A comprehensive assessment of in situ and remote sensing soil moisture data assimilation in the APSIM model for improving agricultural forecasting across the U.S. Midwest

Marissa Kivi[1], Noemi Vergopolan[2], Hamze Dokoohaki[1*]

1 Crop science department, University of Illinois at Urbana-Champaign, Urbana, IL, USA
2 Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA

*Corrpospoding author: Hamze Dokoohaki; hamzed@illinois.edu

Abstract. Today, the most popular approaches in agricultural forecasting leverage process-based crop models, crop monitoring data, and/or remote sensing imagery. Individually, each of these tools has its own unique advantages but is, nonetheless, limited in prediction accuracy, precision, or both. In this study we integrate in situ and remote sensing (RS) soil moisture observations with APSIM model through sequential data assimilation to evaluate the improvement in model predictions of downstream state variables across 5 experimental sites in the U.S Midwest. Four RS data products and in-situ observations spanning 19 site-years were used through two data assimilation approaches namely Ensemble Kalman Filter (EnKF) and Generalized Ensemble Filter (GEF) to constrain model states at observed time steps and estimate joint background and observation error matrices. Then, the assimilation's impact on estimates of soil moisture, yield, NDVI, tile drainage, and nitrate leaching was assessed across all site-years. When assimilating in situ observations, the accuracy of soil moisture forecasts in the assimilation layers was improved by reducing RMSE by an average of 17% for 10cm and ~28% for 20 cm depth soil layer across all site-years. These changes also led to improved simulation of soil moisture in deeper soil layers by an average of 12%. Although crop yield was improved by an average of 23%, the greatest improvement in yield accuracy was demonstrated in site-years with higher water stress, where assimilation served to increase available soil water for crop uptake. Alternatively, estimates of annual tile drainage and nitrate leaching were not well constrained across the study sites. Trends in drainage constraint suggest the importance of including additional data constraint such as evapotranspiration. The assimilation of RS soil moisture showed weaker constraint of downstream model state variables when compared to the assimilation of in situ soil moisture. The median reduction in soil moisture RMSE for observed soil layers was lower, on average, by a factor of 5. However, crop yield estimates were still improved overall with a median RMSE reduction of 17.2%. Crop yield prediction was improved when assimilating both in-situ and remote sensing soil moisture observations and there is strong evidence that yield improvement was higher when under water-stressed conditions. Comparisons of system performance across different combinations of remote sensing data products indicated the importance of high temporal resolution and accurate observation uncertainty estimates when assimilating surface soil moisture observations.

Keywords: Model-data integration, Sequential Data Assimilation, APSIM, soil moisture

1. Introduction

To effectively address pressing global food security challenges, agricultural forecasting tools must exhibit high accuracy and precision across spatial and temporal scales. As process-based crop models offer a system-level representation of many soil and crop processes, they are increasingly recognized as practical forecasting tools in agricultural research (Silva and Giller, 2021; Fer at al., 2021). However, their weakness comes from many unaccounted uncertainties, such as those related to model parameters, initial conditions, and weather (Dokoohaki et al., 2021). Prior studies have shown state data assimilation (SDA) to be a powerful tool to overcome this weakness in process-based crop models (e.g. Dokoohaki et al., 2022a). SDA enables a temporally-continuous, high-dimensional scaffold in which a variety of observations can be smoothly integrated using one of many robust, systematic algorithms, such as the Ensemble Kalman Filter (EnKF; Dietze et al., 2017; Huang et al., 2019; Liu et al., 2021; Dokoohaki et al., 2022a; Kivi et al., 2022). Through SDA, uncertainty around spatially-heterogenous and dynamic properties in agricultural systems can be constrained, thereby increasing precision and accuracy in estimates while decreasing dependence on extensive site-level model calibration (Mishra et al., 2021).

Numerous past studies have used SDA to constrain crop model estimates, using observations on leaf area index (e.g., Nearing et al., 2012; Ines et al., 2013; Ma et al., 2013; Chen et al., 2018; Lu et al., 2021), soil moisture (Kivi et al., 2022), biomass (e.g., Linker and Ioslovich, 2017) and evapotranspiration (e.g., Huang et al., 2015). For example, a synthetic study by Zhu et al. (2017) found that the assimilation of coarse resolution surface soil moisture data into a coupled soil water-groundwater numerical model constrained soil moisture estimates in the first 50 cm of the soil profile despite explicitly unaccounted spatial heterogeneity in soil properties. These studies showed how SDA can partially account for the spatial variability in soil hydraulic conductivity across broad regions without explicit model calibration. In addition to incorporating spatial heterogeneity in soil properties, Kivi et al. (2022) demonstrated that the assimilation of high quality and frequent in-situ soil moisture observations can substantially improve downstream model predictions of tile drainage, nitrate (NO_3) leaching, and root-zone soil moisture (RZSM) for maize and soybeans in the APSIM model. However, collecting field measurements of soil moisture for different cropping systems, soils, and environments is expensive, extremely laborious, and time-consuming.

Alternatively, the assimilation of high-resolution Remote Sensing (RS) data products dramatically increases SDA applications' range beyond in situ data availability by effectively capturing the spatiotemporal variability of many agricultural state variables, such as vegetation cover and soil moisture, with consistency and high temporal frequency (Peng et al., 2017). As a result, RS observations could be invaluable to constraining model predictions at the regional scale and have been increasingly applied for agricultural forecasting in the data assimilation literature, as demonstrated in literature reviews by Dorigo et al. (2007), Huang et al. (2019), and Weiss et al. (2020). The application of RS soil moisture data products has been especially popular and successful in data assimilation-focused agricultural forecasting studies. These data products, which characterize soil moisture content in the first 5 cm of the soil profile, pull information from active and/or passive sensors of microwave reflectance. Due its high sensitivity to surface soil moisture, many data products have been developed around available L-band microwave sensor information collected by NASA's SMAP Mission (Kumar et al., 2018). The SMAP-HydroBlocks data products merges SMAP data with the HydroBlocks land surface model to increase spatial resolution in the final estimates and improve scalability

69 (Vergopolan et al., 2021b), while the SMAP-Sentinel1 data product pairs SMAP data with Sentinel-1 radar
70 information to achieve similar goals (Das et al., 2019). Others, like the ESA-CCI data product (Dorigo et al., 2017),
71 compile information from multiple sensors, including the SMAP passive sensor, to allow for greater temporal
72 coverage. However, this approach comes at the cost of coarser spatial resolution.

73 Nonetheless, as demonstrated in past studies, the assimilation of RS soil moisture data has its limitations. First,
74 uncertainty and biases in RS data products are typically poorly defined (Huang et al., 2019). RS-based data products
75 are based on empirical relationships, and, as they are predicted as a function of surface reflectance, uncertainties in
76 the raw radiance will propagate unsupervised into final estimates (Weiss et al., 2020). Additionally, RS estimates
77 characterize soil moisture in only the top 5 cm of the soil profile and, thus, rely on models or empirical
78 parameterizations to describe the root zone soil profile. Among others, De Lannoy et al. (2007) and Monsivais-
79 Huertero et al. (2010) both found the assimilation of in-situ near-surface soil moisture observations to be far less
80 effective than that of in-situ root-zone soil moisture observations in constraining estimates of the greater soil water
81 profile. Yet, since the surface layer is typically the layer where fertilizers are added, the accurate estimation of surface
82 layer state variables is essential for today's agroecosystems (Verburg and CSIRO, 1996). To overcome relatively
83 coarse spatial resolution in RS data products, past studies have explored downscaling approaches (e.g.,
84 ChakrabartiChakrabarti et al., 2014) or leveraged additional in-situ datasets (e.g., Liu et al., 2021) to overcome
85 "mismatch" challenges and downscale RS soil moisture estimates to more accurately reflect field scale measurements
86 (Vergopolan et al., 2021a). However, the reliance on in situ observations of these approaches can limit system
87 transferability across broad regions (Peng et al., 2017). Moreover, as described by Crow et al. (2012), it can be difficult
88 to properly evaluate coarse soil moisture estimates with point-scale ground measurements due to unknown and often
89 significant sampling uncertainty. Data assimilation with process-based models has been previously applied as a robust
90 and scalable way to leverage information in coarse resolution soil moisture estimates (e.g. Vergopolan et al., 2021b).
91 Despite the immense theoretical potential of SDA with both in situ and RS observations, past studies have reported
92 inconsistent SDA performance in modeling crop yields. For example, de Wit and van Diepen (2007) observed
93 inconsistencies in yield constraint when assimilating soil wetness index (SWI) derived from 0.25° ERS1/2 microwave
94 radiance information into the WOFOST model across agricultural regions of Spain, Germany, France, and Italy. They
95 partially attributed poor predictions in certain regions to irrigation processes that were not captured by the model nor
96 coarse resolution SWI observations. Lu et al. (2021) also saw year-to-year variability in assimilation performance
97 when assimilating in situ observations of canopy cover and soil moisture for 6 site-years in Nebraska. When
98 assimilating soil moisture independently, canopy cover estimates were better constrained in drier years. They
99 suspected this to result from the canopy cover's lower sensitivity to soil moisture in the model when water is in surplus
100 (i.e., due to energy-limited conditions). We further suspect that SDA's inconsistent performance is related to the
101 misrepresentation of model processes linking soil moisture to crop- and soil-related variables (e.g., soil nitrogen, leaf
102 expansion, crop water uptake). As a result, direct upstream improvement of model state variables with SDA does not
103 always translate into improvement in downstream results. To understand the role of soil moisture data assimilation in
104 improving crop yields and better pinpoint areas for future improvement, a comprehensive assessment that investigates
105 performance across time and different genetic (G), environmental (E), and management (M) spaces is required.

106 Although a growing body of studies has attempted to quantify the impact of soil moisture assimilation in crop models,
107 such a comprehensive evaluation of in situ and RS soil moisture SDA in crop models across GxExM spaces is still
108 lacking (Folberth et al. 2016b).

109 To bridge this knowledge gap, we present a comprehensive assessment of soil moisture data assimilation as a method
110 for constraining crop model predictions across the U.S. Midwest. Building on the assimilation framework in Kivi et
111 al. (2022), we independently assimilated both in situ and RS soil moisture observations in the APSIM crop model at
112 five experimental sites in the U.S Midwest. With field data covering 19 site-years of corn and soybean cropping
113 systems across the region, this study tests the data assimilation system across a broader GxExM inference space and
114 quantifies the benefit of assimilating different RS soil moisture products in comparison to the in-situ soil moisture
115 observations. The main objectives of this study were:

116   1.   To quantify how in situ soil moisture observations can constrain crop model forecasts of downstream estimates,
117   including root-zone soil moisture, crop yield, crop phenology via NDVI, tile drainage flow, and NO3 leaching
118   through SDA.
119   2.   To quantify the added benefit of RS soil moisture observations in improving crop model predictions of root-
120   zone soil moisture, crop yield, and crop phenology via NDVI through SDA.

2. Methods

122 Sections 2.1 and 2.2 describe the five experimental sites and the in-situ observations employed in this study for model
123 set-up, SDA, and evaluation. Section 2.3 outlines the four different RS soil moisture data products that were
124 assimilated, and Section 2.4 presents the data-assimilation system used in this study. Lastly, Section 2.4.5 defines the
125 different simulation experiments performed.

2.1 Study sites

127 This study focused on five experimental sites across the U.S. Midwest with in-situ observations of soil moisture, crop
128 yield, nitrate load, and tile drainage flow for 19 years between 2011 and 2019. Site IL was the Energy Farm, a well-
129 monitored experimental site in central Illinois that was the focus of the development and initial evaluation of the
130 employed data-assimilation system (Kivi et al., 2022). Site IN, MN, OH, and SD were available through the
131 Transforming Drainage (TD) project (Chighladze et al., 2021). The TD project database is publicly-available and
132 contains high-quality data from 39 tile-drained research sites with data spanning over 200+ site-years. The available
133 observations include data on tile drainage, yield, water table, water quality, and soil characteristics, among many
134 others. Though numerous sites were available as part of the project, the experimental design and data available for
135 each site-year varies widely in the database. For consistency, this work required that each site-year include a plot with:
136 (1) a free tile drainage system, (2) available NO3 load and tile flow data at the plot level, (3) available in situ soil
137 moisture observations, (4) maize or soybean crops, and (5) a rain-fed system. We identified only 17 site-years across
138 five sites in the database which satisfied all these criteria.

139 To properly set up the APSIM model for each of the five sites, we included all available site information on each year,
140 cropping system, residue type, planting and harvesting details, tillage practices, and fertilizer applications as constants

in the simulations. Following updated information available through Moore et al. (2021), the IL siteincludes tillage practices in the model set-up and increased nitrogen (N) fertilizer from 64.6 kg N/ha, to 202 kg N/ha. Detailed information on the plot and management information for all five sites are included in the Supplementary Materials (Table A1). Study sites will be referred to by their given study IDs in Figure 1.

