

## Article

# Data Assimilation of Satellite-Based Soil Moisture into a Distributed Hydrological Model for Streamflow Predictions

Navid Jadidoleslam <sup>\*</sup>, Ricardo Mantilla  and Witold F. Krajewski 

Iowa Flood Center, The University of Iowa, Iowa City, IA 52242, USA; ricardo-mantilla@uiowa.edu (R.M.); witold-krajewski@uiowa.edu (W.F.K.)

\* Correspondence: navid-jadidoleslam@uiowa.edu

**Abstract:** The authors examine the impact of assimilating satellite-based soil moisture estimates on real-time streamflow predictions made by the distributed hydrologic model HLM. They use SMAP (Soil Moisture Active Passive) and SMOS (Soil Moisture Ocean Salinity) data in an agricultural region of the state of Iowa in the central U.S. They explore three different strategies for updating model soil moisture states using satellite-based soil moisture observations. The first is a “hard update” method equivalent to replacing the model soil moisture with satellite observed soil moisture. The second is Ensemble Kalman Filter (EnKF) to update the model soil moisture, accounting for modeling and observational errors. The third strategy introduces a time-dependent error variance model of satellite-based soil moisture observations for perturbation of EnKF. The study compares streamflow predictions with 131 USGS gauge observations for four years (2015–2018). The results indicate that assimilating satellite-based soil moisture using EnKF reduces predicted peak error compared to that from the open-loop and hard update data assimilation. Furthermore, the inclusion of the time-dependent error variance model in EnKF improves overall streamflow prediction performance. Implications of the study are useful for the application of satellite soil moisture for operational real-time streamflow forecasting.



**Citation:** Jadidoleslam, N.; Mantilla, R.; Krajewski, W.F. Data Assimilation of Satellite-Based Soil Moisture into a Distributed Hydrological Model for Streamflow Predictions. *Hydrology* **2021**, *8*, 52. <https://doi.org/10.3390/hydrology8010052>

Academic Editor: Lorena Liuzzo

Received: 28 February 2021

Accepted: 19 March 2021

Published: 20 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** satellite soil moisture; data assimilation; Ensemble Kalman Filter; flood prediction

## 1. Introduction

Accurate rainfall-runoff partitioning is one of the most critical factors in predicting the magnitude of streamflow fluctuations. Soil moisture is the main state variable in hydrologic models that determines estimated runoff magnitudes. However, the value of the hydrologic model states at any point in time are subject to uncertainties, as they encode the history of all the variables and hydrometeorological input forcings that ultimately determine them. Additionally, epistemic decisions such as model structure, model parameters, closure equations, initial conditions, among others (e.g., [1–3]), also play a factor in determining the value on state variables everywhere and every time as flow equations are integrated. Regional-scale satellite-based soil moisture observations provide new research opportunities for validation and correction of predicted soil moisture states in hydrologic models.

Previous data-driven modeling studies have shown that satellite-based soil moisture estimations provide useful information on runoff production. For example, Crow et al. [4] found a significant relationship between antecedent SMAP soil moisture and runoff ratio for low vegetation regions. Jadidoleslam et al. [5] followed-up [4] and showed SMAP satellite-based soil moisture provides important information on event-scale runoff production in a heavily agricultural region, where satellite-based soil moisture retrievals are more sensitive to Vegetation Optical Depth (VOD) and soil surface roughness in agricultural regions [6,7].

Other studies have explored data assimilation of field sensor and satellite-based soil moisture observations in hydrologic models and their potential on streamflow predictions.

Examples range from simple replacement of observed soil moisture with the model (e.g., [8]) to ensemble-based methods such as Ensemble Kalman Filter (e.g., [9–14]) and evolutionary algorithms (e.g., [15–17]).

More recently, Mao et al. [13] developed a diagnostic framework for assessing the impact of satellite-based soil moisture data assimilation on daily streamflow predictions. Abbaszadeh et al. [18] used the WRF-Hydro model [19] to assimilate SMAP satellite-based soil moisture and daily streamflow observations during hurricane Harvey in 2017. A recent survey of data assimilation techniques and their application in earth science and hydrometeorology is presented in [20,21].

Previous studies provide useful insights on the satellite-based soil moisture data assimilation and its effect on streamflow predictions. These studies adopt calibration as part of hydrologic model simulations. However, as highlighted by Clark et al. [22], “classical” calibration approaches could lead to parameters that do not represent realistic physical values and could change the sensitivity of a given model state to an input forcing. Furthermore, these studies are generally conducted over a small number of basins with a low vegetation cover where satellite-based soil moisture is less prone to retrieval uncertainties.

In this study, our goal is to explore the potential of SMAP and SMOS satellite-based soil moisture data assimilation to improve streamflow predictions using a calibration-free approach following [23]. We conduct our study using a state-wide distributed operational hydrologic model developed at the Iowa Flood Center [24]. We assess the impact of satellite-based soil moisture products on streamflow predictions in a large domain dominated by cropland. For this purpose, we estimate the potential error variance in satellite-based soil moisture using field sensor soil moisture observations. Then, we use three different experiments to examine the impact of satellite-based soil moisture data assimilation on streamflow predictions. Finally, we assess the performance of data assimilation experiments and satellite soil moisture products on streamflow predictions.

This study is organized as follows; first, we describe the study region and data; then, we provide details of our methodology and experiments; and finally, we present results and follow a discussion on their relevance to previous studies and implications on the streamflow predictions.

## 2. Study Region and Data

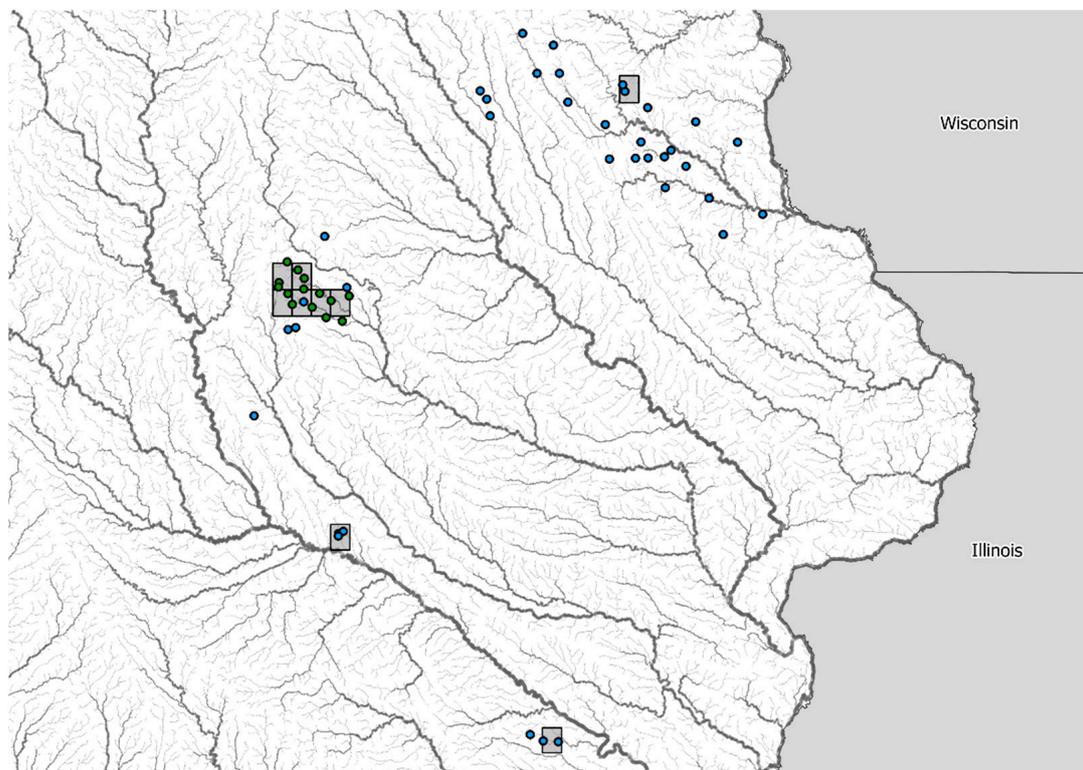
Our study region is the state of Iowa, United States, located in a heavily agricultural region with more than 70 percent of its surface area covered by cropland (mainly corn and soybean) [25]. During the past three decades, devastating floods have cost Iowans about \$18 billion in damages to properties and crops [26].

The data available for our study region includes field sensor soil moisture, satellite-based soil moisture, radar rainfall, and USGS (United States Geological Service) streamflow observations. We use four years of data (2015–2018) from 1 April 1 to 31 October. We exclude data for the cold season to avoid the frozen surface’s impacts on the field sensors and satellite-based soil moisture observations.

We use hourly field sensor soil moisture observations at 5 cm depth from Iowa Flood Center (IFC) and USDA-ARS networks, shown as blue and green circles in Figure 1, respectively. ARS (Agricultural Research Service) sensors are located at the South Fork watershed in the north-central part of the state of Iowa. South Fork is also one of the SMAP Core Validations Sites (CVS) of the SMAPVEX16 field campaign [27].

For satellite-based soil moisture, we use Enhanced SMAP Level 3 Version 2 (L3\_P\_E) soil moisture [28] and SMOS satellite-based soil moisture [29]. Enhanced SMAP data is provided on a 9-km EASE-Grid version 2 [30], and SMOS is posted on the ISEA grid [31] with an approximate resolution of 43 km in space.

For rainfall forcing of the hydrologic model, we use rain gauge bias-corrected Stage IV hourly radar-based rainfall [32] posted on a grid with approximately 4 km resolution [33]. We use climatological estimates from North American Land Data Assimilation (NLDAS) [34] for Evapotranspiration (ET) forcing.



**Figure 1.** Soil moisture sensors from IFC (blue), USDA-ARS (green), and SMAP pixels collocated with at least two field sensors over the State of Iowa, U.S.

