil thickness from 2 m in POLARIS-calibrated WRF-Hydro to 40 cm based on SSURGO depth-to-bedrock data to better replicate the effects of the entire soil column on streamflow production in California. Streamflow fidelity significantly improves in our POLARIS-calibrated with 40 cm soil experiment - seven of the nine USGS gages see an increased NSE and the mean of NSEs at the seven improved gages reaches 0.63. Our data-informed calibration method uses open access, spatially-distributed soil physical information available over the CONUS to constrain our hydrological model's parameter uncertainties. Our calibration method does not require iterative model simulations which highlights its simplicity and potentially wide applicability to improve soil moisture and streamflow simulations in fully-distributed hydrological models, which could facilitate studies in a wide range of disciplines in data-scarce areas.

Despite the generally high model fidelity, we note that there are still considerable differences between WRF-Hydro soil moisture simulations and the observations at some stations. Consistent with other studies, we are able to explain a large portion of the uncertainties in our soil moisture simulations with the uncertainties of precipitation forcing (Alfieri et al. 2012; Hapuarachchi et al. 2011). The gauge-corrected MRMS precipitation we use in this study is found to substantially deviate from the in-situ precipitation at stations sld and gry, which largely explains the differences between the soil moisture simulation and observation at these locations (Section 4.2). Despite the uncertainties in the gauge-corrected MRMS, it provides gridded precipitation at high spatial (1 km) and temporal (hourly) resolutions, making it a valuable forcing for high-resolution hydrological models. More details regarding MRMS uncertainties can be found in the Appendix A of Li et al. (2022). Additional uncertainties can be traced to the soil parameters. By using the POLARIS dataset, an observation-based statistical soil property dataset with higher accuracy and spatial resolution, we have constrained some of the parameter uncertainties of soil porosity and saturated hydraulic conductivity. Nevertheless, the performance of our calibrated model is negatively affected by dry-season simulations as we discussed in Section 4.2. Specifically, POLARIS-calibrated WRF-Hydro tends to underestimate the speed of water drainage during the transition period from the wet to dry season (Figs. 5&6). Among various factors that could cause the model's underestimation of drainage speed, the parameter *bexp* that controls the speed of flows through the soil column is likely the main cause. To derive *b*, we use the clay fraction from POLARIS and a linear regression model from Cosby et al. (1984), which may result in the propagation and accumulation of uncertainties. Indeed, Cosby et al. (1984) also documented the uncertainties of the calculated *b* coefficient. From the perspective of flooding and landslide hazard assessment and control, simulation during wet season is of particular importance. For drought monitoring, agriculture, and water resource management, however, dry-season soil moisture simulation is also critical. To improve dry-season soil moisture simulations in WRF-Hydro, the uncertainties of *bexp* also need to be considered when implementing the calibration method.

In addition to prediction uncertainties, the differences between simulation and observation could originate from other factors including the comparison approach and the possible instrumental errors in soil moisture measurements. To compare with point-scale observations, we use the soil moisture simulated at the grid point that is located closest to the in-situ site. In addition, the WRF-Hydro surface soil moisture is a depth average of the 0–10 cm soil layer while the in-situ soil moisture is measured at 10 cm depth for PSL stations and 5 cm depth for WRCC stations. We also note that the observations might be subject to errors. For example, the observed soil moisture at station lgs was abnormally high during the wet season, exceeding 60%. However, the maximum surface porosity in proximity of station lgs only achieves 0.55 according to the 30-m POLARIS. By referring to the soil moisture sensor instruction manual and consulting with the experts that operate and maintain the PSL stations, we found that the soil moisture sensor at station lgs was likely submerged in ponded water due to the large amount of accumulated precipitation during the wet season and the soil moisture was likely substantially overestimated (Fig. 5b).