2.2 Observation data

In situ soil moisture

Across the study site-years, sub-daily soil moisture (SM) observations were collected at various soil depths between 10 and 105 cm using soil sensors; the measured depths and sensor type varied by site. All observations are available in units of volumetric water fraction (VFW; mm/mm). For the 4 TD sites, SM observations were only available as daily averages. For consistency, SM observations at IL (available at 15-minute intervals) were aggregated to daily averages when at least 40 15-minute observations were available. Observations from the winter months (December-March) were excluded due to the influence of freezing soils. Across all site-years, in situ SM assimilation was performed with available observations for the 10- and 20-cm soil depths, which hereinafter will be referred to as SM3 and SM4, respectively. All other available SM observations for deeper soil layers were used to evaluate model root-zone SM estimates. SM observations were paired with an APSIM soil layer based on the recorded sensor depth and the site soil profile. In the case that more than one observation was available for a given APSIM soil layer, the average SM was computed for each day and layer with the assumption of uniform SM in the layer.

Harvested maize and soybean yields

Data on harvested yield for the TD sites were available for each site-year with 1-3 replicated measurements. These replicated observations were averaged and converted from grain at standard moisture content (i.e., 15.5% for maize and 13% for soybean) to dry-grain weight for best comparison with the APSIM model output. Observations for IL were already recorded as dry-grain weights and given in units of kg/ha. Across 12 maize site-years, observed yields ranged from 6.51 to 13 Mg/ha with an average yield of 9.93 Mg/ha. The 7 soybean site-years had observed yields ranging from 2.78 to 4.15 Mg/ha with an average yield of 3.50 Mg/ha.

Remotely sensed Normalized Difference Vegetation Index (NDVI)

The normalized difference vegetation index (NDVI) can be used to quantify vegetation greenness and reasonably track the phenological development of crops (Gao and Zhang, 2021). In this study, NDVI observations from Landsat between 2011 and 2019 were used to evaluate APSIM's performance in predicting crop phenology for each site-year. NDVI time series were extracted at each site location from Landsat 7 and 8 remote sensing imagery courtesy of the U.S. Geological Survey via Google Earth Engine and derived from the red (RED) and near-infrared (NIR) spectral bands using the following equation:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

175

In situ measurements of tile drainage and nitrate load

Daily observations of tile drainage flow (mm) and NO3 load (kg NO3-N ha-1) were available for all 19 site-years. Any missing daily drainage values for the TD sites had been imputed previously and used to approximate missing values of daily NO3 load, as described by Helmers et al. (2022). Methods and instrumentation used to collect and process the TD sites and IL data are presented by Helmers et al. (2022) and Kivi et al. (2022), respectively. In this study, daily values for tile drainage flow and NO3 load were summed to annual values for comparison with model output. For the purposes of this analysis, we assumed any day with NA tile drainage flow values in the data had no drainage and no NO3 loss.

2.3 Remote sensing soil moisture

To assess the performance of SM data assimilation with satellite-based observations, we included 4 RS data products that span different temporal and spatial resolutions (Table 1). These observations were extracted at the point level for the study sites and serve to represent the first 5 cm of the soil profile or surface SM. Observations from the winter months (i.e., December-March) were removed to avoid issues with snow cover and freezing soils. The product IDs provided in Table 1 will be used to identify each data product.

ESA-CCI

The RS dataset with the coarsest spatial resolution in this study was the ESA-CCI SM product. Each year, the European Space Agency Climate Change Initiative (ESA CCI) algorithmically merges information from 3 active (e.g., ASCAT A/B) and 10 passive (e.g., SSM/I, AMSR-E, SMOS, SMAP) microwave sensors to estimate daily surface SM globally for over 40 years. Dorigo et al. (2017) provide complete documentation on how these data products are produced. Here we used the combined product (version v06.1), which includes daily uncertainty estimates. Several past studies have assimilated this data product into process-based models with varying levels of success (e.g., Zhou et al., 2016; Liu et al., 2017; Liu et al., 2018; Naz et al. 2019).

SMAP-HydroBlocks

The SMAP-HydroBlocks surface SM dataset has the highest spatial resolution in this study. It was introduced by Vergopolan et al. (2021b) by combining the HydroBlocks land surface model, a Tau-Omega radiative transfer model, machine learning, in situ SM observations, and SMAP remotely sensed satellite observations to estimate surface SM with 30-meter resolution across the contiguous United States. In specific, the Hydroblocks model was coupled with a Tau-Omega radiative transfer model (HydroBlocks-RTM) and used to simulate SM, soil temperature, and brightness temperature at a 3-hour, 30-meter resolution. Brightness temperature estimates from NASA's Soil Moisture Active Passive (SMAP) mission were then merged with the HydroBlocks-RTM estimates using a spatial cluster-based Bayesian merging scheme (Vergopolan et al., 2020). Using the inverse HydroBlocks-RTM, SM was estimated at SMAP overpass time at 30-m spatial resolution. Vergopolan et al. (2021b) reported an RMSE of 0.07 mm3/mm3 after comparing SMAP-Hydroblocks estimates to in situ observations from 233 independent experimental sites. This study

is the first to assimilate SMAP-HydroBlocks SM estimates into a crop model. SM morning and afternoon retrievals were aggregated to a daily resolution, and site-level estimates were computed as the mean value of any data point within 0.0005° of the given site location. The uncertainty estimate for each observation was calculated based on the spatial variability of selected data points for that time step and the reported standard error (SE = 0.07 mm3/mm3) as :

$$Var(Y_{s,t}) = Var(y_t) + SE^2 \qquad (2)$$

where, for site s at the tth available time step, Y represents the site-level SM estimate, and y presents SM estimates within 0.0005° of the site location.

SMAP-Sentinel1

The SMAP-Sentinel1 SM product was produced by merging information collected by the SMAP L-band radiometer and the Copernicus Project Sentinel-1 C-band radar. After the malfunction of the SMAP radar in 2015, Sentinel-1 active microwave data were used with passive microwave sensor information from the still-operating SMAP radiometer to estimate surface SM content globally using the active-passive algorithm. Although the merged product increased the revisit interval from 3 to 12 days, it enabled retrievals at two different spatial resolutions (i.e., 1 km and 3 km; Lievens et al., 2017). Upon comparing the estimates with in situ SM measurements, Das et al. (2019) reported RMSE for SMAP-Sentinel1 SM estimates as roughly 0.05 m3/m3. In this study, this value was applied as the standard error for SM estimates at both spatial resolutions and at all available time steps. Retrievals were available for all TD site-years but were unavailable for IL for unknown reasons.

2.4 Data-assimilation system

This study uses the data-assimilation system developed and evaluated in Kivi et al. (2022). The original system leveraged the pSIMS platform, APSIM crop model, Ensemble Kalman Filter (EnKF), and an algorithm presented by Miyoshi et al. (2013) to estimate and propagate uncertainties, perform sequential data assimilation, and generate daily agricultural forecasts at the field scale. The workflow is illustrated in Figure 2. APSIM management variables that were known include planting and harvest dates, fertilizer amount, type, and timing, tillage type, depth, and timing, crop type, row spacing, sowing density, and, if available, planting depth.

2.4.1 Model parameter priors

Initial soil water, cultivar, and residue weight were randomized across model ensembles for each site to incorporate uncertainty around initial conditions. If unavailable in the management data, planting depth was also randomized and drawn from different prior distributions for each crop. These distributions represented reasonable planting depth ranges for the two crops in the Midwest, as described in extension websites produced by the University of Missouri (Luce, 2016) and Michigan State University (Staton, 2012). Using a uniform prior distribution, planting depths ranged from 1.5 to 2.5 inches for maize and 1 to 2 inches for soybean.

244   Prior distributions were also set to incorporate uncertainty around cultivar. For maize, nine cultivar parameters were

245   ensembled, including the six cultivar parameters (i.e., tt_flower_to_maturity, tt_flower_to _start_grain,

246   tt_maturity_to_ripe, tt_emerg_to_endjuv, head_grain_no_max, grain_gth_rate). The other three parameters (i.e.,

247   largestLeafParams1, leaf_init_rate, leaf_app_rate1) were drawn from Dokoohaki et al. (2022b), who identified maize

248   cultivar parameters that were influential for estimates of leaf area index (LAI) in the APSIM Maize module and

249   optimized their value distributions using a hierarchical Bayesian optimization approach across the U.S. Midwest.

250   Table A.2 gives more detailed information on all randomized parameters and their prior distributions. We completed

251   a preliminary assessment of the Maize module at each of the study sites and found that, under the given parameter

252   value ranges, APSIM was capable of appropriately simulating the phenological development and grain yield for maize

253   at each site.

254   The selection of soybean cultivars for each site was determined using a semi-systematic approach. First, a range of

255   maturity groups was determined for each site based on a study by Mourtzinis and Conley (2017), which delineated

256   soybean maturity groups across the U.S. We defined the upper and lower maturity group bounds for each site using

257   the bounding zone contour lines for each site location in Figure 4 of Mourtzinis and Conley (2017). Then, initial

258   APSIM simulations were performed for each site using all APSIM-defined soybean cultivars falling within the

259   prescribed maturity group range. The model results were compared to the observed soybean yields at each site, and

260   the best-performing maturity group (MG) for each site was determined. The final range for each site was

261   approximately MG $\pm$ 0.5. In each ensemble, the cultivar for each crop at each site was assumed to be constant across

262   all site-years.

263   **2.4.2 Weather and soil model drivers**

264   To incorporate uncertainty around soil and weather into our simulations, a Monte Carlo sampling approach was used

265   to randomly assign ensembles of weather and soil drivers to model ensembles. For each study site, ten weather

266   ensembles from the ERA5 reanalysis data product were employed to characterize solar radiation, maximum air

267   temperature, minimum air temperature, precipitation, and wind speed at the daily resolution and at each site location.

268   ERA5 is a global gridded reanalysis data product from the European Centre for Medium-Range Weather Forecasts

269   (ECMWF), which characterizes the weather state variables at hourly time steps with associated uncertainties

270   (Hersbach et al., 2020). In addition, 25 soil ensembles were generated from the SoilGrids global gridded soil database

271   (Hengl et al., 2014) for each site location. These ensembles cover 30 soil properties (including available water lower

272   limit, bulk density, drained upper limit, organic carbon, soil class, and pH) and were created by sampling from each

273   soil parameter mean and uncertainty values available in the SoilGrids dataset.

274   **2.4.3 PROSAIL model**

275   Since APSIM does not currently estimate NDVI, APSIM was coupled with the PROSAIL model described in

276   Dokoohaki et al. (2022b) to estimate daily NDVI values and enable the appropriate evaluation of the model's

277   simulation of crop phenology at the study sites. The PROSAIL model is a radiative transfer tool that combines

278   PROSPECT, a leaf optical properties model, and SAIL, a canopy bidirectional reflectance model, to estimate spectral

279   reflectance for a given vegetative area based on soil and plant/canopy properties (Jacquemoud et al., 2009). In this

280 study, APSIM's daily forecasts of soil and plant variables were transformed and used as inputs into the PROSAIL

281 model to compute the spectral reflectance for each ensemble. Then, for each day and ensemble, the estimated spectral

282 information was used to estimate NDVI using the vegetation index function within the hsdar R library (Lehnert et al.,

283 2019). Further details on the coupling protocols can be found in Dokoohaki et al., (2022b).

284

2.4.4 Ensemble Kalman filter with the Miyoshi algorithm

286 The data-assimilation system(which we will call EnKF-Miyoshi hereinafter) employs the ensemble Kalman filter

287 (EnKF) to assimilate SM observations into the APSIM model. The EnKF merges information from the model

288 ensemble forecast distribution and observations (with associated uncertainty) at each time step to optimally estimate

289 the state of the system (Evensen, 2003). The system also leverages the Miyoshi algorithm in series with the EnKF to

290 improve estimates of the two system uncertainty matrices (i.e., Pf and R) and improve filter performance. Based on

291 diagnostic innovation statistics, the Miyoshi algorithm estimates a forecast inflation scalar (Δ) and observation

292 uncertainty (R) at each analysis time step. At time step t with available data, the system follows the following steps:

1. The mean (Xf,t) and the variance-covariance matrix (Pf,t) of the model forecast ensemble are computed to define the forecast distribution, which is assumed to follow a Normal distribution.

2. The observed distribution (Yt) is also assumed to be Normal with mean yt and variance-covariance matrix Rt, where Rt = R* from the previous analysis time step or R1 = Σ. Σ is a diagonal matrix that assumes 10% standard error for each observed state variable.

