A Comparison of NOAA Modeled and In Situ Soil Moisture Estimates Across the Continental United States

Peter James Marinescu[1], Kyle Hilburn[1], Daniel Abdi[2], and Isidora Jankov[3]

[1]Cooperative Institute for Research in the Atmosphere / Colorado State University
[2]Cooperative Institute for Research in Environmental Sciences
[3]Global Systems Laboratory, NOAA/OAR

November 21, 2022

Abstract

Three estimates of soil moisture from National Oceanic and Atmospheric Administration (NOAA) programs are compared. The estimates are from a high-resolution atmospheric model with a land surface model, a hydrologic model and in situ observations. Both models demonstrate wetter soil moisture in dry regions and drier soil moistures in wet regions, as compared to the in situ observations. These soil moisture differences occur at most soil depths but are larger at the deeper depths below the surface (100 cm). In terms of soil moisture variance, both models generally have lower standard deviations as compared to the in situ observations, except for near the surface where the in situ and high-resolution, land surface model compare well. These NOAA soil moisture estimates are used for a variety of forecasting and societal applications, and understanding their differences provides important context for their applications and can lead to model improvements.

Peter J. Marinescu[1], Kyle Hilburn[1], Daniel Abdi[2], Isidora Jankov[3]

[1] Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado.

[2] Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado.

[3] Global Systems Laboratory, NOAA/OAR, Boulder, Colorado.

Corresponding author: Peter J. Marinescu (*peter.marinescu@colostate.edu*)

Key Points:

- NOAA modeled soil moisture is drier in wet regions and wetter in dry regions as compared to in situ observations.
- NOAA modeled soil moisture has lower variance than in situ observations.
- The differences between NOAA modeled soil moisture and in situ data are larger at deeper soil depths as compared to near the surface.

Abstract

Three estimates of soil moisture from National Oceanic and Atmospheric Administration (NOAA) programs are compared. The estimates are from a high-resolution atmospheric model with a land surface model, a hydrologic model and in situ observations. Both models demonstrate wetter soil moisture in dry regions and drier soil moistures in wet regions, as compared to the in situ observations. These soil moisture differences occur at most soil depths but are larger at the deeper depths below the surface (100 cm). In terms of soil moisture variance, both models generally have lower standard deviations as compared to the in situ observations, except for near the surface where the in situ and high-resolution, land surface model compare

well. These NOAA soil moisture estimates are used for a variety of forecasting and societal applications, and understanding their differences provides important context for their applications and can lead to model improvements.

Plain Language Summary

Soil moisture is an essential variable coupling the land surface to the atmosphere. Accurate estimates of soil moisture are important for predicting where clouds will form, assessing drought and fire weather risks, and assisting with decisions for agricultural production. There are multiple estimates of soil moisture available, and in this study, we compare three different types of soil moisture estimates from the National Oceanic and Atmospheric Administration (NOAA), including two types of soil moisture models and direct observations, as well as direct observations from the United States Department of Agriculture. Both models have soil moisture values that are too wet in dry regions of the United States and too dry in wet regions. The modeled soil moisture also does not change as rapidly from day-to-day as they do in the direct observations of soil moisture. Further analysis at different soil depths shows which depths have the largest differences between these three soil moisture estimates. Understanding these NOAA soil moisture estimate differences is important for forecasters and decision makers and can be useful for the further development of atmospheric and land surface models.

1 Introduction

Knowledge of soil moisture is essential for many Earth system applications, such as forecasting cloud formation (e.g., Ek and Holtslag, 2004), monitoring drought, flood and fire risks (e.g., Svoboda et al., 2002; Rigden et al., 2020), and providing instrumental information for agricultural production (e.g., Madadgar et al., 2017). As such, several advancements in the estimation and utilization of soil moisture have recently transpired. For example, in efforts to improve the accuracy of numerical weather prediction (NWP) and climate models, model developers have focused on increasing the coupling between the land surface and atmosphere components of their data assimilation systems to eliminate persistent atmospheric prediction biases (e.g., Benjamin et al., 2022). Furthermore, NWP models are also beginning to explore the direct assimilation of new soil moisture observations (e.g., Carrera et al., 2019; Muñoz-Sabater et al., 2019; Lin and Pu, 2020).