We use hourly (instantaneous) streamflow observations at 131 USGS gauge locations at the state of Iowa obtained from USGS National Water Information System [35].

### 3. Methods

#### 3.1. Experimental Setup

We use the Hillslope Link Model (HLM) implemented at the Iowa Flood center [24] for hydrologic model predictions over the state of Iowa. HLM decomposes the landscape into hillslopes and links (i.e., river segments) [36] that represent the hillslope processes and channel flow, respectively. Each hillslope-link consists of four states: channel flow, ponded surface water depth, water storages at the top layer, and subsurface. The mass conservation equations for water flux exchange between these storages are defined by nonlinear ordinary differential equations and solved by a parallelized implementation of Runge-Kutta methods [37]. For further details of the model structure and formulation, we refer readers to [38].

We conduct numerical model simulations using an a-priori determined parameter set that applies to the study region. Therefore, the model simulations are not calibrated to any specific model input or hydrometeorological forcing or to match observations in any location in particular. Furthermore, we use the same initial conditions for different experiments for the same year. This allows us to assess the impact of different data assimilation approach on streamflow predictions and create a comparable set of simulations for each year.

We use three strategies to incorporate the satellite-based soil moisture into the hydrologic model: First, we replace the top-layer soil moisture with the satellite-based soil moisture. This method is also called “hard update” (e.g., [8]). Second, we use Ensemble Kalman Filter (EnKF) [39] to account for the potential errors in the model and satellite-based soil moisture. Finally, we include time-dependent error variances for satellite-based soil moisture in EnKF. We use open-loop streamflow predictions as the baseline for hydrologic model performance evaluations. We describe our experiments as follows:

1. Open-loop: We conduct the hydrologic model simulations by integrating the HLM equations using Stage IV rainfall as forcing, and region averaged monthly ET data. These simulations are the baseline for comparisons to other experiments.
2. Hard update: In this scheme, we replace the hydrologic model's top-layer soil moisture with the satellite-based soil moisture estimations at every satellite observation time where soil moisture is available.
3. Ensemble Kalman Filter (EnKF): Model and observation errors are incorporated in EnKF by perturbations with a zero mean and estimated standard deviations. We estimated the error standard deviations for model top-layer and satellite-based soil moisture products using soil moisture field sensors. At each satellite observation time, we update the initial soil moisture in the model with perturbations that follow constant variances for the model and satellite-based soil moisture. We provide more details of the EnKF routine in Section 3.2.
4. EnKF with time-dependent observational error variance (EnKFV): We conduct a similar numerical simulation as EnKF, but we perturb the satellite-based soil moisture observations based on time-dependent error variances calculated for each satellite-based soil moisture product (Section 3.3).

Our simulations span over four years (2015–2018) with contrasting hydroclimatological conditions, which provide more insights into the performance of streamflow prediction after data assimilation. We characterize 2015 as a typical hydrologic year, 2017 as a dry year, and 2016 and 2018 as wet years across the state.

### 3.2. Ensemble Kalman Filter

Ensemble Kalman Filter (EnKF) was proposed by Evensen [40] and later clarified by [41]. Data assimilation (DA), in our study, aims to determine the information about the hydrologic model's state (e.g., soil moisture) given the satellite-based soil moisture observation and update the model state accordingly. EnKF uses an ensemble of model realizations to estimate the covariance of the state vector. In this study, we only update the model's top-layer soil moisture at every observation time. We describe the EnKF steps and formulation as follows:

Let  $X$  be the ensemble predictions of soil moisture at time  $t$ . We define the perturbed model predictions as  $\hat{U}$  as

$$\hat{U} = X + \eta \quad (1)$$

where  $\eta \sim N(0, \sigma)$  is prediction error following a gaussian distribution with zero mean and  $\sigma$  standard deviation. The ensemble mean of the predictions is calculated as

$$\hat{m} = \frac{\sum \hat{U}}{N} \quad (2)$$

where  $N$  is the number of ensemble members and  $N = 30$ . Let  $\hat{C}$  be the covariance of model predictions defined as

$$\hat{C} = \frac{(\hat{U} - \hat{m})(\hat{U} - \hat{m})^T}{N - 1}. \quad (3)$$

Let  $s(t)$  be satellite-based soil moisture observation at time  $t$  at a given satellite pixel. We perturb satellite-based soil moisture observation with a gaussian noise  $\xi_0$  following a zero mean and estimated variance. In EnKF, the measure of mismatch between predicted state and the observations is called innovation  $d$  defined as

$$d = s(t) + \xi_0 - H\hat{U} \quad (4)$$

where  $H$  is the measurement operator and in our case  $H = 1$ .

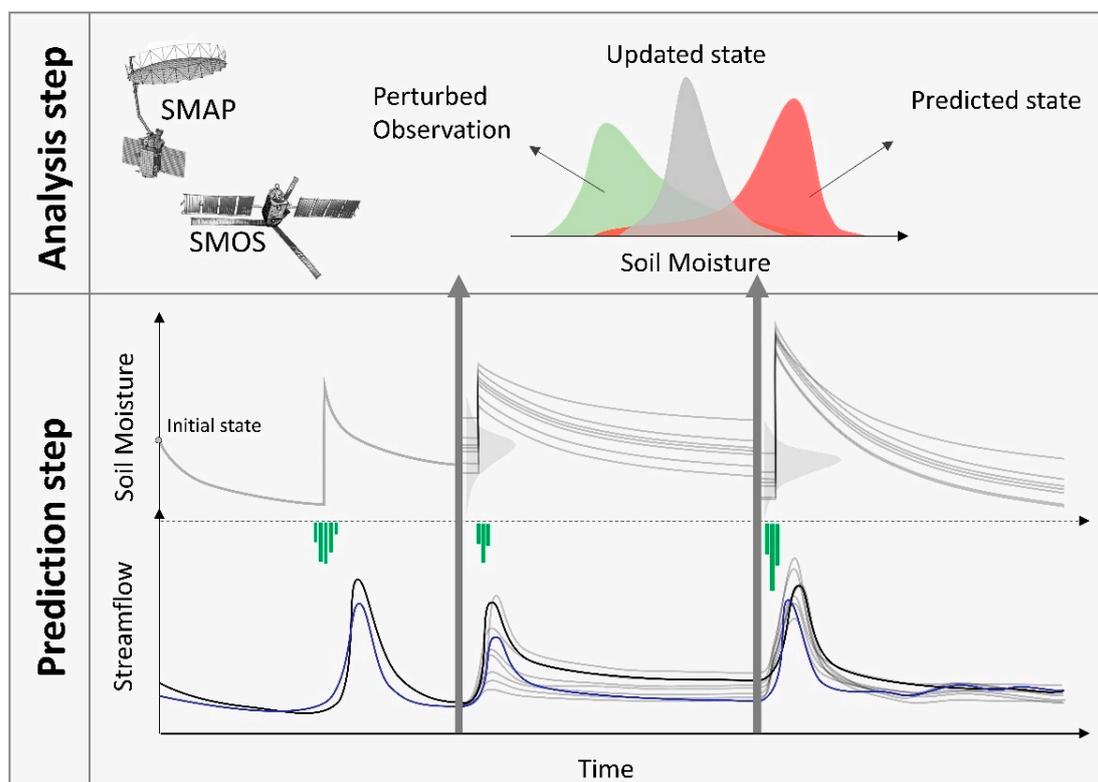
For an optimal update in EnKF, we calculate the matrix of weights  $K$  (Kalman gain) given by

$$K = \hat{C}H^T(H\hat{C}H^T + \Gamma)^{-1} \quad (5)$$

where  $\Gamma$  is covariance reconstructed from perturbed satellite-based soil moisture observation. Finally, the updated top-layer state vector  $U$  is calculated by

$$U = \hat{U} + Kd \quad (6)$$

and used as the initial condition for the ensemble members. The new initial conditions are evolved through the hydrologic model until new satellite observation becomes available. Figure 2 shows an illustration of the EnKF assimilation framework used in this study.



**Figure 2.** An illustration of Ensemble Kalman Filter for satellite-based soil moisture data assimilation in this study.

Hydrologic model states such as soil moisture are initialized with an initial state (e.g., soil moisture) on 1 April. The prediction step evolves the hydrologic model's initial states until a new satellite-based soil moisture observation becomes available, and the analysis step is triggered. In the analysis step, we use EnKF to update the estimated soil moisture by perturbing soil moisture observation from SMAP or SMOS satellites and estimated soil moisture variance corresponding to observations and model soil moisture. After that, new initial conditions (updated state) are evolved in the hydrologic model.

### 3.3. Error Standard Deviation Estimation

We estimate satellite-based soil moisture error standard deviation using the field sensor-average soil moisture average of the field sensors collocated with the SMAP or SMOS pixels. We select the pixels that collocate with at least two field sensors. For example, Figure 1 shows selected SMAP pixels used for comparisons with the sensor-average soil moisture. We compare the sensor-average and satellite-based soil moisture from 2015 to 2018 for each month (April to November).

Figures A1 and A2 show comparisons between the soil moisture average of the collocated field sensors with SMAP and SMOS soil moisture. As shown, the agreement between sensor-average soil moisture and SMAP satellite estimation varies each month.

We calculate the standard deviation of the difference between satellite-based and sensor-averaged soil moisture over four years (2015–2018) by selecting the soil moisture data within a month time window on each day of the year. Figure 3 shows the calculated error standard deviation for SMAP and SMOS satellite-based soil moisture between April and November. SMAP soil moisture error variance is lower than SMOS during April and from August to November. This analysis provides a basis for incorporating the potential errors from satellite-based soil moisture into our data assimilation framework.



**Figure 3.** Error standard deviation estimations [ $\text{cm}^3/\text{cm}^3$ ] for SMAP and SMOS satellite-based soil moisture.