3. The Kalman Gain (K) is computed as follows, where $\Delta t = \Delta^*$ or $\Delta 1 = I$ (I is the identity matrix) and H is the observation operator:

$$K_t = \Delta_t P_{f,t} H^T (R_t + H\Delta_t P_{f,t} H^T)^{-1} \tag{3}$$

4. The analysis distribution, which assumes a Normal distribution, is determined with mean (Xa,t) and variance-covariance matrix (Pa,t).

$$X_{a,t} = X_{f,t} + K_t(Y_t - HX_{f,t}) \tag{4}$$
$$P_{a,t} = (I - K_t H)P_{f,t}$$

5. The model ensemble is updated at each time step according to the analysis distribution based on each ensemble's likelihood within the forecast distribution.

6. Δ^* and R* are recomputed using the following series of equations, where d_{o-a} and d_{o-f} represent the observation-analysis and observation-forecast innovations for the current time step, respectively, E denotes the expectation operator, and ρ is a user-defined weight given to the new estimate. A lower bound of 1 is imposed on each entry in Δest and only the diagonal entries of Rest are maintained.

$$E(d_{o-a} d_{o-f}^T) = R_{est}$$
$$\Delta_{est} = \frac{d_{o-f}^T d_{o-f} - R_{est}}{H\Delta_t P_{f,t} H^T} \tag{5}$$

$$R^* = (\rho)R_{est} + (1-\rho)R_t$$
$$\Delta^* = (\rho)\Delta_{est} + (1-\rho)\Delta_t$$

310

2.4.5 Generalized ensemble filter

However, the EnKF-Miyoshi workflow as established cannot robustly handle observation operators (H) that change dimensions over time. However, to reduce information loss within the system, H must be able to adapt according to the number of observations available. To increase flexibility in system configuration, an alternative sequential data assimilation approach was tested in this work to replace the EnKF-Miyoshi method. The new method, hereinafter called the Generalized Ensemble Filter (GEF), comprises a fully numerical Bayesian approach to estimating the analysis distribution and an inflation scalar. The model resembles the approach presented by Raiho et al. (2020) and Dokoohaki et al., (2022a) and has the following form at analysis time step t:

$$Q \sim U(0.001, 5)$$
$$X_A \sim N(X_{f,t}, P_{f,t} + (Q-1) * diag(P_{f,t})) \tag{6}$$
$$Y_t \sim N(X_A, R_t)$$

319

where Q is the estimated forecast inflation scalar and XA is a drawn sample from the analysis distribution. The estimation of XA and Q was completed using a Markov Chain Monte Carlo (MCMC) approach by leveraging the nimble R library (de Valpine et al., 2017). Though not explored in this study, this approach also allows for the definition and estimation of more complex relationships between observations and model forecasts (e.g., nonlinear observation operators).

In this study, the GEF was applied over the EnKF-Miyoshi workflow when (1) more than one observation was assimilated for a single state variable at a given time step or (2) the number of available observations varied throughout a simulation (i.e., changing H). Conversely, the GEF approach was ineffective for cases where only one observation was available at a given time step, as the MCMC algorithm did not converge due to limited data. The EnKF-Miyoshi was applied in these settings.

2.4.6 Simulation schemes

All simulations in this study were performed with 100 ensembles and with a 4-month initialization period starting on 1 Jan of the first year at each site. There were nine different simulations performed for each site in this study which varied in terms of observations assimilated and assimilation method applied. First, two "baseline" runs were completed across all 19 site-years to establish system performance benchmarks. As a lower bound on performance, a free model simulation was performed with no data assimilation. SM sensor observations were also assimilated into the model to represent a reasonable benchmark data assimilation setting. Next, two groups of runs were performed to test the assimilation of RS SM data products: "individual" and "additive" runs. In the "individual" runs, all 4 RS data products were assimilated independently within the system. These runs were performed to compare the value of different RS data products directly. Then, in the "additive" runs, observations from multiple RS data products were jointly assimilated into the system following an additive approach. The first iteration included only ESA observations, and

each subsequent iteration added another data product until all 4 data products were included (i.e., ALL). Data products were added in succession based on availability, such that the first data product tested had the highest average number of observations per year. By sequentially adding new data products, the additional impact of each RS data product could be evaluated. To allow for the application of the GEF in runs with more than one data product, a minimum of 2 observations per day were required for the "additive runs" to ensure the convergence of the MCMC algorithm. For all runs where RS data were assimilated, only site-years after 2014 were investigated due to the limited temporal extent of RS data products.

2.5 System evaluation

This study applied the year-average ensemble weighting strategy, as presented in Kivi et al. (2022), to leverage all available information from the simulations and evaluate the results more accurately. In each site-year simulation, daily weights were assigned to each ensemble as the likelihood of producing the daily estimate given the analysis distribution, and ensemble weights were normalized across the model ensemble for each day. Finally, the average annual weight for each ensemble was computed for each site-year. The application of annual weights in the analysis was the most robust for evaluating yearly estimates (e.g., yield, cumulative NO3 load, cumulative tile drainage).

To evaluate the accuracy and precision of model forecasts for each site-year simulation, we utilized the root mean squared error (RMSE), spectral norm, and weighted variance. RMSE was calculated for each run to quantify changes in accuracy between runs, while the spectral norm and weighted variance were employed to quantify changes in precision. Additionally, to help standardize accuracy measures across site-years, a normalized RMSE (nRMSE) was calculated as :

$$nRMSE\ (\%) = 100 * \frac{RMSE}{\bar{Y}} \tag{7}$$

where $\bar{Y}$ is the average observed value. Changes in accuracy and precision between the free model and SDA were quantified by computing the relative change in each metric for the two runs. For example, for calculating the change in RMSE, we computed :

$$\Delta RMSE\ = \frac{RMSE_{SDA} - RMSE_{FREE}}{RMSE_{FREE}} \tag{8}$$

The coefficient of determination (R2) was used to compare model performance for each state variable more effectively across all observed time points. It was calculated as :

$$R^2 = 1 - \frac{\sum_{t=1}^{T}(Y_t - \bar{X}_t)^2}{\sum_{t=1}^{T}(Y_t - \bar{X}_t)^2 + \sum_{t=1}^{T}(\bar{X}_t - \bar{Y})^2} \tag{9}$$

where Yt is the observed value at the tth observed time step and is the simulated weighted mean at the tth observed time step. All observations (n = T) from all site-years were included in this calculation. Separate R2 values were computed for the Free and SDA results. Weighted mean estimates were computed using annual ensemble weights. In addition spectral norm, and weighted variance were estimated as follows:

$$||Pf||_2 = \sqrt{Maximum\ Eigenvalue\ of\ P_f^H P_f} \tag{10}$$

Where P_f^H represents the conjugate transpose of P_f.

371

$$Variance = \frac{\sum_{i=1}^{N}(w_i - (x_i - \bar{x}w)^2)}{\frac{(N-1)}{N}} \tag{11}$$

372 Where N is the number of ensembles, w_i is the average weight of the ith ensemble, $\bar{x}w$ is the weighted mean across

373 ensembles, and x_i is the forecasted value of the ith ensemble.

374 To identify and quantify relationships between variables, one of two correlation statistics was employed depending

375 on the sample size of the data. When comparing data with a sufficiently large sample size (n > 30), the Pearson

376 correlation coefficient (r) was calculated to determine the direction and strength of the linear relationship between two

377 variables.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} * \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{12}$$

378 When comparing data at the site-level (n ≤ 19), the Spearman rank-order correlation coefficient (rs) was applied,

379 which is a nonparametric measure of the strength and direction of the monotonic relationship between two variables.

380 Though the sample size in this case is still too small for proper application, the Spearman coefficient was applied as

381 its assumptions are less strict than the Pearson coefficient. It is calculated as :

$$r_s = 1 - \frac{6\sum_{i=1}^{n}d_i^2}{n(n^2-1)} \tag{13}$$

382 where the di is the distance between the two ranks of the ith complete pair (i.e., xi and yi). For both coefficients, a test

383 for association between paired samples was used to determine significance.

3. Results

385   The results in section 3.1 evaluate the forecast accuracy and precision of in situ SM SDA in comparison to

386 the free model. Section 3.2 investigates changes in forecast accuracy and precision when assimilating SM RS

387 observations. The individual runs are assessed with regard to their data characteristics (i.e., retrieval interval and single

388 vs. multi-sensor development), and the additive runs are evaluated in succession to determine the relative impact of

389 added observations. Lastly, the impact of RS-based SDA on the forecast accuracy and precision of state variables is

390 investigated and compared.

391

3.1 Assimilation of in situ soil moisture

3.1.1 Impact on soil moisture

394   Across all assimilation time steps, the free model tended to overpredict SM within the two assimilation layers

395 (Fig. 3). Therefore, the adjustment in the SDA analysis step typically reduced the total amount of water in the soil

396 profile. In SM forecasts for the two assimilation layers (i.e., SM3 and SM4), SDA performed as well or better than

397 the free model in accuracy across all site-years. The median change in RMSE due to SDA was -17% and -28% for

398   SM3 and SM4, respectively (Fig. 4). Average forecast precision for SM3 and SM4 was also increased with SDA in

399   84% of cases and by 23% on average.

400   The three site-years where precision was not increased in SDA include OH in 2013 and 2014 and MN in

401   2013. Interestingly, these site-years were among those with the most remarkable improvement in accuracy. This

402   relationship is intuitive considering the nature of the Miyoshi algorithm, which systematically inflates model forecast

403   uncertainty at time steps when observed and forecasted SM distributions differ substantially. At the cost of reduced

404   forecast precision, such inflation allows for the filter to pull the model forecast toward the observed distribution and

405   improve accuracy in future predictions.

406   SDA's constraint of SM3 and SM4 also led to the indirect constraint of SM in deeper soil profile layers. Across all

407   site-years with available data, the median change in RMSE for SDA estimates of SM5, SM6, and SM7 was -14%, -

408   8%, and -14%, respectively. In terms of precision, SDA had an overall positive impact on lower layer SM estimates.

409   The average change in weighted variance was -16%, -6%, and -20% for estimates of SM5, SM6, and SM7,

410   respectively.

3.1.2. Impact on NDVI and crop yield

412   Overall, in comparison to the free model, SDA improved yield estimates by explaining 17.7% more variation

413   in observed yield values and improving yield accuracy in 63% of site-years (Table 3). SDA accuracy was most

414   effective in site-years facing greater water stress. In those cases where yield estimates were improved, SDA often

415   increased available soil water at critical points in crop development, reducing crop soil water deficit factors and

416   increasing yield compared to the free model (Fig. A1). The most evident example of SDA yield improvement is IN in

417   2012, where the free model estimated complete maize crop failure (i.e., no grain yield) due to leaf senescence in mid-

418   July, but SDA estimated a harvestable crop due to increased soil water in the early season (Fig. 5). However, SDA's

419   impact on yield precision was inconsistent; roughly 53% of site-years saw reduced precision in yield estimates.

420   Overall, the free model accurately captured the phenological development of the cropping systems simulated

421   in this study, as demonstrated by the good agreement between observed and simulated NDVI (Fig. A2). SDA's impact

422   on NDVI accuracy was similar to its impact on yield accuracy, such that it typically either increased accuracy due to

423   lessened water stress or did not substantially affect the model performance. A comparison of R^2 values demonstrates

424   that SDA helped to explain 4.8% more variation in observed NDVI values compared to the free model. Intuitively,

425   the site-years with the greatest jumps in NDVI accuracy also usually showed great improvement in yield accuracy,

426   highlighting a well-defined physiological relationship between vegetation and grain yield in APSIM's Maize and Plant

427   modules. SDA's impact on NDVI precision was inconsistent, such that 63% of site-years reduced precision in

428   estimates.

3.1.3 Impact on tile drainage and nitrate load

430   Across the 19 site-years, the free model and SDA showed overall poor performance in estimating annual

431   drainage with nRMSE values ranging from 18-215% with a median value of 54.3% for SDA and from 20-250% in

432   the free model with a median value of 52.4% . In the site-years with the lowest accuracy, APSIM often overpredicted

433   drainage in both the free model and SDA. However, these cases of considerable overestimation in drainage were also

434   among those site-years that were most improved by SDA. 8 of the 11 site-years where SDA improved estimates of

435 annual drainage were cases where the free model overestimated tile flow. In these scenarios, SDA functioned to
436 remove available water from the soil profile and correctly lower the amount of water lost from the system. In the
437 remaining site-years where SDA did not improve drainage accuracy, SDA increased RMSE values by 32% on average.
438 SDA's impact on precision for annual drainage estimates was highly variable. 63% of site-years saw improvement in
439 precision, but four site-years saw an immense reduction in precision (i.e., between 107-146% reduction).
440 APSIM also struggled to accurately estimate the annual NO3 load for the tested site-years in this study (Fig. A3). For
441 the free model, nRMSE values ranged from 23-681% with a median value of 83.7% and, for SDA, nRMSE values
442 ranged from 17-833% with a median value of 86.9%. Considering the SDA constraint, estimates of annual NO3 load
443 were the most poorly constrained in terms of accuracy and precision. SDA's impact on precision was split, increasing
444 precision in 53% of site-years. Accuracy was improved for just 32% of site-years. Among those six site-years where
445 SDA increased NO3 load accuracy, SDA typically reduced estimates compared to the free model. Improved sites were
446 often maize years characterized by high input winter precipitation (Jan-Apr). No clear environmental nor agronomic
447 trend was identified among those 11 site-years where SDA reduced accuracy.