To enhance the capability of WRF-Hydro to simulate soil moisture, the utility of POLARIS dataset can be further explored. In addition to median values, POLARIS also provides a range of soil property statistics including the mean, mode, 5P, and 95P, which can facilitate an investigation on parameter uncertainties in WRF-Hydro associated with individual parameters. In addition, due to a lack of in-situ soil moisture data of deeper soils, this study is focused on calibrating and validating the surface soil moisture. Accurate surface soil moisture is most important for predicting the occurrence of flooding events via a control on rainfall partitioning (Aubert et al. 2003; Crow et al. 2018; Houser et al. 2003; Kerr 2007) and it also provides initial conditions for slope instability models to predict slope failures (Cai et al. 2019; Di Matteo et al. 2018). Nevertheless, soil water content of deeper soils is critical for ecology, agriculture, drought monitoring, and water and energy fluxes. Both the soil parameters in POLARIS and the soil moisture simulations from WRF-Hydro have multiple soil layers that extend to as deep as 2 m below ground, and different parameter values can be assigned to different layers in WRF-Hydro. To calibrate the soil moisture for all soil layers in WRF-Hydro, POLARIS soil properties from other soil layers will also be needed.

Compared to soil moisture, streamflow is the variable that has been more extensively calibrated and used in WRF-Hydro (Lahmers et al. 2019; Wang et al. 2019) which makes the implication of calibrating soil moisture on streamflow simulation an important topic to cover. Our results show that by just calibrating the surface soil moisture the correlation increases and some biases are reduced but streamflow simulation is not significantly improved (Fig. 8). In contrast, in the experiment that adjusts the total soil thickness according to SSURGO, streamflow simulations across most gages are much improved without diminishing the soil moisture model fidelity. This indicates that the 2 m soil column in the default setting of WRF-Hydro largely overestimates the soil thickness in our model domain (Supplemental Fig. 3) and adjusting the total soil thickness is more efficient than calibrating surface soil moisture to improve streamflow fidelity. This is further proved by running an additional experiment, i.e., default WRF-Hydro with 40 cm soil, in which we find the POLARIS-calibrated WRF-Hydro with 40 cm soil still yields the best results on average and default WRF-Hydro with 40 cm soil yields the second best in terms of streamflow simulation (Supplemental Table 5). Compared to other streamflow calibration studies which focus on variables that control the discharge volume and hydrograph shape, such as the water retention depth coefficient (*REFKDT*), bottom openness (*SLOPE*), and Manning's coefficient (*n*), our method is only focused on soil-related parameters and we are able to achieve similar model performance to simulate streamflow (Wang et al. 2019; Yucel et al. 2015). Nevertheless, we acknowledge that applying a spatially homogeneous total soil thickness to a large domain can introduce bias in the simulation of discharge magnitude at some locations (e.g., Gages 11152050 and 11152000) but the current version of WRF-Hydro is not capable of assigning spatially-distributed total soil thickness. Model developments to enable spatially-varying soil thickness would therefore be advantageous. In addition, the frequent and widespread wildfires in the Coast Ranges of central California and their impacts on downstream hydrology have added additional complexities for streamflow predictions (Li et al. 2022). Accordingly, we suggest users consider many factors to replicate the real-world conditions before intensively calibrating the streamflow parameters to avoid overfitting.

Given the simplicity of the concept underlying our data-informed calibration method, we argue for its extendibility to other hydrological models that deal with spatially-distributed soil parameters and other geographic areas. Indeed, the applicability of our method to other geographic locations is only limited by the availability of reliable and updated soil hydraulic parameter data. Over the CONUS, POLARIS and SSURGO are open access databases, and for studies outside the U.S., the Global Soil Dataset, for example, provides gridded soil hydraulic parameters for use in Earth Systems Models around the globe (Shangguan et al. 2014).

Financial support

This research has been supported by the National Science Foundation (grant no. 1848683).