We include estimated error variance in EnKF and EnKFV as the perturbations to satellite-based soil moisture satellite-based observations. Furthermore, we estimate model top-layer soil moisture error variance using the soil moisture sensors located in different hillslopes across the study region. The error variance for model soil moisture is given as  $0.05 (\text{cm}^3/\text{cm}^3)$ .

### 3.4. Evaluation Metrics

We evaluate streamflow predictions for the four simulation schemes described in Section 3.1. We use the mean of the streamflow ensemble predictions using EnKF and EnKFV schemes for performance evaluations.

We assess the performance of the streamflow predictions with Kling-Gupta Efficiency (KGE) [42] and Peak Difference Ratio (PDR) for each year. KGE is defined as

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (7)$$

where  $r$  is the Pearson correlation coefficient between observed streamflow and modeled streamflow;  $\alpha$  represents the standard deviation ratio of modeled and observed streamflow; and  $\beta$  is the ratio between the mean of the simulations and the mean of the observations. Peak Difference Ratio (PDR) is defined as

$$PDR = \frac{Q_{p_{sim}} - Q_{p_{obs}}}{Q_{p_{obs}}} \quad (8)$$

where  $Q_{p_{sim}}$  and  $Q_{p_{obs}}$  are the peak of modeled and observed streamflow.  $PDR$  is useful to assess the streamflow prediction performance in capturing the observed streamflow peak of the year. The positive and negative value of  $PDR$  represents overestimation and underestimation of the observed peak.

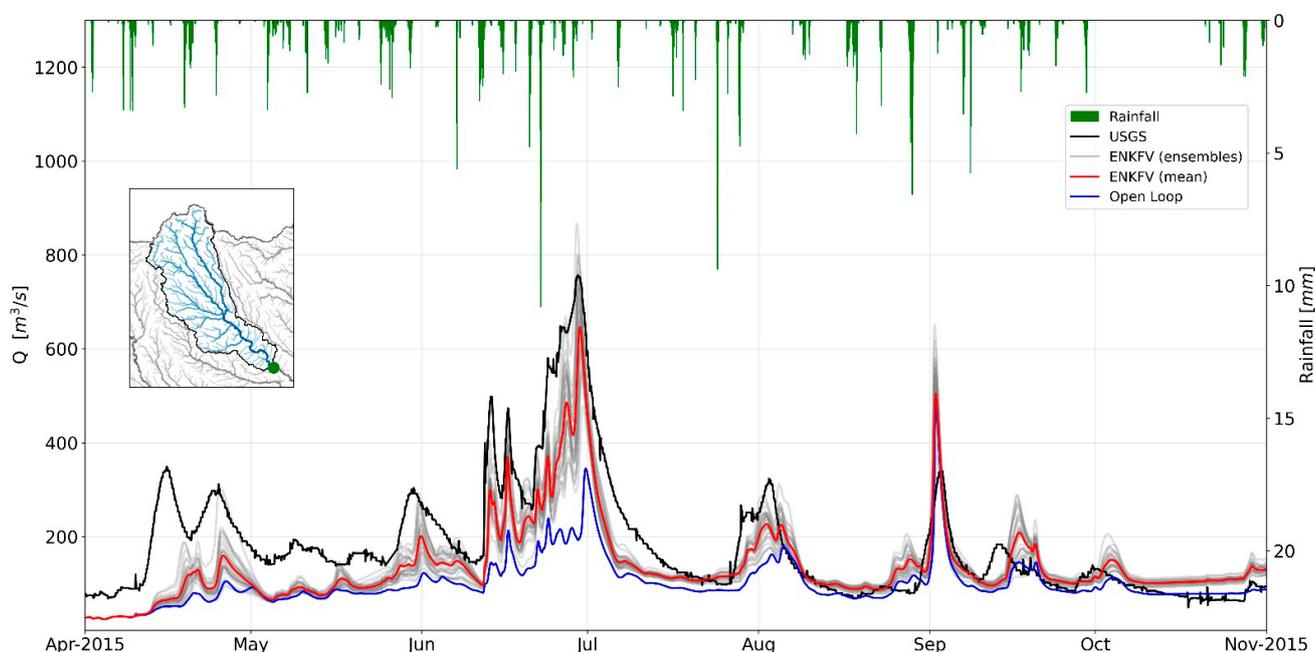
To evaluate the overall prediction performance for different experiments, described in Section 3.1, we construct Kernel Density Estimations ( $KDE$ ) using a gaussian kernel for each performance evaluation metric. We present the streamflow prediction performance for different experiments and the baseline simulations to assess the impact of different data assimilation schemes.

#### 4. Results

In this section, we present results from data assimilation experiments described in the previous section. First, we provide insights on the impact of satellite-based soil moisture data assimilation using a specific example. In the following two subsections, we present streamflow prediction performance maps for hard update and Ensemble Kalman Filter with time-dependent variance (EnKFV). We also show streamflow prediction performance metrics and compare different experiments mentioned in Section 3.1.

We used Hydrovise [43] to share our streamflow prediction results for the study region in an interactive web-based environment. Streamflow prediction time-series and the evaluation metrics for each experiment can be visualized from <http://hydrovise.com/app/?config=da2021/config.json> (accessed on 19 March 2021).

Figure 4 shows example streamflow predictions from an open-loop model and EnKF data assimilation scheme that includes the time-varying error variances (EnKFV) for SMAP satellite-based soil moisture observations. Streamflow data correspond to USGS gauge location at Cedar River at Cedar Rapids, Iowa. We selected this case because the open-loop model exhibited a clear underestimation of runoff production, which is connected to low values of predicted soil moisture.



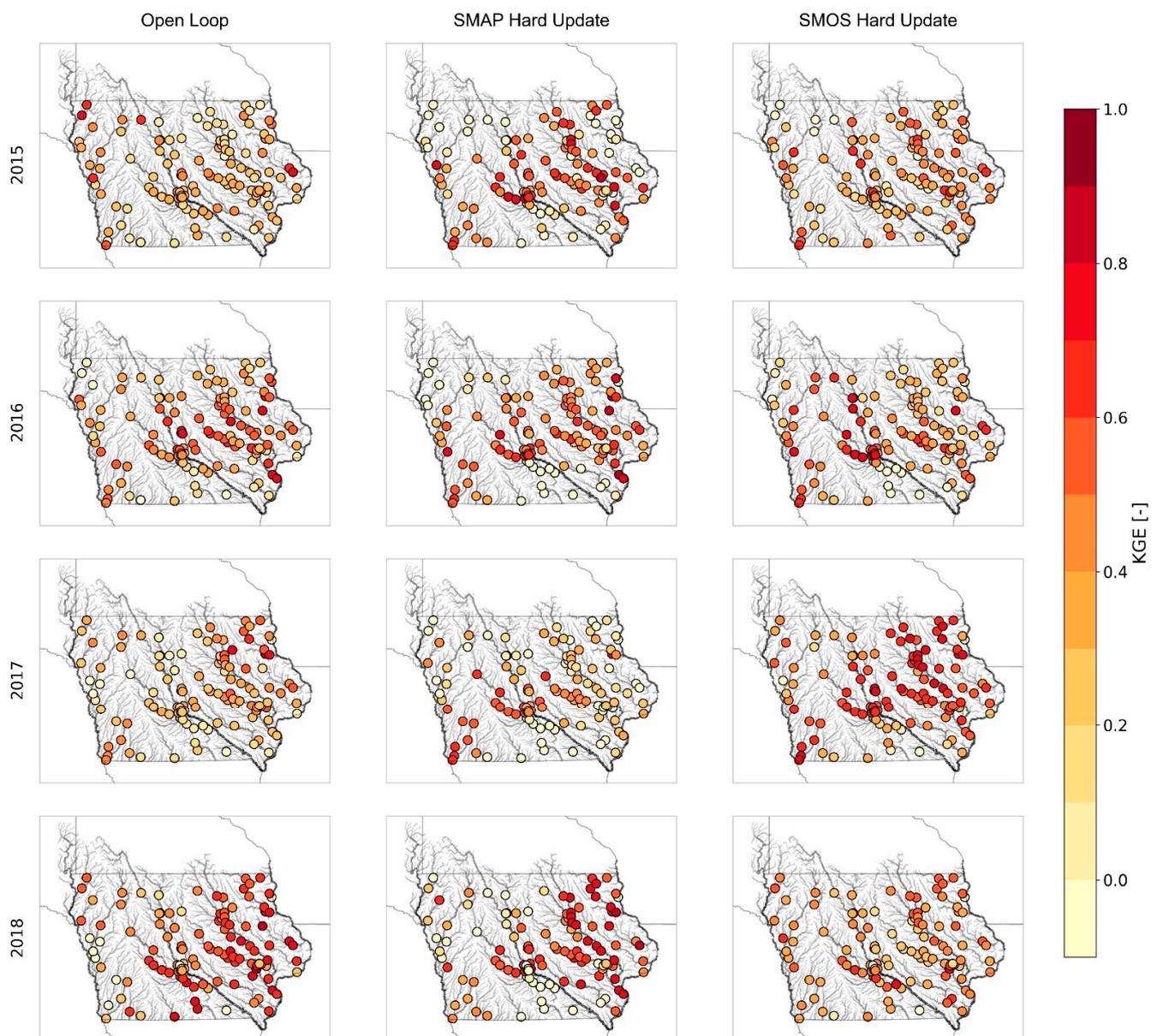
**Figure 4.** Time-series plot for USGS streamflow [ $m^3/s$ ] observations (black), Open-loop (blue), streamflow ensemble predictions from EnKFV (gray), and EnKFV ensemble mean (red) at Cedar River at Cedar Rapids (USGS code: CIDI4-05464500). Basin-averaged Stage IV radar rainfall [mm] time-series are shown in green.

As mentioned in Section 3, we use the same initial conditions for the model states on 1 April. Figure 4 indicates that streamflow ensembles exhibit higher variability after a relatively large rainfall event during the second week of April. This example highlights

that the impact of satellite-based soil moisture on streamflow predictions is pronounced after rainfall-runoff events. Ensemble predictions from EnKFV show higher variability during the year's peak event (June–July) and capture streamflow observations.