3.2 Assimilation of remote sensing soil moisture products

3.2.1 Individual assimilation runs

450 As expected, the individual influence of each RS data product was heavily dependent on its multi- or single-
451 sensor design and temporal availability. ESA, the most widely available data product, had the greatest impact on both
452 assimilation and downstream state variables. In contrast, assimilation with 1KM and 3KM imposed only slight
453 changes in estimates when compared to the free model. However, ESA did not always lead to improvements in model
454 performance. As demonstrated in Figure 6, ESA results were more variable across site-years in terms of the accuracy
455 of state variable estimates, in some cases leading to great improvement and, in other cases, leading to reduced
456 performance. ESA reduced accuracy in predicting SM3 and SM4 in most site-years (i.e., 80-90%) but was the most
457 effective in improving accuracy in estimates of annual yield, SM6, and SM7. ESA also outperformed the other 3 RS
458 data products in constraining forecast precision for all state variables, improving precision in 70-100% of site-years.
459 Importantly, it showed the greatest reduction in the spectral norm of the SM covariance matrix when compared to the
460 free model, indicating the best constraint of SM precision across the entire profile .
461 Alternatively, the assimilation of SMAP-HB, another temporally frequent RS data product, demonstrated
462 more conservative performance than ESA across state variables. For almost all state variables,
463 it also performed similarly or better than the free model. However, any improvements (or reductions) in forecast
464 accuracy were more moderate than observed with ESA. For example, accuracy in yield estimates was improved more
465 consistently with SMAP-HB (90%) compared to ESA (70%), but the maximum improvement in a tested site-year was
466 a 53% accuracy increase compared to a 95% increase with ESA. This trend in the results highlights an important trade-
467 off when assimilating more certain observations (i.e., ESA-CCI) at a coarse spatial resolution over less certain
468 observations at high spatial resolution (i.e., SMAP-HB) when both data products have unknown biases. In terms of
469 forecast precision, SMAP-HB was overall quite effective in constraining state variable predictions, especially when
470 compared to 1KM and 3KM. However, SMAP-HB underperformed compared to ESA in this regard. 1KM and 3KM

471 both underperformed in accuracy constraint when compared to ESA and SMAP-HB, showing little to no change in
472 RMSE compared to the free model.

473 Considering the four individual runs, more frequent assimilation time steps also led to a more robust
474 performance of the EnKF-Miyoshi workflow. Filter divergence (i.e., when the observed mean falls outside of the 95%
475 credibility interval of the analysis distribution) occurred at 52% and 59% of analysis time steps for 1KM and 3KM,
476 respectively, but occurred at only 44% and 30% of analysis time steps for SMAP-HB and ESA, respectively. For
477 estimates of observation uncertainty, the Miyoshi algorithm predicted greater uncertainty for most RS observations
478 than what is reported in the literature. The average standard error in ESA observations was reported to be 0.02 ± 0.004
479 mm3/mm3 but estimated in this study as 0.05 ± 0.01 mm3/mm3. Standard errors in 1KM and 3KM estimates were
480 reported as 0.05 m3/m3 but estimated by the system to be 0.07 ± 0.02 mm3/mm3 and 0.06 ± 0.01 mm3/mm3,
481 respectively. Miyoshi estimated similar uncertainty values for SMAP-HB observations as reported in the literature
482 (i.e., 0.07 ± 0.02 mm3/mm3).

3.2.2 Additive runs

484 The baseline run for the additive RS-SDA runs was ESA, which demonstrated inconsistent constraint of
485 forecast accuracy and strong constraint of forecast precision. The second most available data product, SMAP-HB, was
486 the next RS data product added to the system. New SMAP-HB observations, on average, imposed a -0.012 mm/mm
487 change in μ_a and a -0.0003 change in P_a for SM1 estimates. For downstream forecast accuracy, the addition of SMAP-
488 HB led to improved and/or more consistent constraints for all state variables except SM7 (Fig. 6). At times, the added
489 information from SMAP-HB dampened the benefit of SDA, reducing accuracy improvement. For forecast precision,
490 +SMAP-HB precision was overall better than the free model but with reduced performance compared to ESA.

491 The subsequent additions of the sparser 1KM and 3KM RS data products were less impactful than the
492 addition of SMAP-HB. New 1KM observations imposed an average -0.0004 mm/mm change in μ_a, and, later, new
493 3KM observations imposed an average -0.0003 mm/mm change in μ_a. These changes were less than 4% of the change
494 imposed by the initial addition of SMAP-HB. Neither additional data product produced a notable average change in
495 P_a. Following these minimal changes in SM1, there was also little change in forecast accuracy and precision for
496 downstream state variables in +1KM and ALL when compared to +SMAP-HB (Fig. 6). Adding 1KM observations to
497 +SMAP-HB did hold some benefit for accuracy and precision in SM3 and SM4, while the effect of the 3KM
498 observations was almost negligible or, even at times, harmful to system performance.

3.2.3 Impact on APSIM model estimates

500 When considering the impact of surface SM data assimilation on downstream model variables, we focus on
501 results where all available RS observations were assimilated for each site . Hereinafter, we refer to the compilation of
502 these runs across the five sites as RS-SDA.

503 Overall, RS-SDA had minor impacts on the soil water profile relative to the free model. Figure 7 demonstrates
504 differences between the free model and RS-SDA in SM1 estimates. For several site-years, RS-SDA estimated
505 significantly higher SM1 values in the early growing season (i.e., May-Jun). In the late season and fall, RS-SDA often
506 estimated lower SM1 values. The impact of these SM1 changes on lower layer SM values seemed to decrease with
507 depth, such that differences between the free model and RS-SDA mean estimates were more subtle in deeper layers.

508 This reduced impact on lower layers is also, in part, a reflection of the increasing total soil water volume represented
509 by soil layers down through the profile (see Table 3 for layer depths). Nonetheless, any differences in SM estimates
510 did not lead to notable changes in accuracy for any SM layer (Table 3). Notable changes were visible in the soil water
511 deficit factors for several growing seasons, such that RS-SDA led to reduced water stress for the growing crop. We
512 speculate that this results from increased available soil water in the root zone during initial periods of crop water
513 uptake (i.e., June). Forecast precision for soil water-related estimates also did not change substantially with
514 assimilation. For SM1 estimates, assimilation substantially reduced variability across site-years (Fig. 7). In many
515 cases, this constraint in the surface soil layer did not propagate into significant changes for precision in lower layer
516 estimates (Fig. 7). However, on average, precision was improved rather than reduced with assimilation, with the most
517 significant downstream constraint in the soil layers closest to the surface.

518 RS-SDA demonstrated partial constraint of aboveground estimates. Considering the R2 values reported in
519 Table 3, RS-SDA explained roughly 4% more variation in yield observations than the free model. All site-years except
520 OH 2015 demonstrated increased yield accuracy, and 60% of sites demonstrated increased yield precision with RS-
521 SDA. Based on these results, there is evidence that surface SM data assimilation can constrain, to some extent,
522 estimates of annual yield. Compared to its effect on yield estimates, RS-SDA was less impactful in its constraint of
523 NDVI. However, since the free model could reasonably predict NDVI (R2 = 0.69), there was less potential for
524 improvement with SM assimilation. 60% of site-years had increased accuracy, and 70% had increased precision for
525 NDVI estimates following SDA.

526 **4. Discussion**

527 **4.1 Sensitivity of APSIM model estimates to in situ soil moisture**

528 In this study, the extent to which in situ SM data assimilation affected APSIM model predictions depended
529 on each state variable's sensitivity to the assimilated state variable (i.e., soil moisture). Deeper layer SM estimates—
530 the most sensitive state variables to SM3 and SM4—were the most strongly constrained. Figure A1 demonstrates the
531 significant linear relationship between daily changes in forecasted SM3 and SM4 due to SDA and daily changes in
532 SM estimates for all deeper soil layers. As expected with a cascading water balance model, the strength of the linear
533 relationship weakens as the vertical distance between soil layers increases. In the model, SM in each layer can
534 influence SM estimates of deeper soil layers, but only indirectly through its influence on the SM in the layer
535 immediately below it. Therefore, the influence of the assimilation layers is reduced by each subsequent SM process
536 down through the soil profile and is weakest in the final soil layer (SM7). Nevertheless, the constraint of SM7 was
537 still quite strong in SDA. By assimilating SM for two upper soil layers, the accuracy of SM estimates improved
538 immensely by simply leveraging the pre-existing model structure (compare to Liu et al., 2017).

539 Crop yield showed the next strongest constraint in SDA. However, as noted in previous studies, its sensitivity
540 to SM SDA was conditional (Lu et al., 2021; Kivi et al., 2022). While changes in SM affected lower layer SM at all
541 analysis time steps, crop yield was only affected when the changes impacted crop water stress. Daily crop water uptake
542 is determined in APSIM as the minimum of crop water demand and soil water supply. Therefore, SDA could only

543   influence crop yield when the soil water adjustment pushed the water supply above or below the demand threshold.

544   For this reason, greater SDA improvement was found in crop yield estimates during water-stressed site-years. Other

545   pathways through which SM can impact crop yield in APSIM, like soil N cycling, did not play a strong role in this

546   study.

547   The impact of SM SDA on APSIM drainage estimates can also be beneficial given certain conditions. As

548   shown in the results, drainage was affected by SM3 and SM4 through 2 pathways: (1) changes in total soil water with

549   assimilation adjustment and (2) changes in crop water uptake due to changes in crop water stress. The role of each of

550   these pathways varied over the year, such that the presence of a growing crop and root system weakened the sensitivity

551   of drainage estimates to changes in the assimilation layers. To quantify this change in sensitivity, we divided daily

552   model forecasts into two categories: with crop water uptake (June-Sept) and without crop water uptake. Then, the

553   relationship between changes in SM3 and SM4 and changes in drainage was analyzed separately for each group. There

554   was no significant linear relationship when looking at SM3 changes in either case. However, the linear relationship

555   between changes in SM4 and changes in daily drainage was stronger when no crop was present ($r = 0.23$, $p = 0.00$)

556   than when a crop was present ($r = 0.14$, $p = 0.00$). This is similar to Hu et al. (2008), who identified notable changes

557   in drainage dynamics during rapid crop growth compared to out-of-season dynamics in SPWS model simulations.

558   Among the state variables considered in SDA, NO3 leaching showed the weakest and most complex

559   relationship with SM3 and SM4 in APSIM. Therefore, logically, the presented system performed most poorly in its

560   constraint of annual NO3 leaching estimates. In APSIM, daily NO3 leaching estimates are computed as the product

561   of two different daily values: estimated NO3 concentration in the lowest soil layer and estimated tile drainage.

562   Therefore, in addition to its impact on drainage, SDA can affect NO3 load estimates through (1) changes in N cycle

563   processes via SM rate factors (see Fig. 2 in Kivi et al., 2022) and (2) changes in the vertical movement of soil water

564   (and N solutes) through the soil profile. In a validation study of APSIM N processes, Sharp et al. (2011) also observed

565   inconsistent model behavior in annual leaching estimates for their experimental site in New Zealand when simulating

566   three years of a potato-rye rotation. Their final calibration of the model only improved one of the annual estimates but

567   did not constrain estimates in the other two years. In fact, many past studies have highlighted nitrate leaching estimates

568   as a broader forecasting challenge (Stewart et al., 2006; Sharp et al., 2011; van der Laan et al., 2014; Brilli et al.,

569   2017). As highlighted already in the literature, missing processes related to snowmelt (Ojeda et al., 2018), and tillage-

570   related infiltration (Malone et al., 2007; Brilli et al., 2017; Ojeda et al., 2018), or preferential flow could help to

571   improve APSIM performance. Though there is still potential for the presented system to improve nitrate leaching

572   estimates, further investigation and constraint of the APSIM N and soil water cycles will be necessary to ensure

573   consistent performance.