For these reasons, it is critical to understand the differences in the available soil moisture estimates from models or observations that are used in science, forecasting or agricultural applications. Different estimates of soil moisture exist across the continental United States (CONUS), each with its own benefits and shortfalls. For example, in situ observations are often considered to be the most accurate and are therefore used as a benchmark. However, they have limited spatial coverage (e.g., Quiring et al., 2016). Other products, such as those from low Earth orbiting satellites, have lower temporal and spatial resolution (e.g., Liu et al., 2016). Soil moisture estimates from models depend on many assumptions and reflect the influence of observation-based data to different degrees (e.g., Smirnova et al., 1997; Huang et al., 1996; Mitchell et al., 2004). Several recent studies have made significant progress in comparing many of the available soil moisture estimates. Some studies have compared soil moisture temporal variability and memory between many large-scale land surface models (LSMs) and in situ soil monitoring networks across the CONUS, noting certain biases and uncertainties in the various estimates (e.g., Robock et al., 2003; Xia et al., 2014; 2015a; Dirmeyer et al., 2016). Recent studies have extended soil moisture comparisons to include soil moisture retrievals from new satellite platforms (e.g., Shellito et al., 2016; Pan et al., 2016; Ford and Quiring, 2019).

While various LSMs have been compared to in situ observations for several decades, these prior studies have primarily focused on models with horizontal resolutions on the scales of 1/8 degree or larger (i.e., the North American Land Data Assimilation System models; Mitchell et al., 2004; Xia et al., 2012). Given the local, mesoscale variability of soil moisture process and the subsequent impacts of soil moisture on atmospheric prediction (e.g., Koster et al., 2004; Taylor et al., 2011), high-resolution models should also be tested. Recently, Min et al. (2021) compared near-surface atmospheric and soil variables from the High-Resolution Rapid Refresh (HRRR) model (Dowell et al., 2022; James et al., 2022), which has horizontal grid spacing of

3 km, to observations from the New York State Mesonet. They found that soil moisture was underestimated in HRRR, which contributed to warm and dry biases in atmospheric forecasts. However, how does soil moisture from the HRRR model (NOAA's current operational, convection-allowing model) compare to soil moisture estimates across all of CONUS?

In this study, we focus on a comparison of soil moisture estimates from NOAA. In particular, we uniquely include soil moisture estimates across CONUS from the NOAA operational HRRR model, which utilizes the RUC land surface model (RUC LSM, Smirnova et al., 1997). We also include the NOAA Climate Prediction Center (CPC) leaky-bucket hydrological model (Huang et al., 1996; van den Dool et al., 2003) and in situ observations from two nationwide networks: the NOAA/NCEI United States (US) Climate Reference Network (USCRN; Bell et al., 2013) and the US Department of Agriculture Soil Climate Analysis Network (SCAN; Schaefer et al., 2007). This work provides an assessment of the similarities and differences of soil moisture amounts and variance across three different products, which are all used in various operational and research applications.

2 Soil Moisture Data

2.1 In Situ Observations

The study uses in situ soil moisture observations from two nationwide networks. USCRN provides climate monitoring measurements of atmospheric and soil properties. To increase the coverage of the in situ observations, SCAN is also included. SCAN uses similar sensors to USCRN (i.e., Hydra Probe sensors) and typically have volumetric soil moisture (VSM; ${m_{water}}^3$ / ${m_{soil}}^3$) measurements at the same soil depths (~5, ~10, ~20, ~50, and ~100 cm) as USCRN. Only data from these five levels are used for consistency. Dirmeyer et al. (2016) also found that these two networks have similar error variances. The observations represent point measurements of the soil moisture at specific sites across the US. Daily data are used in this study and represent the average VSM of the entire 24-hour period based on local standard time.

2.2 HRRR Model

The HRRR model is NOAA's operational, convection-allowing model, which has 3 km horizontal grid spacing and covers CONUS with a one-hour temporal refresh rate (Dowell et al., 2022; James et al., 2022). In this study, we use HRRRv3, which was operational between 12 July 12 2018 and 2 December 2020. The HRRR model utilizes a one-dimensional land surface model (RUC LSM; Smirnova et al., 1997), which predicts heat and moisture transfer vertically throughout the soil column. The RUC LSM has undergone several enhancements over the years, including increasing its resolution and incorporating new features, such as snow and ice models (Smirnova et al., 2000; 2016). The current version predicts VSM at nine vertical levels (0, 1, 4, 10, 30, 60, 100, 160 and 300 cm) and utilizes cycling of soil conditions over several years to better capture the soil moisture state. The HRRR utilizes moderately coupled land data assimilation, meaning that near-surface atmospheric data assimilation increments are used to adjust the soil analysis (e.g., Benjamin et al., 2022). Given the recent and continued development of the HRRR data assimilation system and land surface model, it is critical for assessments of HRRR's soil moisture to other estimates. This study provides a benchmark for HRRR soil moisture estimates in support of the continued development of NOAA's land surface prediction capabilities in the Unified Forecast System. The focus of this study is on the analyzed soil moisture field, rather than forecast fields, and thus our results are most directly relevant to the LSM and data assimilation system development.