Data availability statement

The NLDAS-2 reanalysis forcing data are publicly available at NASA GES DISC: https://doi.org/10.5067/6J5LHHOHZHN4 (Xia et al., 2009). The MRMS gauge-corrected precipitation estimate is archived at https://mtarchive.geol.iastate.edu/. POLARIS dataset can be downloaded at http://hydrology.cee.duke.edu/POLARIS/PROPERTIES/v1.0/ (Chaney et al. 2016). SSURGO dataset is available at https://websoilsurvey.nrcs.usda.gov/ (Soil Survey Staff, 2021). The PSL in situ soil moisture data are publicly available at https://psl.noaa.gov/data/obs/datadisplay/ (NOAA PSL, 2021). WRCC soil moisture data is available at https://wrcc.dri.edu/weather/index.html (WRCC 2021). SMAP Level 4 version 6 soil moisture data is available at https://nsidc.org/data/spl4smau/versions/6 (Reichle et al. 2021). A more recent version (version 7) of SMAP Level 4 is available at https://nsidc.org/data/spl4smau/versions/7 (Reichle et al. 2022). NLDAS-2 Noah, VIC, and Mosaic soil moisture datasets are available at https://disc.gsfc.nasa.gov/datasets?keywords=NLDAS&page=1&measurement=Soil%20Moisture%2FWater%20Content (NCEP/EMC, 2009, 2012, 2014; Xia et al. 2012). The USGS streamflow is publicly available at https://doi.org/10.5066/F7P55KJN (USGS, 2016). All processed data required to reproduce the results of this study are archived on Zenodo at https://doi.org/10.5281/zenodo.7487179 (Li 2022).

References

Acharya S, Kaplan DA, McLaughlin DL, Cohen MJ (2022) In-Situ Quantification and Prediction of Water Yield From Southern US Pine Forests 58:e2021WR031020 doi:https://doi.org/10.1029/2021WR031020

Al-Yaari A et al. (2017) Evaluating soil moisture retrievals from ESA's SMOS and NASA's SMAP brightness temperature datasets 193:257-273

Alfieri L, Salamon P, Pappenberger F, Wetterhall F, Thielen J (2012) Operational early warning systems for water-related hazards in Europe Environmental Science & Policy 21:35-49 doi:https://doi.org/10.1016/j.envsci.2012.01.008

Alvioli M, Baum RLJEM, Software (2016) Parallelization of the TRIGRS model for rainfall-induced landslides using the message passing interface 81:122-135

Aubert D, Loumagne C, Oudin LJJoh (2003) Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall–runoff model 280:145-161

Baum RL, Savage WZ, Godt JW (2008) TRIGRS: a Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, version 2.0. US Geological Survey Reston, VA, USA,

Becker R, Koppa A, Schulz S, Usman M, aus der Beek T, Schüth C (2019) Spatially distributed model calibration of a highly managed hydrological system using remote sensing-derived ET data Journal of Hydrology 577:123944 doi:https://doi.org/10.1016/j.jhydrol.2019.123944

Berg A, Sheffield J (2018) Climate Change and Drought: the Soil Moisture Perspective Current Climate Change Reports 4:180-191 doi:10.1007/s40641-018-0095-0

Boucher MA, Quilty J, Adamowski JJWRR (2020) Data assimilation for streamflow forecasting using extreme learning machines and multilayer perceptrons 56:e2019WR026226

Cai J-S, Jim Yeh T-C, Yan EC, Tang R-X, Hao Y-H, Huang S-Y, Wen J-C (2019) Importance of variability in initial soil moisture and rainfalls on slope stability Journal of Hydrology 571:265-278 doi:https://doi.org/10.1016/j.jhydrol.2019.01.046

Campbell Scientific, INC. CS616 and CS625 water content reflectometer instruction manual. Retrieved on Feb. 2021 from https://psl.noaa.gov/data/obs/instruments/SoilWaterContent.pdf

Campbell Scientific, INC. CS615 water content reflectometer instruction manual. Retrieved Feb. 2021. https://s.campbellsci.com/documents/us/manuals/cs615.pdf

Carrão H, Russo S, Sepulcre-Canto G, Barbosa P (2016) An empirical standardized soil moisture index for agricultural drought assessment from remotely sensed data International Journal of Applied Earth Observation and Geoinformation 48:74-84 doi:https://doi.org/10.1016/j.jag.2015.06.011