574   **4.2   Impact of remote sensing soil moisture data assimilation**

575   The assimilation of RS surface SM observations imposed a far weaker constraint on APSIM state variables

576   compared to the assimilation of the soil sensor observations. For example, the median reduction in SM RMSE ranged

577   from 7-27% across different layers of the soil profile with soil sensor observations, but, with RS observations in RS-

578   SDA, it ranged from roughly 1-5% (Table 3). The weakened constraint with RS-SDA was likely more than an issue

579 of observation inaccuracies. Instead, there is greater evidence to show that changes in SM1 simply had less influence
580 on downstream state variables than changes in SM3 and SM4. This is due, in part, to the increased vertical distance
581 between the surface SM layer (SM1) and other observed soil layers (i.e., SM3-7). The APSIM SoilWat module
582 operates as a cascading water balance model to estimate the movement of water and solutes between and across soil
583 layers (Dokoohaki et al., 2018). Thus, the assimilation adjustment of the SM1 estimate would not be as strongly tied
584 to lower layer estimates when using a top-down approach. Yet, surface SM data assimilation notably changed SM2
585 estimates, the SM estimates for the layer just below it. This result reflects the findings of Lu and Steele-Dunne (2019),
586 who assimilated RS surface SM observations into a surface energy balance model. They found that SDA improved
587 SM estimates in the second layer to a greater extent than in lower layers when comparing estimates to observations.
588 Since observations were not available for SM2 at the study sites, this hypothesis could not be tested within this work.

589   The two assimilation protocols (i.e., assimilation of SM1 vs. assimilation of SM3 and SM4) were also
590 markedly different in the quantity of soil water associated with their assimilation adjustments. Where soil layers 3 and
591 4 corresponded to almost 14% of the soil profile (20 cm depth), the near-surface soil layer only corresponded to about
592 3.6% of the soil profile (5 cm depth). Thus, when considering the top-down effect of SM assimilation on lower layers,
593 each adjustment with RS assimilation had just 25% of the impact of the previous system given the same adjustment
594 in volumetric soil water content. This 5-fold reduction in potential impact closely mirrors the change in RMSE
595 reduction for SM layers highlighted above (i.e., 7-27% to 1-5%). One way to overcome this limitation of surface SM
596 is to leverage the strong covariance between SM1 and SM in nearby layers (i.e., SM2) to directly nudge their values
597 within the analysis time step using, for example, an augmented state vector (e.g., Kivi et al., 2022) or exponential filter
598 approaches (e.g., Albergel et al., 2008).

599

600   RS surface SM data assimilation still demonstrated strong potential for improving APSIM forecasts within
601 this study. First, the assimilation of surface SM improved estimates of crop yield overall when compared to the free
602 model, with a median RMSE reduction of 17.2%. Past RS SM data assimilation studies had similar success in
603 improving crop yield estimates, and several attributed the improvement to increased surface SM and reduced crop
604 water stress with SM assimilation (e.g., Ines et al., 2013; Chakrabarti et al., 2014). We speculate that the model
605 performance indicate that water stress likely played an important role. Although direct observations are not available
606 for crop water uptake to test this hypothesis, we suspect RS-SDA accurately increased available soil water at critical
607 growth stages and, thus, increased crop water uptake.

4.3 Comparison of remote sensing soil moisture data products

609   The four different RS SM data products varied quite broadly in spatial resolution, varying from 30 meters to
610 0.25°. However, their individual assimilation performance seemed to be most closely tied to the temporal availability
611 of observations. ESA with a multi-sensor nature had an average, 219 observations per growing season and showed
612 the best overall constraint of forecast precision and good constraint of forecast accuracy in downstream state variables.
613 Alternatively, the 1KM and 3KM data products, which each had an average of 7 observations per growing season,
614 had almost no impact on forecast accuracy and only a slight impact on forecast precision. Although this study was not

designed to independently test the impact of temporal and spatial resolution on performance, it echoes the findings of Lu and Steele-Dunne . (2019), who found a high temporal resolution to be far more important to assimilation performance than high spatial resolution. They suspected that increased time between assimilation adjustments allowed errors in model structure, inputs, and/or parameters to go unchecked for more extended periods of time, thereby allowing the magnitude of simulation errors to become large and unreasonable. More frequent assimilation helps mitigate the impact of such model errors and improve overall crop model predictions by correcting errors more often (De Lannoy et al., 2007; Pauwels et al., 2007; Lu et al., 2021). Alternatively, in the case of low temporal resolution, a recalibration-based assimilation approach or the inclusion of a bias correction method might be more appropriate (De Lannoy et al., 2007; Curnel et al., 2011).

When comparing RS data products in this study, it is important to recognize that all data products considered in this work are based, in part, on SMAP radiometer data. SMAP-HB merged SMAP brightness temperature data with the HydroBlocks-RTM model, ESA includes SMAP as one of its ten passive microwave sensors, and 1KM and 3KM rely on SMAP for passive microwave information within their derivation. In the first iteration, ESA contributed most of the information provided by the SMAP radiometer to the model and, therefore, imposed large changes in SM1 estimates. Then, with each additional data product, the overall impact on the analysis distribution weakened as much of the new information had already been provided to the system. It is also important to note that given that all data products directly or indirectly are based SMAP, the successive assimilation of these data prodcuts can introduce error covariances between the model runs and the observations. This may potentially result in an over or under estimation of the uncertainty, thereby affecting the performance of the filter. Therefore, further investigation into the impact of including these error covariances between the data products is deemed necessary in order to enhance the accuracy of the EnKF filter.

The Miyoshi algorithm often estimated higher observation uncertainty (R) than the values reported in the literature. This is unsurprising as RS SM data products, like most RS data products, often have poorly characterized uncertainties (Peng et al., 2021). For each data product, uncertainty is typically reported as a standard error value after comparing the data product to a limited set of observations. This estimate does not capture all possible sources of uncertainty and cannot be easily generalized to different places or time points (Huang et al., 2019). Yet, in the additive runs, these uncertainty values were applied uniformly across time and space. Future applications of the GEF scheme could benefit from additional terms in the model that could capture R or the use of the Miyoshi algorithm. These approaches may better estimate observation uncertainties within the system's context.

5. Conclusions

In the study, we assessed the extent to which soil moisture data assimilation can improve APSIM model forecasts. We used a generalizable and novel data-assimilation system to assimilate RS and in situ soil moisture measurements across the U.S. Midwest 19 site-years, and evaluated how direct soil moisture constraint affected downstream model estimates of root-zone soil moisture, crop yield, tile flow, and nitrate leaching. Our results highlighted the capacity of soil moisture data assimilation to improve model estimates of crop yield in water-limited conditions, increasing crop water uptake at critical points in the growing season. Soil moisture data assimilation also improved estimates of soil

651 moisture throughout the profile in most cases but did not well constrain nitrate leaching or tile drainage. This indicates
652 a need for better constraint of both the soil water and soil nitrogen cycles in the APSIM model.
653 This work also lays the groundwork for future regional applications of soil moisture data assimilation. Importantly,
654 our findings reaffirmed soil moisture data assimilation's ability to "localize" gridded weather estimates of precipitation
655 to reflect observed values more accurately. Since cropping systems are highly sensitive to precipitation inputs, this is
656 a strong advantage of soil moisture data assimilation for forecasting applications where coarse-resolution weather
657 drivers are employed. Though RS soil moisture data assimilation could be an effective way to overcome limited
658 availability of in situ data, our work shows that assimilation of in situ surface soil moisture is not as powerful as the
659 assimilation of in situ root-zone soil moisture values in terms of model constraint. If the former is applied, additional
660 constraints or an augmented state-vector approach would be necessary to achieve higher system performance. When
661 selecting a RS soil moisture data product for data assimilation applications, high temporal resolution due to multi-
662 sensor satellite availability and accurately estimated observation uncertainty are two critical components for optimal
663 system performance. To that same point, combining several data products at different spatial resolutions can help to
664 reduce assimilation intervals within the system. Further investigation is needed to independently test the impact of
665 observation sample size (i.e., number of data products), temporal resolution, spatial resolution, and uncertainty on
666 system performance. Moreover, the data products considered in this study do not represent the full range of RS soil
667 moisture data products that are available publicly. This work should be expanded to evaluate data products derived
668 from other satellites/derivations both individually and in combination with other sources to exhaust all available
669 options.

670 **6. Code and data availability**

671 Code and observational data used in this study will be provided upon request.

672 **7. Author contribution**

673 MK was responsible for code development, performing the simulations and writing the manucript. NV contributed to
674 revising the manuscript and providing SMAP-HB dataset. HD was responsible for developing the initial idea, code
675 development, writing and supervising the study.

676 **8. Competing interests**

677 The contact author has declared that neither they nor their co-authors have any competing interests.

678 **9. Acknowledgements**

679 The authors would like to thank all those on the Energy Farm team who made the presented case study possible. In
680 particular, we would like to thank Carl Bernacchi, Bethany Blakely, Michael Masters, Grace Andrews and Heather

Goring-Harford, who made the Energy Farm dataset available and performed the analyses for the nitrate leaching data, and Konrad Taube and Haley Ware, who helped with water collection and water filtering in 2018 and 2019. We also want to thank Caitlin Moore and Evan Dracup, who helped to collect and process much of the other data from the plot. Additionally, we wanted to acknowledge those funding sources that supported the work of the Energy Farm team. First, the data used in this study was funded in part by (1) the Leverhulme Centre for Climate Change Mitigation, funded by the Leverhulme Trust through a Research Centre award (RC-2015-029), (2) the Center for Advanced Bioenergy and Bioproducts Innovation (CABBI) at the University of Illinois, and (3) the Global Change and Photosynthesis Research Unit of the USDA Agricultural Research Service.

10. References

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., & Martin, E. (2008). From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. Hydrology and Earth System Sciences, 12(6), 1323–1337. https://doi.org/10.5194/hess-12-1323-2008.

Chakrabarti, S., Bongiovanni, T., Judge, J., Zotarelli, L., & Bayer, C. (2014). Assimilation of SMOS soil moisture for quantifying drought impacts on crop yield in agricultural regions. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(9), 3867–3879. https://doi.org/10.1109/JSTARS.2014.2315999.

Chen, Y., Zhang, Z., & Tao, F. (2018). Improving regional winter wheat yield estimation through assimilation of phenology and leaf area index from remote sensing data. European Journal of Agronomy, 101, 163–173. https://doi.org/10.1016/j.eja.2018.09.006.

Chighladze, G., Abendroth, L. J., Herzmann, D., Helmers, M., Ahiablame, L., Allred, B., Bowling, L., Brown, L., Fausey, N., Frankenberger, J., Jaynes, D., Jia, X., Kjaersgaard, J., King, K., Kladivko, E., Nelson, K., Pease, L., Reinhart, B., Strock, J., & Youssef, M. (2021). Transforming Drainage Research Data (USDA-NIFA Award No. 2015-68007-23193). National Agricultural Library – ARS – USDA. https://doi.org/10.15482/USDA.ADC/1521092.

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., de Rosnay, P., Ryu, D., & Walker, J. P. (2012). Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products: UPSCALING SOIL MOISTURE. Reviews of Geophysics, 50(2). https://doi.org/10.1029/2011RG000372.

Das, N. N., Entekhabi, D., Dunbar, R. S., Chaubell, M. J., Colliander, A., Yueh, S., Jagdhuber, T., Chen, F., Crow, W., O'Neill, P. E., Walker, J. P., Berg, A., Bosch, D. D., Caldwell, T., Cosh, M. H., Collins, C. H., Lopez-Baeza, E., & Thibeault, M. (2019). The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product. Remote Sensing of Environment, 233, 111380. https://doi.org/10.1016/j.rse.2019.111380.

Dokoohaki, H., Miguez, F.E., Archontoulis, S. and Laird, D., 2018. Use of inverse modelling and Bayesian optimization for investigating the effect of biochar on soil hydrological properties. Agricultural Water Management, 208, pp.268-274.

715 Dokoohaki, H., Kivi, M. S., Martinez-Feria, R., Miguez, F. E., & Hoogenboom, G. (2021). A comprehensive
716 uncertainty quantification of large-scale process-based crop modeling frameworks. Environmental Research Letters,
717 16(8), 084010. https://doi.org/10.1088/1748-9326/ac0f26.

718 Dokoohaki, H., Morrison, B.D., Raiho, A., Serbin, S.P., Zarada, K., Dramko, L. and Dietze, M., 2022a. Development
719 of an open-source regional data assimilation system in PEcAn v. 1.7. 2: application to carbon cycle reanalysis across
720 the contiguous US using SIPNET. Geoscientific Model Development, 15(8), pp.3233-3252.

721 Dokoohaki, H., Rai, T., Kivi, M., Lewis, P., Gomez-Dans, J. and Yin, F., 2022b. Linking Remote Sensing with APSIM
722 through Emulation and Bayesian Optimization to Improve Maize Yield Prediction in the US Midwest.

723 Dorigo, W. A., Zurita-Milla, R., de Wit, A. J. W., Brazile, J., Singh, R., & Schaepman, M. E. (2007). A review on
724 reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. International
725 Journal of Applied Earth Observation and Geoinformation, 9(2), 165–193. https://doi.org/10.1016/j.jag.2006.05.003

726 Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber,
727 A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., …
728 Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future
729 directions. Remote Sensing of Environment, 203, 185–215. https://doi.org/10.1016/j.rse.2017.07.001.