2.3 CPC Leaky-Bucket Model

The CPC soil moisture product utilizes a leaky-bucket model that solves the time tendency equation in soil moisture over a region from several inputs: precipitation minus evapotranspiration, net streamflow divergence and net groundwater loss (Huang et al., 1996; van den Dool et al., 2003). These inputs to the time tendency equation for soil moisture have been improved over the years with new observations and parameterizations (Fan and van den Dool, 2004; Arevalo et al., 2021). The CPC model provides 1.6 m deep integrated soil moisture (ISM, mm), and these estimates are provided daily for each of the NOAA climate divisions across the United States (Guttman and Quayle, 1996). There are typically about 7-10 climate divisions per state, although there are fewer for states with smaller geographical areas, like those in the Northeast United States. The CPC soil moisture data are used as an input to the United States Drought Monitor (Svoboda et al., 2002) and continues to be used as a reference data set in various soil moisture application studies, from assessing soil moisture impacts on carbon fluxes (e.g., Yao et al., 2021) to understanding climate impacts on agricultural production (e.g., Atiah et al., 2022).

3 Methods

3.1 ISM and VSM Comparisons

Since the CPC product only provides 1.6 m ISM, VSM values from the in situ and HRRR data are vertically integrated in order to compare 1.6 m ISM in all three datasets. The VSM values are assumed to represent the mean value over a depth between the midpoints of the specified levels, as has been done in other studies (e.g., Dirmeyer et al., 2016; Ford and Quiring, 2019). For example, the 10 cm VSM observation in the in situ data is assumed to represent the average VSM for the layer between 7.5 cm (i.e., the midpoint between 5 and 10 cm) and 15 cm (i.e., the midpoint between 10 and 20 cm). The 100 cm VSM in the in situ data is also assumed to be constant to the depth of 160 cm. The HRRR ISM calculations are better constrained than the in situ ISM calculations, since the HRRR VSM data span 3.0 m below ground using 9 levels. VSM values are also compared between the HRRR and in situ data to glean whether certain vertical levels are driving differences between these two datasets. An understanding of soil moisture differences at varying depths is also critical since soil moisture's role in Earth system processes is depth dependent. In terms of spatial comparisons, the HRRR and CPC data are linearly interpolated to the locations of the in situ stations. The analysis is completed over a ˜2.4 year period from 12 July 2018 through 2 December 2020, which is the timeframe that HRRRv3 was operational. By confining the analyses to this time frame, uncertainties associated with model version changes are avoided.

3.2 Quality Control

In situ data provide the most direct physical estimate of soil moisture, but it is important to ensure that the in situ data are of the highest quality. As such, a variety of quality control procedures are undertaken. First, in situ data are only included if they have VSM values available at all five vertical levels (˜5cm, ˜10cm, ˜20cm, ˜50cm, and ˜100cm), since missing data could lead to larger uncertainties in the ISM calculation. Second, in situ stations directly along the coast are removed due to unphysical spatial interpolations from the CPC and HRRR data. From the remaining in situ data, we estimate the ratio of error variance to ISM variance using the method defined in Robock et al. (1995) and used in more recent soil moisture comparisons studies (e.g., Dirmeyer et al., 2016). In essence, soil moisture can be well approximated by a red-noise process (i.e., first-order Markov process; e.g., Delworth and Manabe, 1988; Vinnikov and Yeserkepova, 1991) with the natural logarithm of the soil moisture autocorrelation (r) decreasing linearly with increased lag times (?). For this study, autocorrelations are computed for ? of 1-30 days for each station's ISM daily anomalies. A linear fit is applied to the ln(r) versus ? data and is extrapolated to ? = 0. Deviations from 1 at ? = 0 can be used to solve for to the ratio of error variance to ISM variance (e.g., Robock et al., 1995). For this study, stations where this ratio is greater than 0.08 are removed. This error ratio threshold is in-line with estimates of the mean error ratio of the USCRN and SCAN networks (Dirmeyer et al., 2016). While this error variance ratio threshold results in the removal of 63 (˜27%) of the 235 in situ stations, it provides more confidence that only the highest quality in situ observations are being used in the analyses. Even with the significant reduction of the in situ data, the stations span the entirety of the CONUS (Figure 1, dots). Different thresholds, autocorrelation lengths and dataset lengths were tested and did not qualitatively impact the results.