Chaney NW, Wood EF, McBratney AB, Hempel JW, Nauman TW, Brungard CW, Odgers NP (2016) POLARIS: A 30-meter probabilistic soil series map of the contiguous United States Geoderma 274:54-67 doi:10.1016/j.geoderma.2016.03.025

Chen F, Crow WT, Bindlish R, Colliander A, Burgin MS, Asanuma J, Aida KJRSoE (2018) Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation 214:1-13

Coe JA, Kinner DA, Godt JW (2008) Initiation conditions for debris flows generated by runoff at Chalk Cliffs, central Colorado Geomorphology 96:270-297 doi:https://doi.org/10.1016/j.geomorph.2007.03.017

Colby B (1956) Relationship of sediment discharge to streamflow. US Dept. of the Interior, Geological Survey, Water Resources Division,

Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R.: A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils, Water Resour. Res., 20(6), 682-690, doi:https://doi.org/10.1029/WR020i006p00682, 1984.

Crow W, Kumar S, Bolten JJH, Sciences ES (2012) On the utility of land surface models for agricultural drought monitoring 16:3451-3460

Crow W, Van den Berg MJWRR (2010) An improved approach for estimating observation and model error parameters in soil moisture data assimilation 46

Crow WT, Chen F, Reichle RH, Xia Y, Liu Q (2018) Exploiting Soil Moisture, Precipitation, and Streamflow Observations to Evaluate Soil Moisture/Runoff Coupling in Land Surface Models Geophys Res Lett 45:4869-4878 doi:https://doi.org/10.1029/2018GL077193

Davy P, Lague DJJoGRES (2009) Fluvial erosion/transport equation of landscape evolution models revisited 114

De Santis D, Biondi D, Crow W, Camici S, Modanesi S, Brocca L, Massari CJWRR (2021) Assimilation of satellite soil moisture products for river flow prediction: An extensive experiment in over 700 catchments throughout Europe 57:e2021WR029643

Di Matteo L et al. (2018) Reliability of water content estimation by profile probe and its effect on slope stability Landslides 15:173-180 doi:10.1007/s10346-017-0895-7

Dottori F, Salamon P, Bianchi A, Alfieri L, Hirpa FA, Feyen LJAiwr (2016) Development and evaluation of a framework for global flood hazard mapping 94:87-102

Douville H, Chauvin F (2000) Relevance of soil moisture for seasonal climate predictions: A preliminary study vol 16.

Fan X, Liu Y, Gan G, Wu G (2020) SMAP underestimates soil moisture in vegetation-disturbed areas primarily as a result of biased surface temperature data Remote Sensing of Environment 247:111914 doi:10.1016/j.rse.2020.111914

Fennessy MJ, Shukla J (1999) Impact of Initial Soil Wetness on Seasonal Atmospheric Prediction J Climate 12:3167-3180 doi:10.1175/1520-0442(1999)012<3167:IOISWO>2.0.CO;2

Ford TW, Quiring SM (2019) Comparison of Contemporary In Situ, Model, and Satellite Remote Sensing Soil Moisture With a Focus on Drought Monitoring 55:1565-1582 doi:https://doi.org/10.1029/2018WR024039

Friend AD, Kiang NYJJoC (2005) Land surface model development for the GISS GCM: Effects of improved canopy physiology on simulated climate 18:2883-2902

Gallagher M, Doherty JJEM, Software (2007) Parameter estimation and uncertainty analysis for a watershed model 22:1000-1020

Gasmo JM, Rahardjo H, Leong EC (2000) Infiltration effects on stability of a residual soil slope Computers and Geotechnics 26:145-165 doi:https://doi.org/10.1016/S0266-352X(99)00035-X

Gochis, D.J., M. Barlage, R. Cabell, M. Casali, A. Dugger, K. FitzGerald, M. McAllister, J. McCreight, A. RafieeiNasab, L. Read, K. Sampson, D. Yates, Y. Zhang (2020). The WRF-Hydro® modeling system technical description, (Version 5.1.1). NCAR Technical Note. 107 pages. Available online at: https://ral.ucar.edu/sites/default/files/public/WRFHydroV511TechnicalDescription.pdf.