#### 4.1. Hard Update

Figure 5 shows streamflow prediction performance in KGE before and after satellite-based soil moisture hard update with model top-layer soil moisture. Hard update model predictions using SMOS data indicate higher KGE values compared to the open-loop model, and SMAP hard update for 2017 that was a dry year. SMAP satellite-based soil moisture assimilation using hard update results in higher KGE values for streamflow predictions in other years.

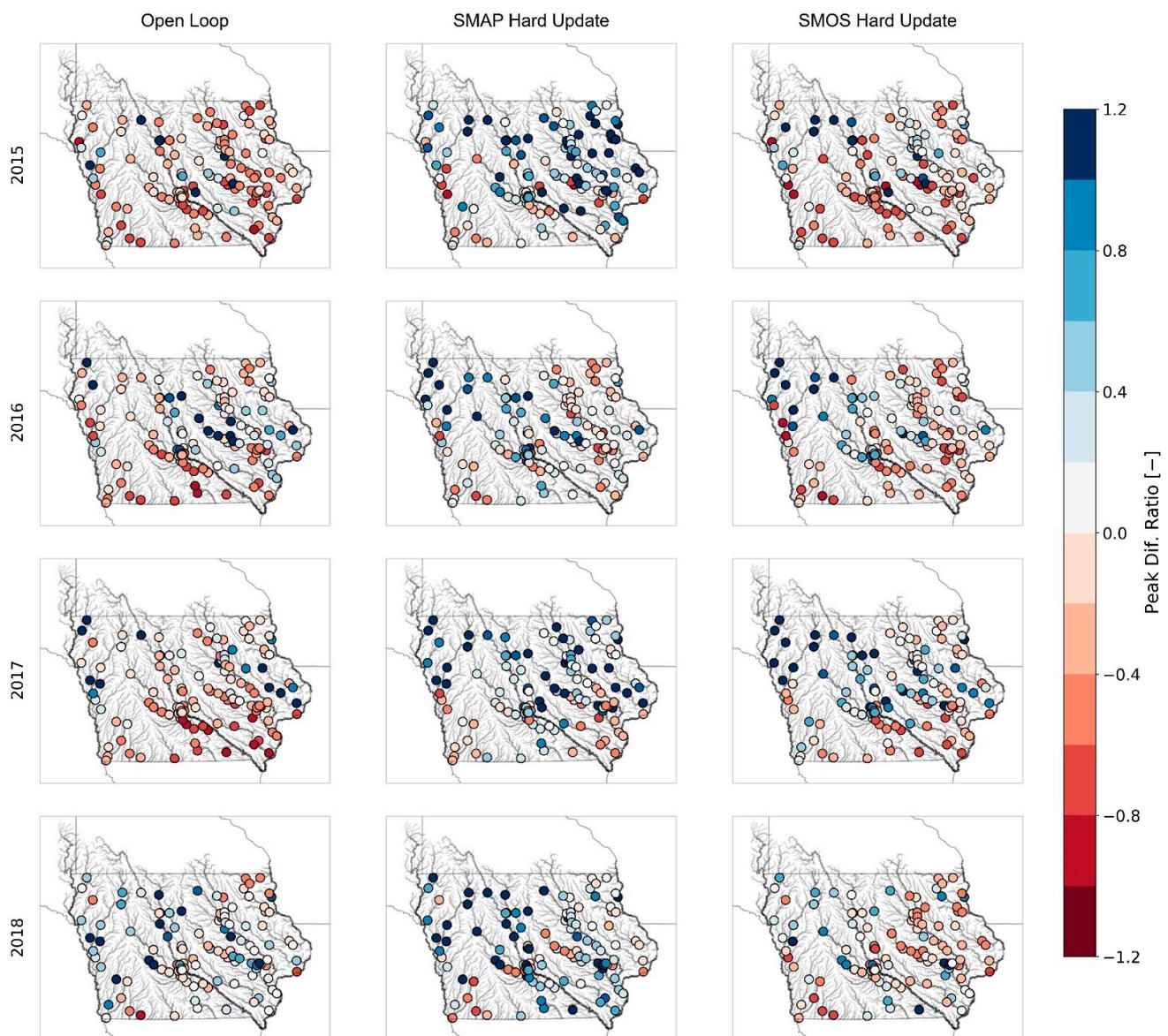


**Figure 5.** Map of streamflow prediction performance in KGE for open-loop and soil moisture data assimilation using hard update for SMAP and SMOS for 2015 to 2018.

Hard update slightly increases streamflow prediction performance in the central and eastern part of the study region, while it reduces the prediction performance in terms of

KGE for most of the locations. More specifically, the KGE values for streamflow predictions are lower after hard update at locations in the north-west (e.g., Little Sioux River) and southern parts of the state (e.g., Chariton River).

Figure 6 shows model streamflow prediction performance in terms of Peak Difference Ratio for open-loop and satellite-based soil moisture hard update for our study period (2015–2018). In this figure, blue and red colors represent overestimation and underestimation of annual peak streamflow relative to the observed annual peak. Lighter colors indicate a smaller peak difference between observations and simulations. Moreover, peaks that do not coincide within 48 h temporal window were excluded from the analysis to compare the peaks that occurred in the same rainfall-runoff event.



**Figure 6.** Map of streamflow prediction performance in terms of Peak Difference Ratio for Open-loop, SMAP Hard Update, and SMOS Hard Update from 2015 to 2018.

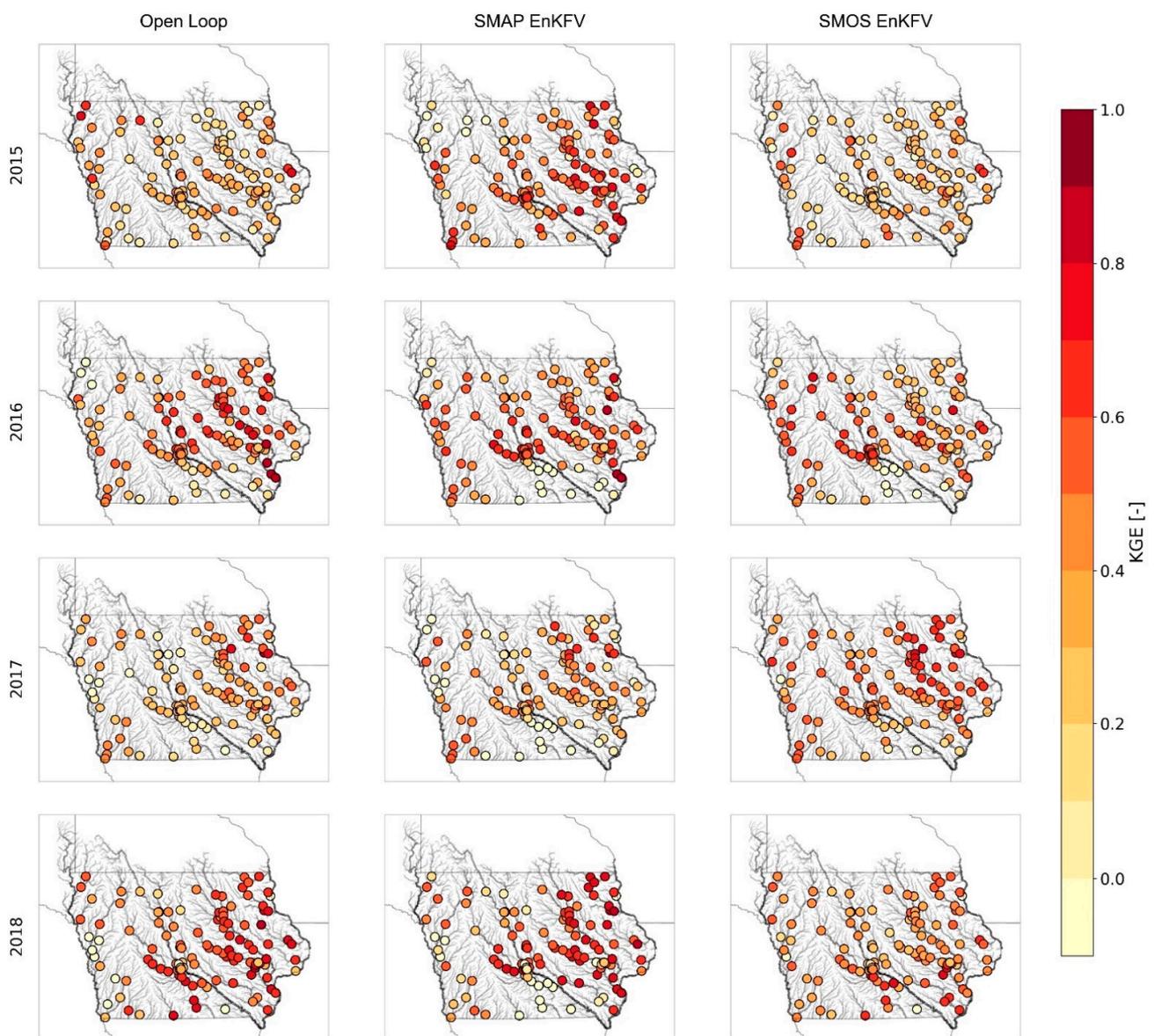
Results from hard update experiment suggest that magnitude of the simulated peaks increases after SMAP hard update. On the other hand, SMOS hard update reduces the peaks in the eastern part of the study region, such as the Cedar River basin.

The open-loop model's predicted peaks are lower than the observed peak of the year, while SMAP soil moisture hard update overestimates annual peaks in the study domain's

western part. SMOS hard update reduces peak differences in the eastern part of the state while it increases the peak overestimations in the western part of the study domain from 2015 to 2017.