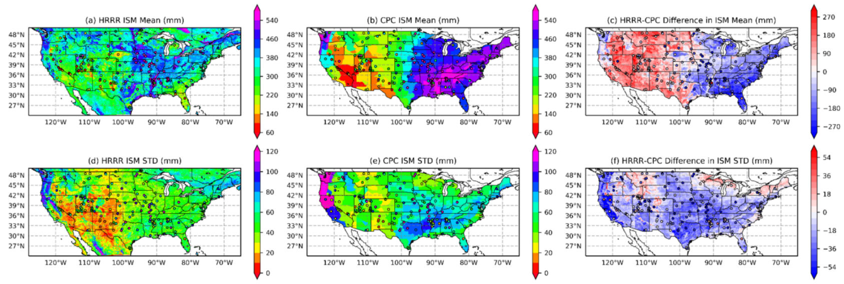


Figure 1: 1.6 m ISM values and standard deviations temporally averaged over the study period. (a) represents the HRRR model, (b) represents the CPC model, and (c) represents their difference. (d-f) are the same as (a-c), except for the ISM standard deviations. Filled circles represent locations of the 172 in situ observations from the USCRN and SCAN networks that pass quality control checks, as described in the text. In (c) and (f), the filled circles represent the HRRR minus in situ differences, while the filled circles represent the in situ station ISM mean and standard deviations, respectively, for the other panels.

3.3 Quintile Analysis

To determine whether differences in soil moisture estimates vary in different soil moisture regimes (i.e., wetter versus drier conditions), we composite the comparisons over locations with similar soil moisture amounts. The ISM or VSM values for each dataset are averaged temporally for each location and are then averaged among the available datasets (i.e., all three datasets for ISM and the in situ and HRRR datasets for VSM). Using this mean, the locations are separated into five quintiles. For example, locations with a mean ISM estimate that is greater than or equal to the 0^{th} percentile ISM and less than the 20^{th} percentile ISM are placed in the lowest quintile of soil moisture amounts (i.e., driest locations). These locations are termed $L_{00\text{-}20}$. Similarly, locations that fall within the 20^{th}-40^{th}, 40^{th}-60^{th}, 60^{th}-80^{th} and 80^{th}-100^{th} percentiles are termed $L_{20\text{-}40}$, $L_{40\text{-}60}$, $L_{60\text{-}80}$and $L_{80\text{-}100}$, respectively and represent dry to wet soil moisture regimes. There are either 34 or 35 locations that are included in each of these quintiles.

4 Soil Moisture Amounts

4.1 ISM Mean Comparisons

Stark, regionally-dependent differences are apparent in the ISM estimates (Figure 1c,f). The HRRR ISM values (Figure 1a) have a more muted range than the CPC ISM values (Figure 1b). The HRRR ISMs are larger (i.e., wetter) than the CPC ISMs across the drier regions of CONUS (most of the western CONUS; Figure 1c) and are smaller (i.e., drier) than the CPC ISMs in the wetter regions of CONUS (eastern half of CONUS and the coastal regions of the Pacific Northwest). This regional dependence is also clearly associated with varying soil moisture amounts.