730 Evensen, G. (2003). The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dyn.
731 53, 343-367. https://doi.org/10.1007/s10236-003-0036-9.

732 Fer, I., Gardella, A. K., Shiklomanov, A. N., Campbell, E. E., Cowdery, E. M., De Kauwe, M. G., Desai, A., Duveneck,
733 M. J., Fisher, J. B., Haynes, K. D., Hoffman, F. M., Johnston, M. R., Kooper, R., LeBauer, D. S., Mantooth, J., Parton,
734 W. J., Poulter, B., Quaife, T., Raiho, A., … Dietze, M. C. (2021). Beyond ecosystem modeling: A roadmap to
735 community cyberinfrastructure for ecological data-model integration. Global Change Biology, 27(1), 13–26.
736 https://doi.org/10.1111/gcb.15409.

737 Gao, F., and Zhang, X. (2021). Mapping Crop Phenology in Near Real-Time Using Satellite Remote Sensing:
738 Challenges and Opportunities. Journal of Remote Sensing, 2021, 1–14. https://doi.org/10.34133/2021/8379391.

739 Helmers, M. J., Abendroth, L., Reinhart, B., Chighladze, G., Pease, L., Bowling, L., Youssef, M., Ghane, E.,
740 Ahiablame, L., Brown, L., Fausey, N., Frankenberger, J., Jaynes, D., King, K., Kladivko, E., Nelson, K., & Strock, J.
741 (2022). Impact of controlled drainage on subsurface drain flow and nitrate load: A synthesis of studies across the U.S.
742 Midwest and Southeast. Agricultural Water Management, 259, 107265. https://doi.org/10.1016/j.agwat.2021.107265

743 Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., Samuel-Rosa, A.,
744 Kempen, B., Leenaars, J. G. B., Walsh, M. G., & Gonzalez, M. R. (2014). SoilGrids1km—Global Soil Information
745 Based on Automated Mapping. PloS ONE, 9(8), e105992. https://doi.org/10.1371/journal.pone.0105992.

746 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R.,
747 Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J.,
748 Bonavita, M., … Thépaut, J. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological
749 Society, 146(730), 1999–2049. https://doi.org/10.1002/qj.3803.

750   Huang, J., Ma, H., Su, W., Zhang, X., Huang, Y., Fan, J., & Wu, W. (2015). Jointly Assimilating MODIS LAI and

751   ET Products Into the SWAP Model for Winter Wheat Yield Estimation. IEEE Journal of Selected Topics in Applied

752   Earth Observations and Remote Sensing, 8(8), 4060–4071. https://doi.org/10.1109/JSTARS.2015.2403135

753   Huang, J., Gómez-Dans, J. L., Huang, H., Ma, H., Wu, Q., Lewis, P. E., Liang, S., Chen, Z., Xue, J.-H., Wu, Y., Zhao,

754   F., Wang, J., & Xie, X. (2019). Assimilation of remote sensing into crop growth models: Current status and

755   perspectives. Agricultural and Forest Meteorology, 276–277, 107609.

756   https://doi.org/10.1016/j.agrformet.2019.06.008.

757   Ines, A. V. M., Das, N. N., Hansen, J. W., & Njoku, E. G. (2013). Assimilation of remotely sensed soil moisture and

758   vegetation with a crop simulation model for maize yield prediction. Remote Sensing of Environment, 138, 149–164.

759   https://doi.org/10.1016/j.rse.2013.07.018.

760   Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., & Ustin, S. L.

761   (2009). PROSPECT+SAIL models: A review of use for vegetation characterization. Remote Sensing of Environment,

762   11.

763   Jiang, H., Hu, H., Zhong, R., Xu, J., Xu, J., Huang, J., Wang, S., Ying, Y., & Lin, T. (2020). A deep learning approach

764   to conflating heterogeneous geospatial data for corn yield estimation: A case study of the US Corn Belt at the county

765   level. Global Change Biology, 26(3), 1754–1766. https://doi.org/10.1111/gcb.14885.

766   Kivi, M. S., Blakely, B., Masters, M., Bernacchi, C. J., Miguez, F. E., & Dokoohaki, H. (2022). Development of a

767   data-assimilation system to forecast agricultural systems: A case study of constraining soil water and soil nitrogen

768   dynamics in the APSIM model. Science of The Total Environment, 820, 153192.

769   https://doi.org/10.1016/j.scitotenv.2022.153192.

770   Kumar, S. V., Dirmeyer, P. A., Peters-Lidard, C. D., Bindlish, R., & Bolten, J. (2018). Information theoretic evaluation

771   of satellite soil moisture retrievals. Remote Sensing of Environment, 204, 392–400.

772   https://doi.org/10.1016/j.rse.2017.10.016.

773   de Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N., & Verhoest, N. E. C. (2007). State and bias estimation for soil

774   moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and frequency. Water Resources

775   Research, 43(6). https://doi.org/10.1029/2006WR005100.

776   Lehnert L. W., Meyer H,, Obermeier W. A., Silva B., Regeling B., Thies B., & Bendix J. (2019). "Hyperspectral Data

777   Analysis in R: The hsdar Package." Journal of Statistical Software, 89(12), 1–23. doi: 10.18637/jss.v089.i12.

778   Lievens, H., Reichle, R. H., Liu, Q., De Lannoy, G. J. M., Dunbar, R. S., Kim, S. B., Das, N. N., Cosh, M., Walker,

779   J. P., & Wagner, W. (2017). Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates.

780   Geophysical Research Letters, 44(12), 6145–6153. https://doi.org/10.1002/2017GL073904.

781   Linker, R., and Ioslovich, I. (2017). Assimilation of canopy cover and biomass measurements in the crop model

782   AquaCrop. Biosystems Engineering, 162, 57–66. https://doi.org/10.1016/j.biosystemseng.2017.08.003.

783   Liu, Y., Wang, W., & Hu, Y. (2017). Investigating the impact of surface soil moisture assimilation on state and

784   parameter estimation in SWAT model based on the ensemble Kalman filter in upper Huai River basin. Journal of

785   Hydrology and Hydromechanics, 65(2), 123–133. https://doi.org/10.1515/johh-2017-0011.

786   Liu, Y., Wang, W., & Liu, Y. (2018). ESA CCI Soil Moisture Assimilation in SWAT for Improved Hydrological

787   Simulation in Upper Huai River Basin. Advances in Meteorology, 2018, 1–13. https://doi.org/10.1155/2018/7301314

788   Liu, Z., Xu, Z., Bi, R., Wang, C., He, P., Jing, Y., & Yang, W. (2021). Estimation of Winter Wheat Yield in Arid and

789   Semiarid Regions Based on Assimilated Multi-Source Sentinel Data and the CERES-Wheat Model. Sensors, 21(4),

790   1247. https://doi.org/10.3390/s21041247.

791   Lu, Y., Chibarabada, T. P., Ziliani, M. G., Onema, J. M. K., McCabe, M. F., & Sheffield, J. (2021). Assimilation of

792   soil moisture and canopy cover data improves maize simulation using an under-calibrated crop model. Agricultural

793   Water Management, 252. https://doi.org/10.1016/j.agwat.2021.106884.

794   Luce, G. A. "Optimum corn planting depth – 'Don't plant your corn too shallow.'" University of Missouri Integrated

795   Pest & Crop Management. 6 Apr. 2016.

796   Lu, Y., Dong, J., and Steele-Dunne, S. C. (2019). Impact of Soil Moisture Data Resolution on Soil Moisture and

797   Surface Heat Flux Estimates through Data Assimilation: A Case Study in the Southern Great Plains. Journal of

798   Hydrometeorology, 20(4), 715–730. https://doi.org/10.1175/JHM-D-18-0234.1.

799   Ma, G., Huang, J., Wu, W., Fan, J., Zou, J., & Wu, S. (2013). Assimilation of MODIS-LAI into the WOFOST model

800   for forecasting regional winter wheat yield. Mathematical and Computer Modelling, 58(3–4), 634–643.

801   https://doi.org/10.1016/j.mcm.2011.10.038.

802   Malone, R. W., Huth, N., Carberry, P. S., Ma, L., Kaspar, T. C., Karlen, D. L., Meade, T., Kanwar, R. S., & Heilman,

803   P. (2007). Evaluating and predicting agricultural management effects under tile drainage using modified APSIM.

804   Geoderma, 140(3), 310–322. https://doi.org/10.1016/j.geoderma.2007.04.014.

805   Mishra, V., Cruise, J. F., & Mecikalski, J. R. (2021). Assimilation of coupled microwave/thermal infrared soil moisture

806   profiles into a crop model for robust maize yield estimates over Southeast United States. European Journal of

807   Agronomy, 123. https://doi.org/10.1016/j.eja.2020.126208.

808   Miyoshi, T., Kalnay, E., & Li, H. (2013). Estimating and including observation-error correlations in data assimilation.

809   Inverse Problems in Science and Engineering, 21(3), 387–398. https://doi.org/10.1080/17415977.2012.712527

810   Monsivais-Huertero, A., Graham, W. D., Judge, J., & Agrawal, D. (2010). Effect of simultaneous state–parameter

811   estimation and forcing uncertainties on root-zone soil moisture for dynamic vegetation using EnKF. Advances in

812   Water Resources, 33(4), 468–484. https://doi.org/10.1016/j.advwatres.2010.01.011.

813   Moore, C. E., Haden, A. C., Burnham, M. B., Kantola, I. B., Gibson, C. D., Blakely, B. J., Dracup, E. C., Masters, M.

814   D., Yang, W. H., DeLucia, E. H., & Bernacchi, C. J. (2021). Ecosystem-scale biogeochemical fluxes from three

815   bioenergy crop candidates: How energy sorghum compares to maize and miscanthus. GCB Bioenergy, 13(3), 445–

816   458. https://doi.org/10.1111/gcbb.12788.

817   Mourtzinis, S., & Conley, S. P. (2017). Delineating Soybean Maturity Groups across the United States. Agronomy

818   Journal, 109(4), 1397–1403. https://doi.org/10.2134/agronj2016.10.0581.

819   Naz, B. S., Kurtz, W., Montzka, C., Sharples, W., Goergen, K., Keune, J., Gao, H., Springer, A., Hendricks Franssen,

820   H.-J., & Kollet, S. (2019). Improving soil moisture and runoff simulations at 3 km over Europe using land surface

821   data assimilation. Hydrology and Earth System Sciences, 23(1), 277–301. https://doi.org/10.5194/hess-23-277-2019

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

825 Peng, J., Loew, A., Merlin, O., & Verhoest, N. E. C. (2017). A review of spatial downscaling of satellite remotely
826 sensed soil moisture: Downscale Satellite-Based Soil Moisture. Reviews of Geophysics, 55(2), 341–366.
827 https://doi.org/10.1002/2016RG000543.

828 Raiho, A., Dietze, M., Dawson, A., Rollinson, C. R., Tipton, J., & McLachlan, J. (2020). Towards understanding
829 predictability in ecology: A forest gap model case study [Preprint]. Ecology.
830 https://doi.org/10.1101/2020.05.05.079871.

831 Silva, J. V., and Giller, K. E. (2021). Grand challenges for the 21st century: What crop models can and can't (yet) do.
832 In Journal of Agricultural Science. Cambridge University Press. https://doi.org/10.1017/S0021859621000150.

833 Staton, M. "Pay close attention to soybean planting depth." Michigan State University Extension. 9 May 2012.

834 de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., Wehrhahn Cortes, C.,
835 Rodrìguez, A., Temple Lang, D., & Paganin, S. (2022). NIMBLE: MCMC, Particle Filtering, and Programmable
836 Hierarchical Modeling. doi: 10.5281/zenodo.1211190, R package version 0.12.2, https://cran.r-
837 project.org/package=nimble.

838 de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Temple Lang, D., & Bodik, R. (2017). "Programming
839 with models: writing statistical algorithms for general model structures with NIMBLE." Journal of Computational and
840 Graphical Statistics, 26, 403-413. doi: 10.1080/10618600.2016.1172487.

841 van der Laan, M., Annandale, J. G., Bristow, K. L., Stirzaker, R. J., Preez, C. C. du, & Thorburn, P. J. (2014).
842 Modelling nitrogen leaching: Are we getting the right answer for the right reason? Agricultural Water Management,
843 133, 74–80. https://doi.org/10.1016/j.agwat.2013.10.017.

844 Verburg, K., and CSIRO Division of Soils (1996). Methodology in soil-water-solute balance modelling: an evaluation
845 of the APSIM-SoilWat and SWIMv2 models. Division of Soils divisional report, no. 131.