#### 4.2. EnKFV

Maps in Figure 7 show KGE performance metrics for streamflow predictions with open-loop and EnKFV using SMAP and SMOS soil moisture for 2015–2018. The streamflow prediction performance (KGE) shows improvement in most streamflow gauge locations over the study domain. At a few USGS gauge locations in the southern part, streamflow predictions show relatively lower performance after data assimilation using EnKFV. These locations are predominantly basins with smaller drainage areas. Higher KGE values are obtained in the eastern part of the state.

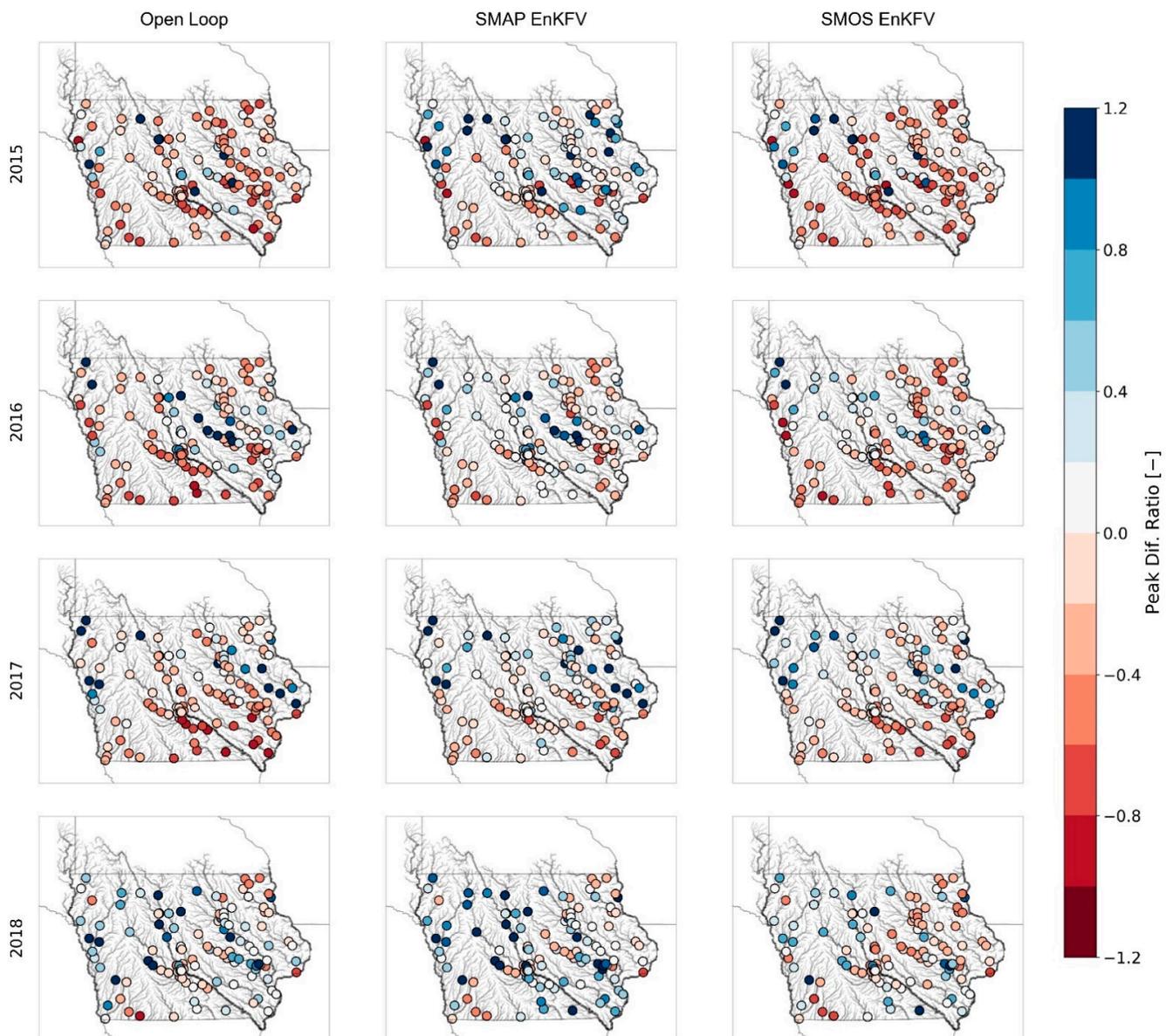


**Figure 7.** Map of streamflow prediction performance in KGE for Open-loop, SMAP EnKFV, and SMOS EnKFV from 2015 to 2018.

SMAP soil moisture assimilation using EnKFV shows larger improvements than SMOS assimilation, specifically in the eastern part of the state. The highest prediction

performances are obtained from SMOS data assimilation during a dry year (2017) and SMAP data assimilation in the other three years.

Figure 8 shows maps of predicted streamflow peak performance in terms of Peak Difference Ratio for simulations conducted by open-loop model and EnKFV with SMAP and SMOS satellite-based soil moisture data assimilation.



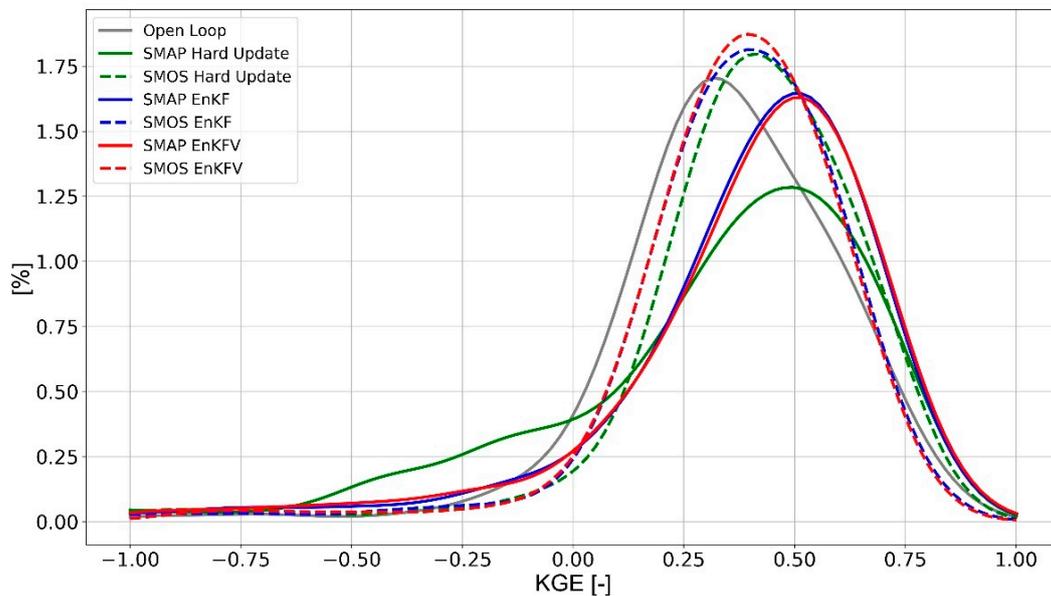
**Figure 8.** Map of streamflow prediction performance in terms of Peak Difference Ratio for Open-loop, SMAP EnKFV, and SMOS EnKFV from 2015 to 2018.

Compared to streamflow predictions with SMOS EnKFV, SMAP soil moisture assimilation with EnKFV achieves a higher reduction in streamflow peak differences in the eastern part of the study domain. Furthermore, SMOS EnKFV improves the Peak Difference Ratio in the western part of the domain.

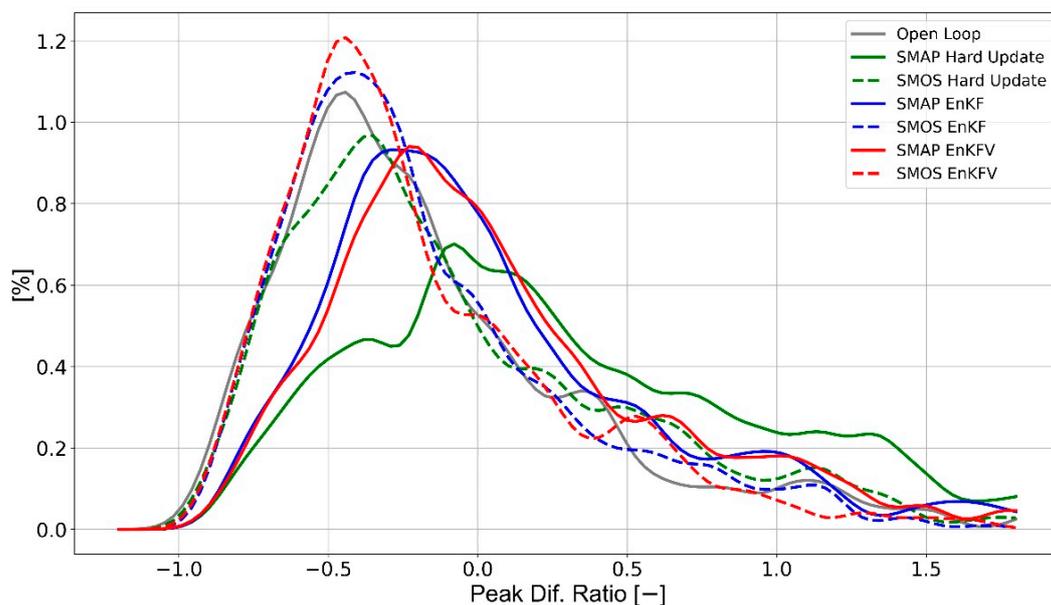
#### 4.3. Prediction Performance Summary

This section presents and compares the overall streamflow prediction performance for open-loop, hard update, EnKF, and EnKFV. Figures 9 and 10 present the Kernel Density

Estimation for KGE and Peak Difference Ratio performance metrics for 131 USGS gauge locations and four years (2015–2018).



**Figure 9.** Kernel density estimations of KGE for streamflow predictions of open-loop, hard update, EnKF, and EnKFV for 131 USGS gauge locations during four years (2015–2018).