Gong G, Wang L, Condon L, Shearman A, Lall UJJJotAWRA (2010) A simple framework for incorporating seasonal streamflow forecasts into existing water resource management practices 1 46:574-585

Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling Journal of Hydrology 377:80-91 doi:https://doi.org/10.1016/j.jhydrol.2009.08.003

Handwerger AL, Huang M-H, Fielding EJ, Booth AM, Bürgmann R (2019) A shift from drought to extreme rainfall drives a stable landslide to catastrophic failure Scientific Reports 9:1569 doi:10.1038/s41598-018-38300-0

Hapuarachchi HAP, Wang QJ, Pagano TC (2011) A review of advances in flash flood forecasting 25:2771-2784 doi:https://doi.org/10.1002/hyp.8040

Holtan, H. N., C. B. England, G. P. Lawless, and G. A. Schumaker, Moisture-tension data for selected soils on experimental watersheds, Rep. ARS 41-144, 609 pp., Agric. Res. Serv., Beltsville, Md., 1968.

Houser PRJHoW, Climate, Water: Atmospheric Chemistry H, Wiley SIJ, Sons IH, New Jersey (2003) Infiltration and soil moisture processes

Johnson K, Sitar NJCGJ (1990) Hydrologic conditions leading to debris-flow initiation 27:789-801

Jong B-T, Ting M, Seager R (2016) El Niño's impact on California precipitation: seasonality, regionality, and El Niño intensity Environmental Research Letters 11:054021 doi:10.1088/1748-9326/11/5/054021

Kang Y, Khan S, Ma X (2009) Climate change impacts on crop yield, crop water productivity and food security – A review Progress in Natural Science 19:1665-1674 doi:https://doi.org/10.1016/j.pnsc.2009.08.001

Kean JW, McCoy SW, Tucker GE, Staley DM, Coe JA (2013) Runoff-generated debris flows: Observations and modeling of surge initiation, magnitude, and frequency Journal of Geophysical Research: Earth Surface 118:2190-2207 doi:https://doi.org/10.1002/jgrf.20148

Kerr YH (2007) Soil moisture from space: Where are we? Hydrogeology Journal 15:117-120 doi:10.1007/s10040-006-0095-3

Kling H, Fuchs M, Paulin M (2012) Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios Journal of Hydrology 424-425:264-277 doi:https://doi.org/10.1016/j.jhydrol.2012.01.011

Koster RD (2004) Realistic initialization of land surface states: Impacts on subseasonal forecast skill vol 5.

Koster RD, Guo Z, Yang R, Dirmeyer PA, Mitchell K, Puma MJ (2009) On the Nature of Soil Moisture in Land Surface Models J Climate 22:4322-4335 doi:10.1175/2009JCLI2832.1

Koster RD et al. (2010) Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment 37

Kumar SV, Dirmeyer PA, Peters-Lidard CD, Bindlish R, Bolten JJRSoE (2018) Information theoretic evaluation of satellite soil moisture retrievals 204:392-400

Lahmers TM, Castro CL, Hazenberg P (2020) Effects of Lateral Flow on the Convective Environment in a Coupled Hydrometeorological Modeling System in a Semiarid Environment J Hydrometeor 21:615-642 doi:10.1175/JHM-D-19-0100.1

Lahmers TM et al. (2019) Enhancing the structure of the WRF-Hydro hydrologic model for semiarid environments 20:691-714

Lahmers, T. M., Kumar, S. V., Rosen, D., Dugger, A., Gochis, D. J., Santanello, J. A., et al. (2022). Assimilation of NASA's Airborne Snow Observatory snow measurements for improved hydrological modeling: A case study enabled by the coupled LIS/WRF-Hydro system. Water Resources Research, 58, e2021WR029867. https://doi.org/10.1029/2021WR029867