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

849 Vergopolan, N., Xiong, S., Estes, L., Wanders, N., Chaney, N. W., Wood, E. F., Konar, M., Caylor, K., Beck, H. E.,
850 Gatti, N., Evans, T., & Sheffield, J. (2021a). Field-scale soil moisture bridges the spatial-scale gap between drought
851 monitoring and agricultural yields. Hydrology and Earth System Sciences, 25(4), 1827–1847.
852 https://doi.org/10.5194/hess-25-1827-2021.

853 Vergopolan, N., Chaney, N. W., Pan, M., Sheffield, J., Beck, H. E., Ferguson, C. R., Torres-Rojas, L., Sadri, S., &
854 Wood, E. F. (2021b). SMAP-HydroBlocks, a 30-m satellite-based soil moisture dataset for the conterminous US.
855 Scientific Data, 8(1), 264. https://doi.org/10.1038/s41597-021-01050-2.

856 Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. Remote
857 Sensing of Environment, 236, 111402. https://doi.org/10.1016/j.rse.2019.111402.

858 de Wit, A. J. W. and van Diepen, C. A. (2007). Crop model data assimilation with the Ensemble Kalman filter for

859 improving regional crop yield forecasts. Agricultural and Forest Meteorology 146(1): 38-56.

860 Zhou, H., Wu, J., Li, X., Geng, G., & Liu, L. (2016). Improving soil moisture estimation by assimilating remotely

861 sensed data into crop growth model for agricultural drought monitoring. 2016 IEEE International Geoscience and

862 Remote Sensing Symposium (IGARSS), 4229–4232. https://doi.org/10.1109/IGARSS.2016.7730102.

863 Zhu, P., Shi, L., Zhu, Y., Zhang, Q., Huang, K., & Williams, M. (2017). Data assimilation of soil water flow via

864 ensemble Kalman filter: Infusing soil moisture data at different scales. Journal of Hydrology, 555, 912–925.

865 https://doi.org/10.1016/j.jhydrol.2017.10.078.

866

867

868
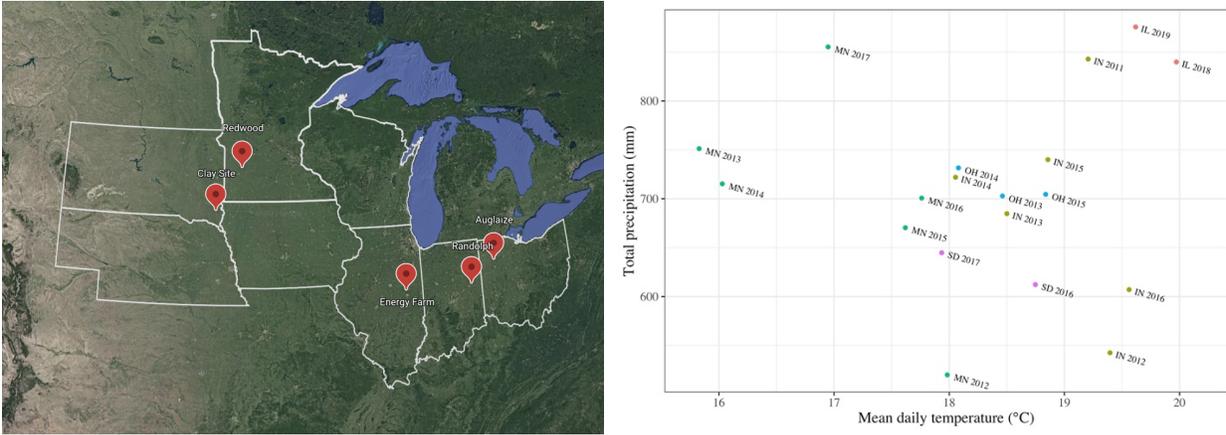
Figure 1. (A) Site map (ESRI) and (B) scatterplot demonstrating site-year total precipitation and average daily temperature (°C) for each site-year between April and October. Climate information was extracted and averaged across the 10 ERA5 weather ensembles for each site-year.

869
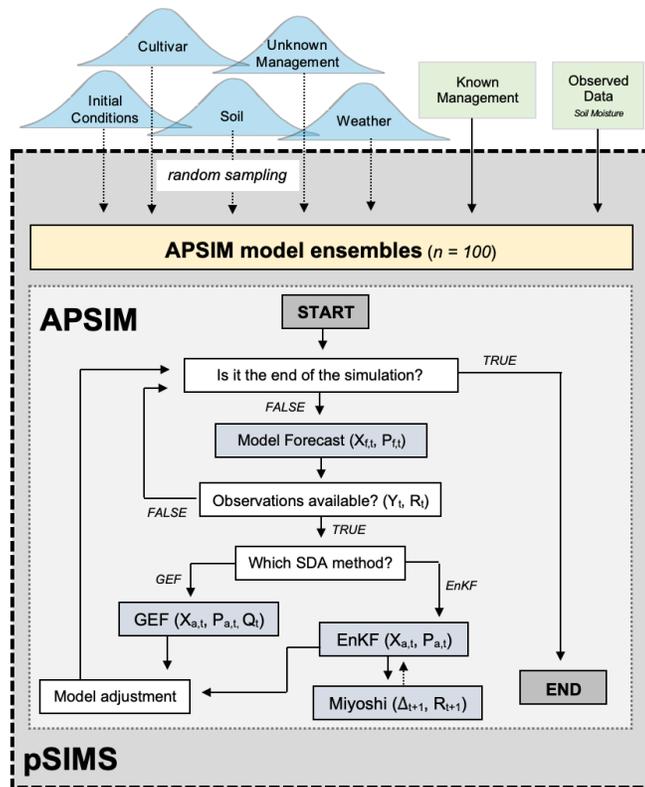
Figure 2. Schematic demonstrating the workflow of the data assimilation system. System inputs represented by blue Normal distributions have incorporated uncertainty in this study, while green rectangles represent known values that were included as constants.

870
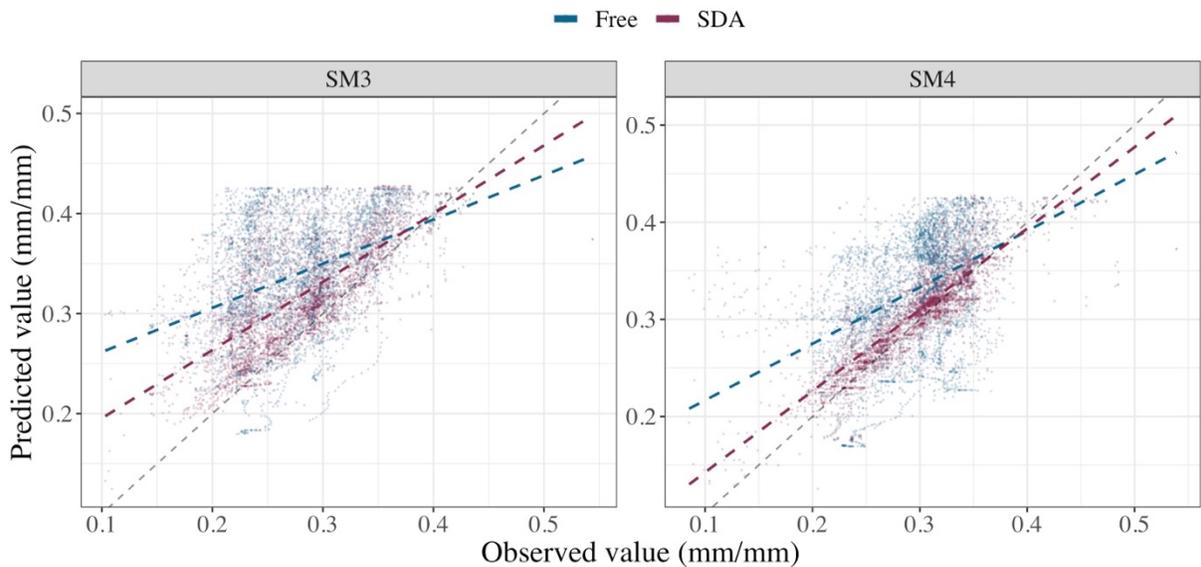
Figure 3. One-to-one plots for soil moisture estimates (mm/mm) in the two assimilation layers for the free model and in situ SDA across all analysis time-steps and site-years. The least-squares regression line is shown for both schemes next to the black dashed 1:1 line, demonstrating a perfect fit.

871
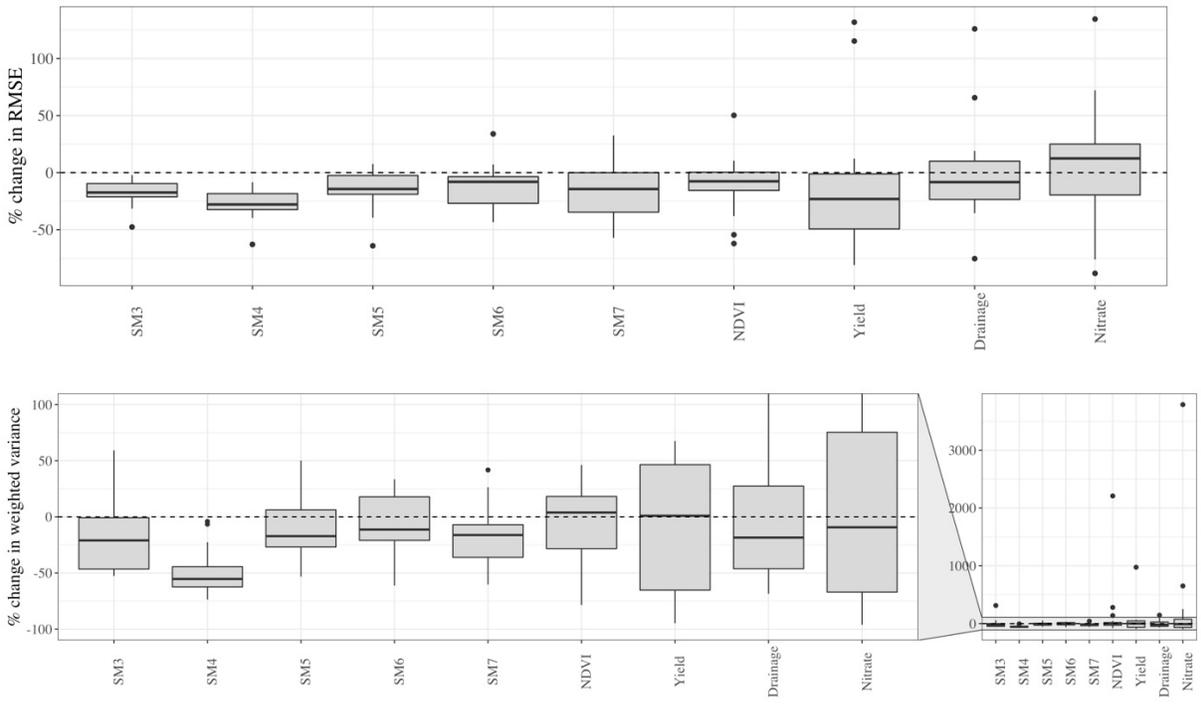
Figure 4. Boxplots demonstrating the distribution of relative change in (a) accuracy (RMSE) and (b) precision (weighted variance) due to in situ SDA for each state variable across all site-years. The relative change is computed with respect to the free model run, with negative values indicating SDA improvement.