**Figure 10.** Kernel density estimations of peak difference ratio for streamflow predictions of open-loop, hard update, EnKF, and EnKFV for 131 USGS gauge locations during four years (2015–2018).

As shown in Figure 9, SMAP hard update shows improvements to the KGE of the predictions, but it also increases the percentage of station-years with lower KGE values. SMOS hard update provides higher KGE values compared to SMAP hard update. The median value of KGE for the open-loop model, SMAP hard update, and SMOS hard update is 0.34, 0.37, and 0.42, respectively.

Accounting for potential observation errors in SMAP soil moisture data assimilation (EnKF) reduces the number of stations with lower KGE values and increases the percentage of station-years with higher KGE values. The model predictions with the EnKFV

scheme results in a slight increase in the KGE for SMAP satellite-based soil moisture data assimilation (0.45). Three schemes for SMOS soil moisture data assimilation show similar improvements in KGE for streamflow predictions compared to the open-loop model.

Figure 10 compares the kernel density estimations for peak difference ratio for open-loop model run and data assimilation experiments.

The open-loop model generally underestimates the peaks with a median value for PDR as  $-0.29$ , while SMAP hard update and SMOS hard update result in median PDR values of  $0.18$  and  $-0.23$ .

SMAP hard update reduces the median value for peak difference ratio, but it also increases the percentage of station-years with larger peak difference ratios. In contrast, SMAP EnKF reduces the median peak difference ratio to  $-0.08$ , and SMAP EnKFV experiments result in more improved predicted peaks with a median value of  $-0.04$  for peak difference ratio (i.e., slight underestimation).

SMOS EnKFV experiment reduces the percentage of station-years with peak overestimations compared to SMOS hard update and SMOS EnKF while it increases the percentage of station-years that underestimate the peaks. SMAP soil moisture data assimilation using EnKFV results in approximately 20% more improved peak streamflow predictions than the SMOS EnKFV approach.

## 5. Discussion

Our analysis on SMAP and SMOS satellite-based soil moisture error variance suggests a good agreement with findings from [44] but results in slightly higher values ( $\sim 0.02 \text{ m}^3/\text{m}^3$ ) than other studies with an average unbiased root mean square error (ubRMSE) ranges from  $0.04$  to  $0.062 \text{ (m}^3/\text{m}^3)$  [45–48]. We also found similar results on error seasonality compared to previous studies that investigated seasonal or time-variant errors in satellite-based soil moisture errors (e.g., [49,50]).

The variability of streamflow ensembles in the EnKF method is higher after larger rainfall events. The effect of soil moisture perturbations on predicted streamflow variability reduces after the peak of runoff event, as found by Niroula et al. [51].

Our results indicate that streamflow predictions improve after satellite-based soil moisture data assimilation using SMAP and SMOS satellite-based soil moisture observations. In terms of the KGE performance metric, SMOS soil moisture assimilation shows better streamflow prediction performance than SMAP during a dry year. Further comparison of the satellite-based soil moisture retrievals with field sensor-average shown in Figures A1 and A2 indicates that SMOS soil moisture is generally drier than SMAP satellite soil moisture.

Peak flow predictions after data assimilation of satellite-based soil moisture improve significantly. In this respect, the least and most improvements correspond to hard update and EnKFV, respectively. Compared to the open-loop mode, the latter case improves overall streamflow peak predictions up to 24%.

The accuracy and spatial resolution of satellite-based soil moisture are other limiting factors in improving streamflow predictions. Assimilating higher spatial resolution SMAP products, such as 3-km SMAP-Sentinel [52] or 1-km resolution [53] products, could provide higher prediction performance. For example, Abbaszadeh et al. [18] showed that assimilating SMAP soil moisture with 1-km resolution [53] results in better streamflow predictions compared to 9-km or 36-km satellite-based soil moisture. Our results also show that SMAP soil moisture ( $\sim 9 \text{ km}$ ) data assimilation provides better results than SMOS ( $\sim 43 \text{ km}$ ). The soil moisture retrievals in agricultural regions such as Iowa are prone to more considerable uncertainties than other regions [7]. We tried to address this issue in our study by including time-dependent variance in satellite-based soil moisture perturbations.

Our results indicate an overall improvement in streamflow predictions after satellite-based soil moisture assimilation. In contrast, previous studies found slight improvements (e.g., [9,18]) and sometimes degradation [13] in streamflow prediction performance after satellite-based soil moisture data assimilation. In addition to soil moisture data assimilation,

some studies have conducted dual state-parameter update or forcing update to further improve their streamflow predictions (e.g., [10,11,54]). A few studies have shown Particle Filtering could provide better performance in improving soil moisture estimates (e.g., [55]) and streamflow predictions (e.g., [56]) compared to EnKF. The differences in the degree of improvements by satellite-based soil moisture data assimilation in streamflow predictions could be related to the study region, data, and methodological aspects of different studies. This study only assimilated satellite-based soil moisture data (SMAP, SMOS) in a state-wide distributed hydrologic model to isolate their impact on streamflow predictions.

## 6. Summary and Conclusions

The goal of this study was to gain insights on the impact of SMAP and SMOS satellite-based soil moisture data assimilation on streamflow predictions. We investigated the utility of the satellite-based soil moisture data assimilation in streamflow predictions by three different assimilation experiments. First, we updated the model soil moisture by hard update (simple replacement) that does not account for potential observation errors. Then, we used Ensemble Kalman Filter (EnKF) to account for the soil moisture observation and modeling errors. For this purpose, we estimated the potential satellite-based soil moisture error variances using field sensor-average soil moisture. Moreover, we tested the impact of time-dependent error variance in the EnKF assimilation scheme for SMAP and SMOS satellite-based soil moisture products. We adopted a calibration-free distributed hydrologic model to conduct our numerical experiments. Finally, we evaluated the streamflow prediction performance with two evaluation metrics for each data assimilation experiment over the state of Iowa for four years and 131 USGS gauges. Based on our findings, the following conclusions could be drawn:

- SMAP and SMOS satellite-based soil moisture data assimilation improve streamflow predictions in terms of Kling-Gupta Efficiency (KGE) and Peak Difference Ratio.
- Hard update routine provided the least improvements to streamflow predictions while accounting for the potential satellite-based soil moisture errors using the EnKF approach, resulting in greater improvements in streamflow predictions.
- Time-dependent perturbations of satellite-based soil moisture observations resulted in better streamflow predictions than a constant error variance in the EnKF assimilation approach. These improvements are pronounced in terms of the streamflow peak predictions.
- SMAP satellite-based soil moisture data assimilation shows overall better performance during the study period (2015–2018) than SMOS satellite data. However, because SMOS satellite-based soil moisture is generally drier than SMAP, SMOS soil moisture data assimilation showed better performance during a dry year (2017).

Our study shows the potential of satellite-based soil moisture data assimilation in improving streamflow predictions with a calibration-free approach. Implications of our findings could be useful for data-scarce regions and ungauged basins where dual state-parameter or rainfall forcing updates are not feasible.

The operational flood forecasting model implemented at the IFC leverages more than 250 streamflow gauge data, including the USGS gauges and IFC bridge sensor observations for streamflow data assimilation. Our results could be useful for improving real-time flood forecasts by data assimilation of streamflow observations and satellite-based soil moisture observations.

Satellite-based soil moisture data assimilation in hydrologic models aims to correct potential historical errors caused by rainfall and evapotranspiration forcings in soil moisture state. We note that top-layer soil moisture, as a bounded variable, has less control than rainfall on the short-term water cycle. Therefore, perturbations in rainfall forcing could have a more considerable impact on streamflow predictions. We plan to investigate the impact of rainfall corrections on our hydrologic model's predictions as a follow-up study.

**Author Contributions:** Conceptualization, N.J., R.M. and W.F.K.; methodology, N.J., R.M. and W.F.K.; software, N.J.; validation, N.J.; formal analysis, N.J.; investigation, N.J.; resources, R.M. and W.F.K.; data curation, N.J.; writing—original draft preparation, N.J.; writing—review and editing, N.J., R.M. and W.F.K.; visualization, N.J.; supervision, R.M. and W.F.K.; project administration, W.F.K.; funding acquisition, R.M. and W.F.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by NASA Science Utilization of the Soil Moisture Active-Passive Mission program, grant number “NNX16AQ57G”.