Leach JM, Kornelsen KC, Coulibaly PJJoH (2018) Assimilation of near-real time data products into models of an urban basin 563:51-64

Lee H, Seo D-J, Koren V (2011) Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the data and initial model soil moisture states Adv Water Resour 34:1597-1615 doi:https://doi.org/10.1016/j.advwatres.2011.08.012

Legates DR, Mahmood R, Levia DF, DeLiberty TL, Quiring SM, Houser C, Nelson FEJPiPG (2011) Soil moisture: A central and unifying theme in physical geography 35:65-86

Li C et al. (2022) Augmentation of WRF-Hydro to simulate overland-flow- and streamflow-generated debris flow susceptibility in burn scars Nat Hazards Earth Syst Sci 22:2317-2345 doi:10.5194/nhess-22-2317-2022

Li C (2022) data-informed calibration of WRF-Hydro data1. Zenodo, [Data set]. doi: https://doi.org/10.5281/zenodo.7487179

Loizu J, Massari C, Alvarez-Mozos J, Tarpanelli A, Brocca L, Casali JJAiWR (2018) On the assimilation set-up of ASCAT soil moisture data for improving streamflow catchment simulation 111:86-104

Ma H, Zeng J, Chen N, Zhang X, Cosh MH, Wang W (2019) Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: A comprehensive assessment using global ground-based observations Remote Sensing of Environment 231:111215 doi:https://doi.org/10.1016/j.rse.2019.111215

Massari C, Brocca L, Moramarco T, Tramblay Y, Lescot J-FDJAiWR (2014) Potential of soil moisture observations in flood modelling: Estimating initial conditions and correcting rainfall 74:44-53

Matgen P et al. (2010) Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept 14:1773-1785

Mohanty BP, Cosh MH, Lakshmi V, Montzka C (2017) Soil Moisture Remote Sensing: State-of-the-Science 16:vzj2016.2010.0105 doi:https://doi.org/10.2136/vzj2016.10.0105

Narasimhan B, Srinivasan R (2005) Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring vol 133.

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I — A discussion of principles Journal of Hydrology 10:282-290 doi:https://doi.org/10.1016/0022-1694(70)90255-6

NCEP/EMC (2012), NLDAS Noah Land Surface Model L4 Hourly 0.125 x 0.125 degree V002, Edited by David Mocko, NASA/GSFC/HSL, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), [Data set], Accessed: 12/02/2021, 10.5067/47Z13FNQODKV

NCEP/EMC (2014), NLDAS VIC Land Surface Model L4 Hourly 0.125 x 0.125 degree V002, Edited by David Mocko, NASA/GSFC/HSL, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), [Data set], Accessed: 12/02/2021, 10.5067/ELBDAPAKNGJ9

NCEP/EMC (2009), NLDAS Mosaic Land Surface Model L4 Hourly 0.125 x 0.125 degree V002, Edited by David Mocko, NASA/GSFC/HSL, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), [Data set], Accessed: 12/02/2021, 10.5067/EN4MBWTCENE5

Niu G-Y et al. (2011) The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements J Geophys Res Atm 116 doi:https://doi.org/10.1029/2010JD015139

NOAA Physical Sciences Laboratory (PSL): Surface Meteorology and Physics Data, NOAA [Data set], https://psl.noaa.gov/data/obs/datadisplay/, access: 28 June 2021.

Rawls, W., P. Yates, and L. Asmussen, Calibration of selected infiltration equations for the Georgia Coastal Plain, Rep. USDA-ARS-S113, 110 pp., Agric. Res. Serv., Beltsville, Md., 1976

Ray RL, Jacobs JM (2007) Relationships among remotely sensed soil moisture, precipitation and landslide events Nat Hazards 43:211-222 doi:10.1007/s11069-006-9095-9