872
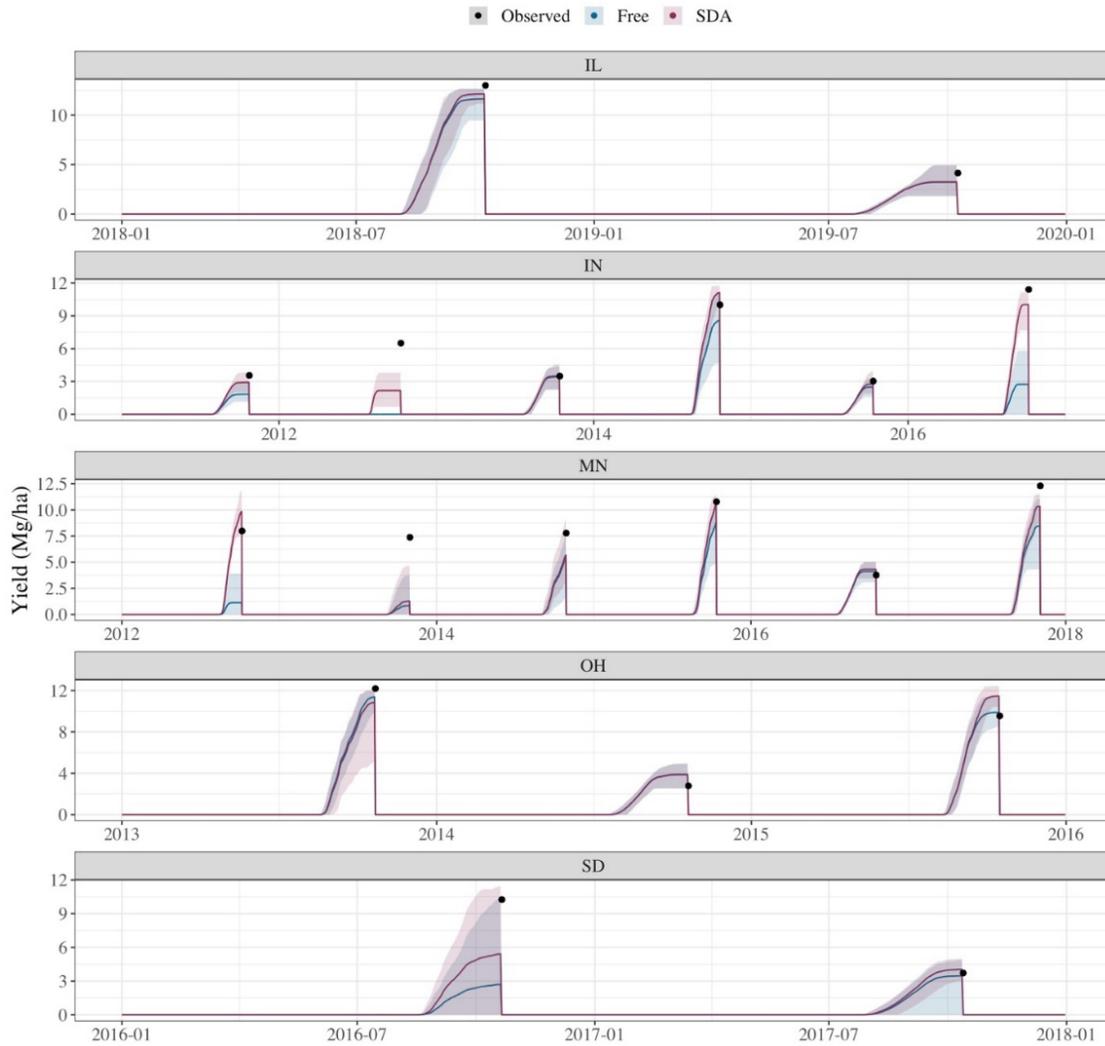
Figure 5. Time series of yield estimates for the free model and in situ SDA with mean daily estimates demonstrated with line graphs and the 95% credible intervals demonstrated by the shaded regions. Black points represent the observed harvest date and yield for each site-year.

873
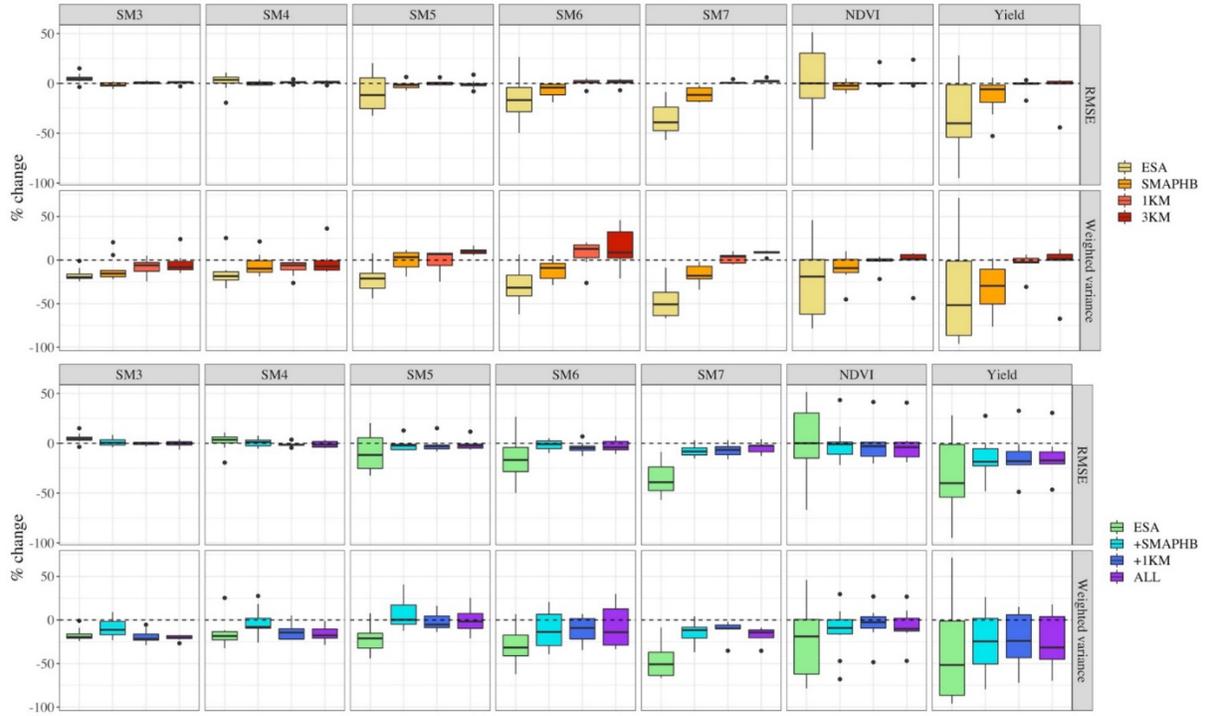
Figure 6. **Boxplots demonstrating the distribution of relative change (%) in state variable accuracy (RMSE) and precision (weighted variance) for the (a) individual and (b) additive runs across all site-years. Change is computed relative to the free model results. Negative values indicate improvement (e.g., (RMSES − RMSEF) / RMSEF).**

874

875
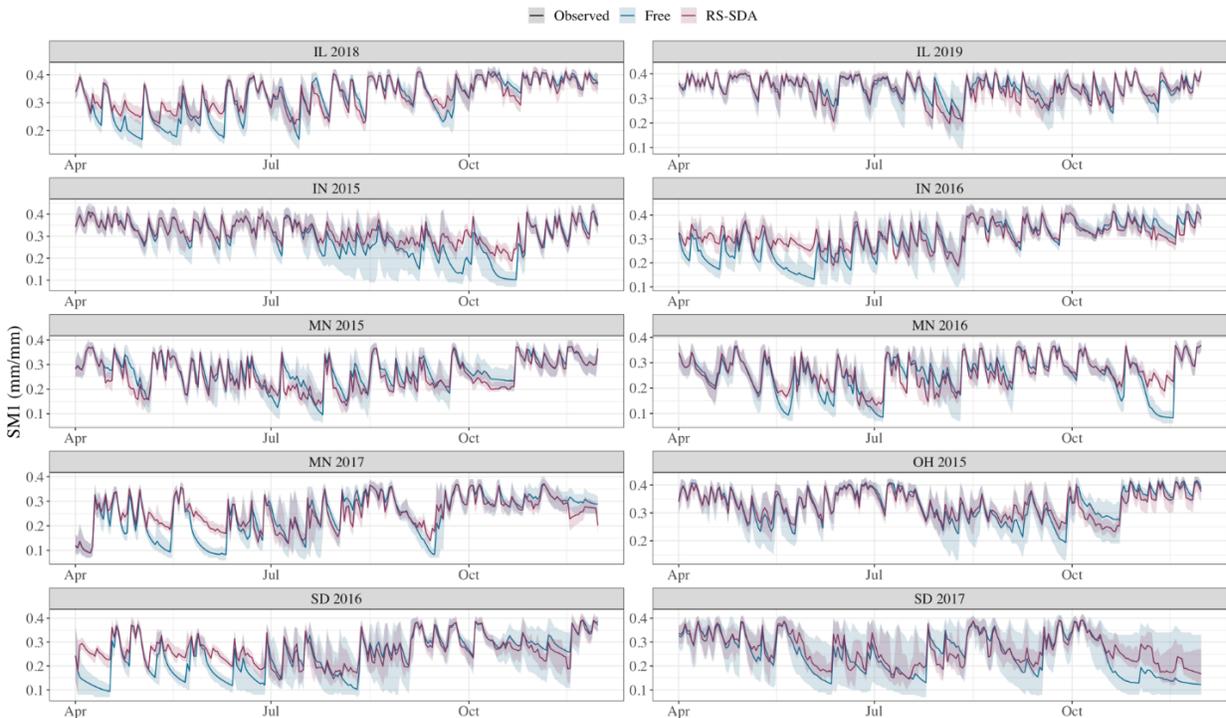
Figure 7. Time series of SM1 estimates from the free model and RS-SDA with the mean daily estimates demonstrated with line graphs. The shaded regions indicate 95% credibility intervals.

876

Table 1. Overview of remote sensing soil moisture data products.

Product	Product ID	Temporal coverage	Temporal frequency	Spatial resolution	Average data availability	Average observation variance	Reference
ESA-CCI	ESA	1978-2019	1-2 days	0.25°	219 days	0.0003	Dorigo et al. (2017)
SMAP-Hydroblocks	SMAP-HB	2015-2019	1-3 days	30 m	127 days	0.0050	Vergopolan et al. (2021b)
SMAP-Sentinel1	1KM/3KM	2015-now	12 days	1 km/3 km	7 days	0.0025	Das et al. (2019)

[a]Availability is calculated after removing observations in the winter months (i.e., Dec-Mar) and is given on a per-year basis.

877

878

Table 2. Overview of system configuration for the nine runs performed in this study. SDA methods include the Ensemble Kalman Filter (EnKF) coupled with the Miyoshi algorithm, and the Generalized Ensemble Filter (GEF). The former method of these two methods provided systematic estimates of R applied within the system, but the latter method used literature values. The state variables included in Xf are given.

Run group	Name	SDA method	R estimates	Temporal extent	State variable(s)	Observation(s)
Baseline	Free	N/A	N/A	2011-2019	N/A	N/A
	SDA	EnKF	Miyoshi	2011-2019	SM3, SM4	In situ soil sensor
Individual Runs	ESA	EnKF	Miyoshi	2015-2019	SM1	ESA
	SMAP-HB	EnKF	Miyoshi	2015-2019	SM1	SMAP-HB
	1KM[a]	EnKF	Miyoshi	2015-2019	SM1	1KM
	3KM[a]	EnKF	Miyoshi	2015-2019	SM1	3KM
Additive Runs	+SMAHB	GEF	Literature	2015-2019	SM1	ESA, SMAP-HB
	+1KM[a]	GEF	Literature	2015-2019	SM1	ESA, SMAP-HB, 1KM
	ALL[a]	GEF	Literature	2015-2019	SM1	ESA, SMAP-HB, 1KM, 3KM

[a] Observations for 1KM and 3KM were not available for IL, and thus simulations were not performed for the site.

879

880

Table 3. Summary statistics to quantify the impact of in situ SDA (IS) and RS-SDA (RS) on forecast accuracy of APSIM state variables. The "Ns" column indicates the number of site-years with available data for each state variable and each run, and the "ns" column indicates the total number of observations across site-years for each run. A subscript (F) denotes a value computed for the free model estimates, a subscript (IS) denotes a value for the in-situ SDA estimates, and a subscript (RS) denotes a value for RS-SDA runs. The median change (D) in RMSE was computed for both runs. Two values for R2F are given for the different data subsets demonstrated in the "N" and "n" columns.

State variable	Depth (cm)	N_{IS} (N_{RS})	n_{IS} (n_{RS})	Δ RMSE$_{IS}$	Δ RMSE$_{RS}$	R^2_F	R^2_{IS}	R^2_{RS}
SM3 *mm/mm*	9.1 – 16.6	19 (10)	12252 (5592)	-17.4%	-0.9%	0.49 (0.48)	0.57	0.48
SM4 *mm/mm*	16.6 – 28.9	19 (10)	12735 (6141)	-27.9%	-2.8%	0.52 (0.43)	0.73	0.43
SM5 *mm/mm*	28.9 – 49.3	17 (8)	11325 (5101)	-14.3%	-2.6%	0.45 (0.45)	0.38	0.45
SM6 *mm/mm*	49.3 – 82.9	19 (10)	12846 (6169)	-8.0%	-1.0%	0.42 (0.43)	0.34	0.42
SM7 *mm/mm*	82.9 – 138	9 (6)	5715 (3265)	-14.3%	-5.4%	0.43 (0.44)	0.34	0.43
NDVI *unitless*	-	19 (10)	244 (134)	-7.6%	-1.8%	0.62 (0.69)	0.66	0.71
Yield *Mg/ha*	-	19 (10)	19 (10)	-23.1%	-17.2%	0.55 (0.53)	0.73	0.59
Annual drainage *mm*	-	19	19	-8.3%	-	0.47	0.46	-
Annual NO$_3$ load *Kg NO$_3$-N/ha*	-	19	19	+12.5%	-	0.42	0.45	-

881