**Institutional Review Board Statement:** Not applicable.

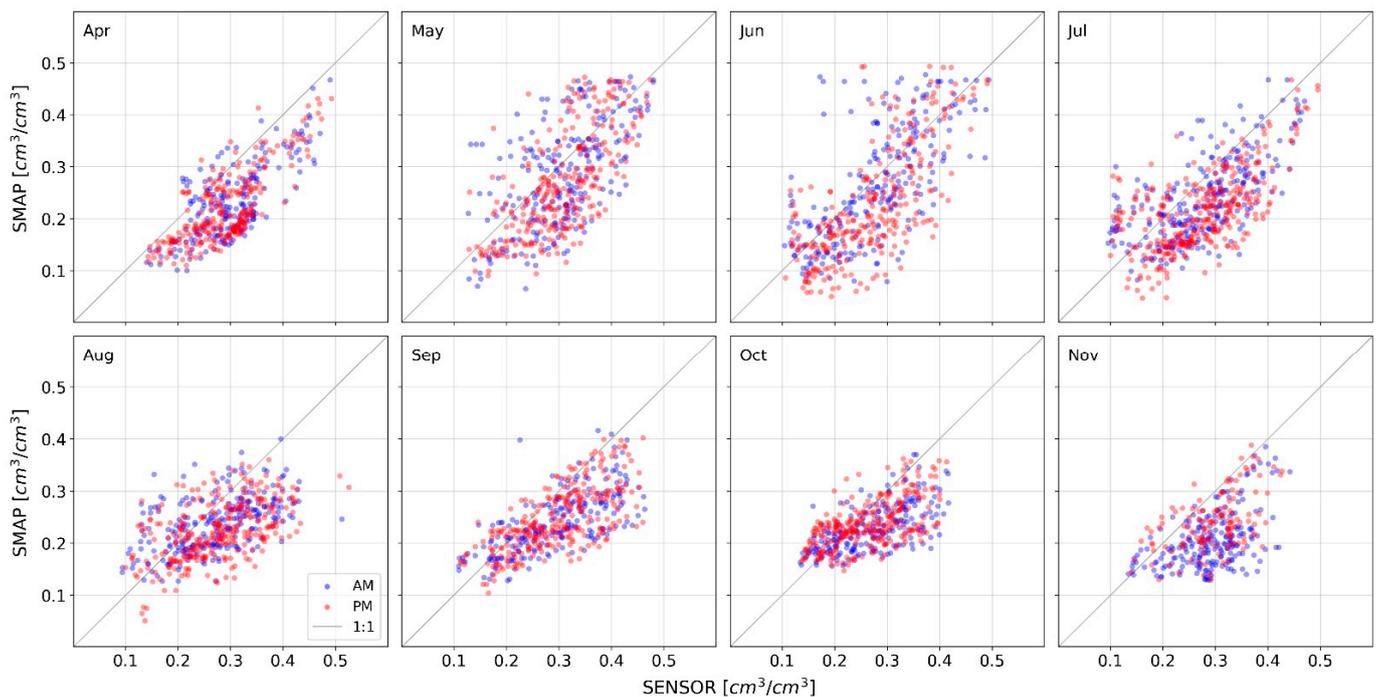
**Informed Consent Statement:** Not applicable.

**Acknowledgments:** We gratefully acknowledge the computational support from the High-Performance Computing group at the University of Iowa. Moreover, the authors acknowledge useful discussions with Brian Hornbuckle from Iowa State University.

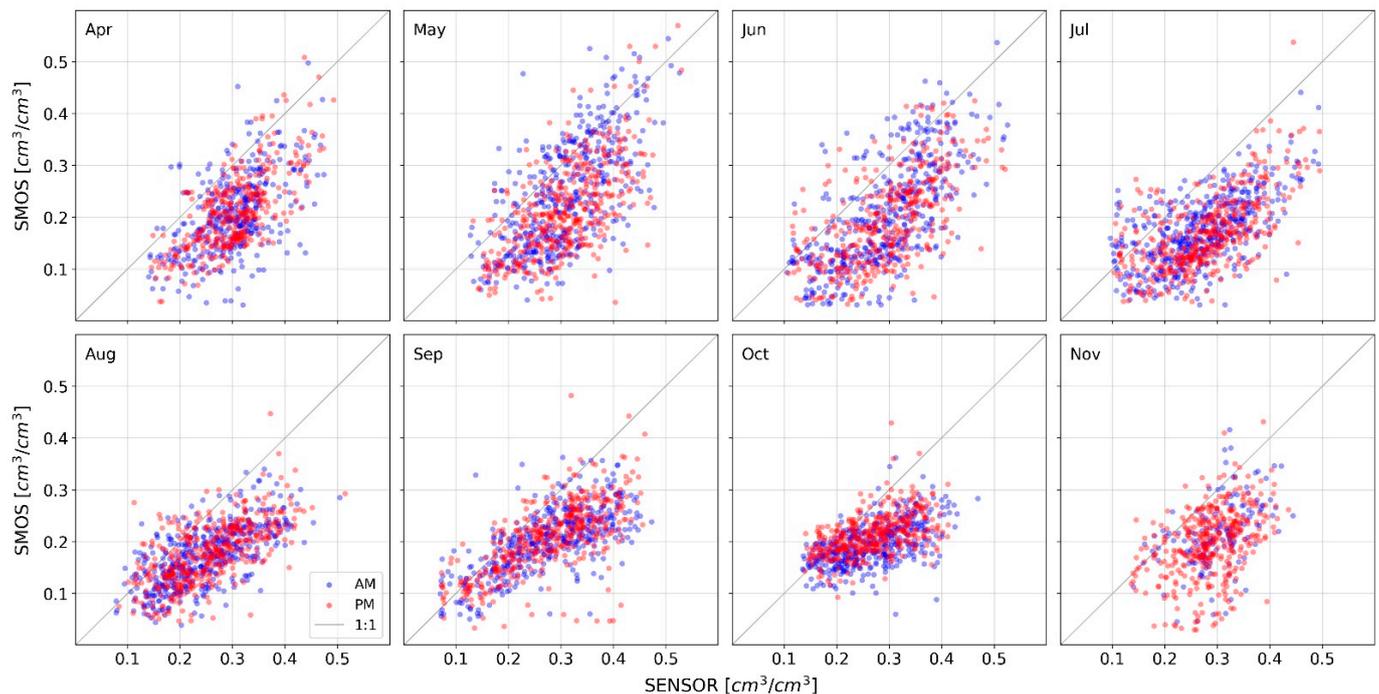
**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

In Figures A1 and A2, we compare the field sensor-average and SMAP and SMOS satellite-based soil moisture, respectively.



**Figure A1.** Scatterplots for soil moisture average of the collocated soil moisture field sensors with SMAP satellite-based soil moisture.



**Figure A2.** Scatterplots for soil moisture average of the collocated soil moisture field sensors with SMOS satellite-based soil moisture.

## References

1. Syed, T.H.; Famiglietti, J.S.; Rodell, M.; Chen, J.; Wilson, C.R. Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resour. Res.* **2008**, *44*. [[CrossRef](#)]
2. Liu, Y.; Gupta, H.V. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.* **2007**, *43*. [[CrossRef](#)]
3. Moradkhani, H.; Hsu, K.L.; Gupta, H.; Sorooshian, S. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resour. Res.* **2005**, *41*, 1–17. [[CrossRef](#)]
4. Crow, W.T.; Chen, F.; Reichle, R.H.; Liu, Q. L band microwave remote sensing and land data assimilation improve the representation of prestorm soil moisture conditions for hydrologic forecasting. *Geophys. Res. Lett.* **2017**, *44*, 5495–5503. [[CrossRef](#)] [[PubMed](#)]
5. Jadidoleslam, N.; Mantilla, R.; Krajewski, W.F.; Goska, R. Investigating the role of antecedent SMAP satellite soil moisture, radar rainfall and MODIS vegetation on runoff production in an agricultural region. *J. Hydrol.* **2019**, *579*, 124210. [[CrossRef](#)]
6. Hornbuckle, B.K.; England, A.W. Radiometric sensitivity to soil moisture at 1.4 GHz through a corn crop at maximum biomass. *Water Resour. Res.* **2004**, *40*. [[CrossRef](#)]
7. Neelam, M.; Mohanty, B.P. Global sensitivity analysis of the radiative transfer model. *Water Resour. Res.* **2015**, *51*, 2428–2443. [[CrossRef](#)]
8. Li, J.; Islam, S. Estimation of root zone soil moisture and surface fluxes partitioning using near surface soil moisture measurements. *J. Hydrol.* **2002**, *259*, 1–14. [[CrossRef](#)]
9. Chen, F.; Crow, W.T.; Ryu, D. Dual forcing and state correction via soil moisture assimilation for improved rainfall-runoff modeling. *J. Hydrometeorol.* **2014**, *15*, 1832–1848. [[CrossRef](#)]
10. Alvarez-Garretón, C.; Ryu, D.; Western, A.W.; Su, C.-H.; Crow, W.T.; Robertson, D.E.; Leahy, C. Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: Comparison between lumped and semi-distributed schemes. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 1659–1676. [[CrossRef](#)]
11. Alvarez-Garretón, C.; Ryu, D.; Western, A.W.; Crow, W.T.; Su, C.-H.; Robertson, D.R. Dual assimilation of satellite soil moisture to improve streamflow prediction in data-scarce catchments. *Water Resour. Res.* **2016**, *52*, 5357–5375. [[CrossRef](#)]
12. Leroux, D.J.; Pellarin, T.; Vischel, T.; Cohard, J.M.; Gascon, T.; Gibon, F.; Mialon, A.; Galle, S.; Peugeot, C.; Seguis, L. Assimilation of SMOS soil moisture into a distributed hydrological model and impacts on the water cycle variables over the Ouémé catchment in Benin. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 2827–2840. [[CrossRef](#)]
13. Mao, Y.; Crow, W.T.; Nijssen, B. A framework for diagnosing factors degrading the streamflow performance of a soil moisture data assimilation system. *J. Hydrometeorol.* **2019**, *20*, 79–97. [[CrossRef](#)]
14. Khaki, M.; Hendricks Franssen, H.J.; Han, S.C. Multi-mission satellite remote sensing data for improving land hydrological models via data assimilation. *Sci. Rep.* **2020**, *10*, 1–23. [[CrossRef](#)] [[PubMed](#)]