Reichle, R. H., de Lannoy, G. J. M., Liu, Q., Ardizzone, J. V., Colliander, A., Conaty, A., Crow, W., Jackson, T. J., Jones, L. A., Kimball, J. S., Koster, R. D., Mahanama, S. P., Smith, E. B., Berg, A., Bircher, S., Bosch, D., Caldwell, T. G., Cosh, M., González-Zamora, Á., Holifield Collins, C. D., Jensen, K. H., Livingston, S., Lopez-Baeza, E., Martínez-Fernández, J., McNairn, H., Moghaddam, M., Pacheco, A., Pellarin, T., Prueger, J., Rowlandson, T., Seyfried, M., Starks, P., Su, Z., Thibeault, M., van der Velde, R., Walker, J., Wu, X., & Zeng, Y. (2017). Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements. Journal of Hydrometeorology, 18(10), 2621–2645. https://doi.org/10.1175/jhm-d-17-0063.1

Reichle, R., G. De Lannoy, R. D. Koster, W. T. Crow, J. S. Kimball, and Q. Liu. (2021). SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Analysis Update, Version 6 [Data Set]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/6P2EV47VMYPC. Date Accessed 12-02-2021.

Reichle, R., G. De Lannoy, R. D. Koster, W. T. Crow, J. S. Kimball, Q. Liu, and M. Bechtold. (2022). SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Analysis Update, Version 7 [Data Set]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/LWJ6TF5SZRG3. Date Accessed 12-27-2022.

Rosenzweig C, Tubiello FN, Goldberg R, Mills E, Bloomfield J (2002) Increased crop damage in the US from excess precipitation under climate change vol 12.

Schaefli B, Gupta HV (2007) Do Nash values have value? Hydrological Processes 21:2075-2080 doi:https://doi.org/10.1002/hyp.6825

Seneviratne SI et al. (2010) Investigating soil moisture–climate interactions in a changing climate: A review Earth-Science Reviews 99:125-161 doi:https://doi.org/10.1016/j.earscirev.2010.02.004

Seneviratne SI et al. (2013) Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment 40:5212-5217 doi:https://doi.org/10.1002/grl.50956

Shangguan W, Dai Y, Duan Q, Liu B, Yuan H (2014) A global soil data set for earth system modeling 6:249-263 doi:https://doi.org/10.1002/2013MS000293

Silver M, Karnieli A, Ginat H, Meiri E, Fredj E (2017) An innovative method for determining hydrological calibration parameters for the WRF-Hydro model in arid regions Environmental Modelling & Software 91:47-69 doi:https://doi.org/10.1016/j.envsoft.2017.01.010

Sofokleous I, Bruggeman A, Camera C, Eliades M (2022) Grid-based calibration of the WRF-Hydro with Noah-MP model with improved groundwater and transpiration process equations Journal of Hydrology:128991 doi:https://doi.org/10.1016/j.jhydrol.2022.128991

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey [Data set]. Available online at https://websoilsurvey.nrcs.usda.gov/. Accessed 12/13/2021.

Sweeney D, Robertson P (1979) A fundamental approach to slope stability problems in Hong Kong. Hong Kong Engineer. October,

Tang H, McGuire LA, Rengers FK, Kean JW, Staley DM, Smith JBJGRL (2019) Developing and testing physically based triggering thresholds for runoff-generated debris flows 46:8830-8839

Tavakol A, Rahmani V, Quiring SM, Kumar SV (2019) Evaluation analysis of NASA SMAP L3 and L4 and SPoRT-LIS soil moisture data in the United States Remote Sensing of Environment 229:234-246 doi:https://doi.org/10.1016/j.rse.2019.05.006

U.S. Geological Survey (USGS) (2016). National Water Information System data, USGS Water Data for the Nation, USGS [Data set], https://doi.org/10.5066/F7P55KJN.

Vergopolan N, Chaney NW, Beck HE, Pan M, Sheffield J, Chan S, Wood EF (2020) Combining hyper-resolution land surface modeling with SMAP brightness temperatures to obtain 30-m soil moisture estimates Remote Sensing of Environment 242:111740 doi:https://doi.org/10.1016/j.rse.2020.111740

Wagner S, Fersch B, Yuan F, Yu Z, Kunstmann HJWRR (2016) Fully coupled atmospheric-hydrological modeling at regional and long-term scales: Development, application, and analysis of WRF-HMS 52:3187-3211

Wang G, Garcia D, Liu Y, de Jeu R, Johannes Dolman A (2012) A three-dimensional gap filling method for large geophysical datasets: Application to global satellite soil moisture observations Environmental Modelling & Software 30:139-142 doi:https://doi.org/10.1016/j.envsoft.2011.10.015

Wang J, Wang C, Rao V, Orr A, Yan E, Kotamarthi R (2019) A parallel workflow implementation for PEST version 13.6 in high-performance computing for WRF-Hydro version 5.0: a case study over the midwestern United States Geosci Model Dev 12:3523-3539 doi:10.5194/gmd-12-3523-2019

Western Regional Climate Center (2021). Current WRCC Weather Data Plots [Data set]. Retrieved from: https://wrcc.dri.edu/weather/ access: 12/06/2021.

Xia Y, Ek MB, Wu Y, Ford T, Quiring SM (2015) Comparison of NLDAS-2 Simulated and NASMD Observed Daily Soil Moisture. Part I: Comparison and Analysis J Hydrometeor 16:1962-1980 doi:10.1175/JHM-D-14-0096.1

Xia Y et al. (2012) Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow J Geophys Res Atm 117 doi:https://doi.org/10.1029/2011JD016051

Xia Y et al. (2014) Evaluation of multi-model simulated soil moisture in NLDAS-2 Journal of Hydrology 512:107-125 doi:https://doi.org/10.1016/j.jhydrol.2014.02.027

Xu L, Abbaszadeh P, Moradkhani H, Chen N, Zhang XJRSoE (2020) Continental drought monitoring using satellite soil moisture, data assimilation and an integrated drought index 250:112028

Yeh TC, Wetherald RT, Manabe S (1984) The Effect of Soil Moisture on the Short-Term Climate and Hydrology Change—A Numerical Experiment Mon Weather Rev 112:474-490 doi:10.1175/1520-0493(1984)112<0474:TEOSMO>2.0.CO;2

Yin D, Xue ZG, Gochis DJ, Yu W, Morales M, Rafieeinasab A (2020) A Process-Based, Fully Distributed Soil Erosion and Sediment Transport Model for WRF-Hydro 12:1840

Yu G, Wright DB, Li Z (2020) The Upper Tail of Precipitation in Convection-Permitting Regional Climate Models and Their Utility in Nonstationary Rainfall and Flood Frequency Analysis 8:e2020EF001613 doi:https://doi.org/10.1029/2020EF001613

Yucel I, Onen A, Yilmaz KK, Gochis DJ (2015) Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall Journal of Hydrology 523:49-66 doi:https://doi.org/10.1016/j.jhydrol.2015.01.042

Yuste J, Baldocchi DD, Gershenson A, Goldstein A, Misson L, Wong S (2007) Microbial soil respiration and its dependency on carbon inputs, soil temperature, and moisture vol 13.

Zhang, J. et al. (2011). National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. Bull. Amer. Meteor. Soc., 92, 1321–1338, https://doi.org/10.1175/2011BAMS-D-11-00047.1.

Zhang, J., Qi, Y., Langston, C., Kaney, B., & Howard, K. (2014). A real-time algorithm for merging radar QPEs with rain gauge observations and orographic precipitation climatology. J. Hydrometeor., 15, 1794–1809, https://doi.org/10.1175/JHM-D-13-0163.1.

Zhang, J. et al. (2016). Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. Bull. Amer. Meteor. Soc., 97, 621–638, https://doi.org/10.1175/BAMS-D-14-00174.1.

Zhuo L, Han D, Dai Q, Islam T, Srivastava PK (2015) Appraisal of NLDAS-2 Multi-Model Simulated Soil Moistures for Hydrological Modelling Water Resources Management 29:3503-3517 doi:10.1007/s11269-015-1011-1