15. Dumedah, G.; Coulibaly, P. Evolutionary assimilation of streamflow in distributed hydrologic modeling using in-situ soil moisture data. *Adv. Water Resour.* **2013**, *53*, 231–241. [CrossRef]
16. Abbaszadeh, P.; Moradkhani, H.; Yan, H. Enhancing hydrologic data assimilation by evolutionary Particle Filter and Markov Chain Monte Carlo. *Adv. Water Resour.* **2018**, *111*, 192–204. [CrossRef]
17. Abbaszadeh, P.; Moradkhani, H.; Daescu, D.N. The quest for model uncertainty quantification: A hybrid ensemble and variational data assimilation framework. *Water Resour. Res.* **2019**, *55*, 2407–2431. [CrossRef]
18. Abbaszadeh, P.; Gavahi, K.; Moradkhani, H. Multivariate remotely sensed and in-situ data assimilation for enhancing community WRF-Hydro model forecasting. *Adv. Water Resour.* **2020**, *145*, 103721. [CrossRef]
19. Gochis, D.J.; Yu, W.; Yates, D. The WRF-Hydro Model Technical Description and User's Guide, Version 3.0. 2015. Available online: [https://ral.ucar.edu/sites/default/files/public/images/project/WRF\\_Hydro\\_User\\_Guide\\_v3.0.pdf](https://ral.ucar.edu/sites/default/files/public/images/project/WRF_Hydro_User_Guide_v3.0.pdf) (accessed on 19 March 2021).
20. Reichle, R.H. Data assimilation methods in the Earth sciences. *Adv. Water Resour.* **2008**, *31*, 1411–1418. [CrossRef]
21. Moradkhani, H.; Nearing, G.; Abbaszadeh, P.; Pathiraja, S. Fundamentals of data assimilation and theoretical advances. In *Handbook of Hydrometeorological Ensemble Forecasting*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 1–26.
22. Clark, M.P.; Vrugt, J.A. Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters. *Geophys. Res. Lett.* **2006**, *33*, 6406. [CrossRef]
23. Quintero, F.; Krajewski, W.F.; Seo, B.C.; Mantilla, R. Improvement and evaluation of the Iowa Flood Center Hillslope Link Model (HLM) by calibration-free approach. *J. Hydrol.* **2020**, *584*, 124686. [CrossRef]
24. Krajewski, W.F.; Ceynar, D.; Demir, I.; Goska, R.; Kruger, A.; Langel, C.; Mantilla, R.; Niemeier, J.; Quintero, F.; Seo, B.-C.; et al. Real-time flood forecasting and information system for the state of iowa. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 539–554. [CrossRef]
25. Summary Report: 2017 National Resources Inventory. Available online: <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/technical/nra/nri/results/> (accessed on 19 March 2021).
26. Eller, D. Iowa Fourth Worst for Flooding: \$18 Billion in Damages Since 1988. Available online: <https://www.desmoinesregister.com/story/money/agriculture/2018/04/29/iowa-flood-center-ranks-disaster-damages-billions-wapsipinicon-river/422336002/> (accessed on 5 February 2021).
27. Colliander, A.; Cosh, M.H.; Misra, S.; Jackson, T.J.; Crow, W.T.; Powers, J.; McNairn, H.; Bullock, P.; Berg, A.; Magagi, R.; et al. Comparison of high-resolution airborne soil moisture retrievals to SMAP soil moisture during the SMAP validation experiment 2016 (SMAPVEX16). *Remote Sens. Environ.* **2019**, *227*, 137–150. [CrossRef]
28. O'Neill, P.E.; Chan, S.; Njoku, E.; Jackson, T.; Bindlish, R. Soil Moisture Active Passive (SMAP) Algorithm Theoretical Basis Document Level 2 & 3 Soil Moisture (Passive) Data Products. 2016. Available online: [https://smap.jpl.nasa.gov/system/internal\\_resources/details/original/316\\_L2\\_SM\\_P\\_ATBD\\_v7\\_Sep2015.pdf](https://smap.jpl.nasa.gov/system/internal_resources/details/original/316_L2_SM_P_ATBD_v7_Sep2015.pdf) (accessed on 19 March 2021).
29. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.-P.; Delwart, S.; Cabot, F.; Boutin, J.; Escorihuela, M.-J.; Font, J.; Reul, N.; Gruhier, C.; et al. The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proc. IEEE* **2010**, *98*, 666–687. [CrossRef]
30. Brodzik, M.J.; Billingsley, B.; Haran, T.; Raup, B.; Savoie, M.H. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. *ISPRS Int. J. Geo Inf.* **2012**, *1*, 32–45. [CrossRef]
31. Talone, M.; Portabella, M.; Martínez, J.; González-Gambau, V. About the optimal grid for SMOS level 1C and level 2 products. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1630–1634. [CrossRef]
32. Lin, Y.; Mitchell, K.E. The NCEP Stage II/IV Hourly Precipitation Analyses: Development and Applications. 2005. Available online: <https://ams.confex.com/ams/pdfpapers/83847.pdf> (accessed on 19 March 2021).
33. Reed, S.M.; Maidment, D.R. Coordinate transformations for using NEXRAD Data in GIS-based hydrologic modeling. *J. Hydrol. Eng.* **1999**, *4*, 174–182. [CrossRef]
34. Mitchell, K.E.; Lohmann, D.; Houser, P.R.; Wood, E.F.; Schaake, J.C.; Robock, A.; Cosgrove, B.A.; Sheffield, J.; Duan, Q.; Luo, L.; et al. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res. Atmos.* **2004**, *109*, D07S90. [CrossRef]
35. U.S. Geological Survey National Water Information System. Available online: <https://waterdata.usgs.gov/nwis> (accessed on 19 March 2021).
36. Mantilla, R.; Gupta, V.K. A GIS numerical framework to study the process basis of scaling statistics in river networks. *IEEE Geosci. Remote Sens. Lett.* **2005**, *2*, 404–408. [CrossRef]
37. Small, S.J.; Jay, L.O.; Mantilla, R.; Curtu, R.; Cunha, L.K.; Fonley, M.; Krajewski, W.F. An asynchronous solver for systems of ODEs linked by a directed tree structure. *Adv. Water Resour.* **2013**, *53*, 23–32. [CrossRef]
38. Quintero, F.; Krajewski, W.F.; Mantilla, R.; Small, S.; Seo, B.C. A spatial-dynamical framework for evaluation of satellite rainfall products for flood prediction. *J. Hydrometeorol.* **2016**, *17*, 2137–2154. [CrossRef]
39. Evensen, G. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dyn.* **2003**, *53*, 343–367. [CrossRef]
40. Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **1994**, *99*, 10143. [CrossRef]
41. Burgers, G.; Jan van Leeuwen, P.; Evensen, G. Analysis Scheme in the Ensemble Kalman Filter. *Mon. Weather Rev.* **1998**, *126*, 1719–1724. [CrossRef]

42. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91. [[CrossRef](#)]
43. Jadidoleslam, N.; Goska, R.; Mantilla, R.; Krajewski, W.F. Hydrovis: A non-proprietary open-source software for hydrologic model and data visualization and evaluation. *Environ. Model. Softw.* **2020**, *134*, 104853. [[CrossRef](#)]
44. Stillman, S.; Zeng, X. Evaluation of SMAP soil moisture relative to five other satellite products using the climate reference network measurements over USA. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6296–6305. [[CrossRef](#)]
45. Chan, S.K.; Bindlish, R.; O'Neill, P.; Jackson, T.; Njoku, E.; Dunbar, S.; Chaubell, J.; Piepmeier, J.; Yueh, S.; Entekhabi, D.; et al. Development and assessment of the SMAP enhanced passive soil moisture product. *Remote Sens. Environ.* **2018**, *204*, 931–941. [[CrossRef](#)] [[PubMed](#)]
46. Colliander, A.; Jackson, T.J.; Chan, S.K.; O'Neill, P.; Bindlish, R.; Cosh, M.H.; Caldwell, T.; Walker, J.P.; Berg, A.; McNairn, H.; et al. An assessment of the differences between spatial resolution and grid size for the SMAP enhanced soil moisture product over homogeneous sites. *Remote Sens. Environ.* **2018**, *207*, 65–70. [[CrossRef](#)]
47. Zhang, R.; Kim, S.; Sharma, A. A comprehensive validation of the SMAP Enhanced Level-3 Soil Moisture product using ground measurements over varied climates and landscapes. *Remote Sens. Environ.* **2019**, *223*, 82–94. [[CrossRef](#)]
48. Walker, V.A.; Hornbuckle, B.K.; Cosh, M.H. A Five-Year Evaluation of SMOS Level 2 Soil Moisture in the Corn Belt of the United States. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 4664–4675. [[CrossRef](#)]
49. Walker, V.A.; Hornbuckle, B.K.; Cosh, M.H.; Prueger, J.H. Seasonal evaluation of SMAP soil moisture in the U.S. Corn Belt. *Remote Sens.* **2019**, *11*, 2488. [[CrossRef](#)]
50. Wu, K.; Ryu, D.; Nie, L.; Shu, H. Time-variant error characterization of SMAP and ASCAT soil moisture using triple collocation analysis. *Remote Sens. Environ.* **2021**, *256*, 112324. [[CrossRef](#)]
51. Niroula, S.; Halder, S.; Ghosh, S. Perturbations in the initial soil moisture conditions: Impacts on hydrologic simulation in a large river basin. *J. Hydrol.* **2018**, *561*, 509–522. [[CrossRef](#)]
52. Das, N.N.; Entekhabi, D.; Kim, S.; Yueh, S.; O'Neill, P. Combining SMAP and sentinel data for high-resolution soil moisture product. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; Volume 2016, pp. 129–131.
53. Abbaszadeh, P.; Moradkhani, H.; Zhan, X. Downscaling SMAP Radiometer soil moisture over the CONUS using an ensemble learning method. *Water Resour. Res.* **2019**, *55*, 324–344. [[CrossRef](#)]
54. Moradkhani, H.; Sorooshian, S.; Gupta, H.V.; Houser, P.R. Dual state-parameter estimation of hydrological models using ensemble Kalman filter. *Adv. Water Resour.* **2005**, *28*, 135–147. [[CrossRef](#)]
55. Lei, F.; Crow, W.T.; Kustas, W.P.; Dong, J.; Yang, Y.; Knipper, K.R.; Anderson, M.C.; Gao, F.; Notarnicola, C.; Greifeneder, F.; et al. Data assimilation of high-resolution thermal and radar remote sensing retrievals for soil moisture monitoring in a drip-irrigated vineyard. *Remote Sens. Environ.* **2020**, *239*, 111622. [[CrossRef](#)] [[PubMed](#)]
56. Dechant, C.; Moradkhani, H. Radiance data assimilation for operational snow and streamflow forecasting. *Adv. Water Resour.* **2011**, *34*, 351–364. [[CrossRef](#)]