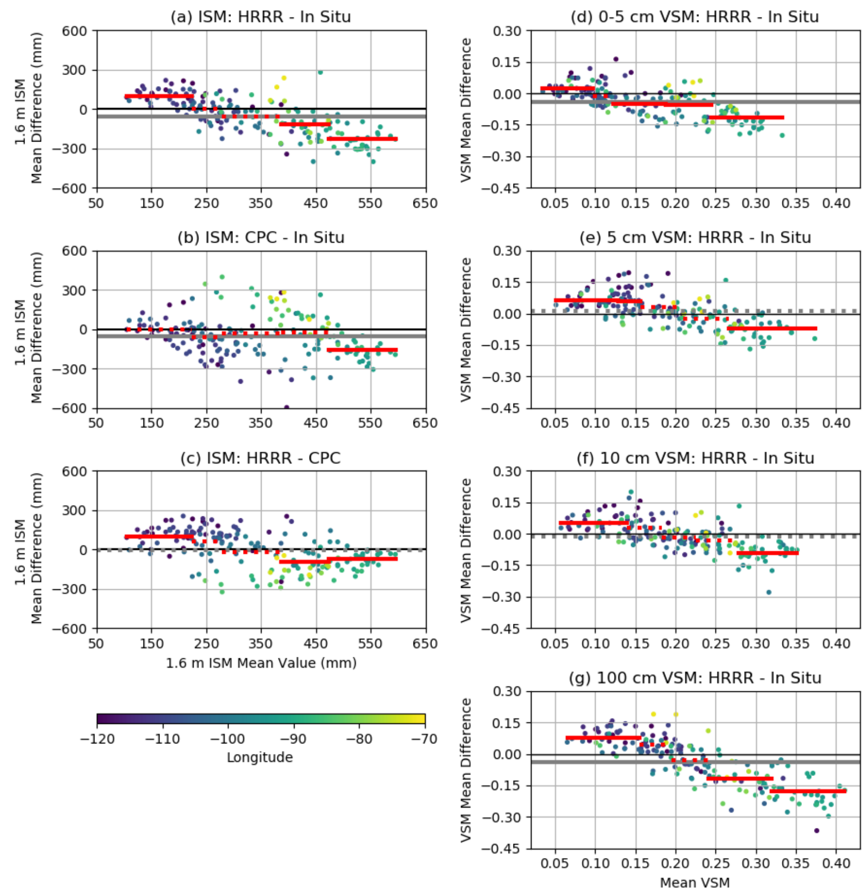


Figure 2: Comparisons of the 1.6 m ISM means between (a) HRRR and in situ values, (b) CPC and in situ values, and (c) HRRR and CPC values. Note, (c) only shows results from the same locations where in situ data are available for consistency. (d-g) shows comparisons of VSM between the HRRR and in situ data for 4 different level combinations: (d) surface in HRRR to 5 cm in the in situ data, (e) 5 cm for both, (f) 10 cm for both, and (g) 100 cm for both. The gray, horizontal line represents the mean difference for the entire dataset, while red horizontal lines represent the mean difference over the five quintiles. Horizontal gray and red lines are solid if their associated data differences pass statistical significance using a paired Student's t-test at the 99.9th percentile.

Figure 2 includes the in situ ISM estimates within the comparisons in order to better quantify the differences between the three datasets. When compared to both in situ ISM (Figure 2a) and CPC ISM (Figure 2c), HRRR is wetter in $L_{00\text{-}20}$ (+101% and +97%, respectively, when taking the mean percentage difference for all locations in $L_{00\text{-}20}$) and drier in $L_{80\text{-}100}$ (-34% and -13%, respectively). Even though there are fewer moisture quintiles with statistically significant differences between the CPC and in situ ISMs, the longitude locations show large regional differences between the CPC and in situ datasets, even within a given quintile range. For example, $L_{40\text{-}60}$and $L_{60\text{-}80}$ show minimal differences between the mean CPC and in situ ISMs (Figure 2b, red lines), but at eastern locations the CPC ISMs are generally larger (Figure 2b., yellow and green dots) and at western locations the CPC ISMs are generally smaller (Figure 2b, blue dots). As was also shown in Figure 1, Figure 2 demonstrates clear regional relationships to these ISM differences between the datasets.

Despite larger, systematic, moisture-regime-dependent differences between HRRR and in situ observations

(Figure 2a), HRRR aligns more closely to the in situ observation than does the CPC for some locations. These occurrences are typically associated with local soil characteristics or topography, which are not captured in the coarser climate division regions used in the CPC product. For example, based on comparisons with the in situ ISM, the HRRR produces similarly dry soil moisture conditions within the Sand Hills region of Nebraska (e.g., -101.4 W, 42.1 N in Figure 1). The temporally averaged mean ISM for this location is 138 mm and 165 mm for the in situ and HRRR data, respectively, as compared to 372 mm in the CPC data (2-3x larger).

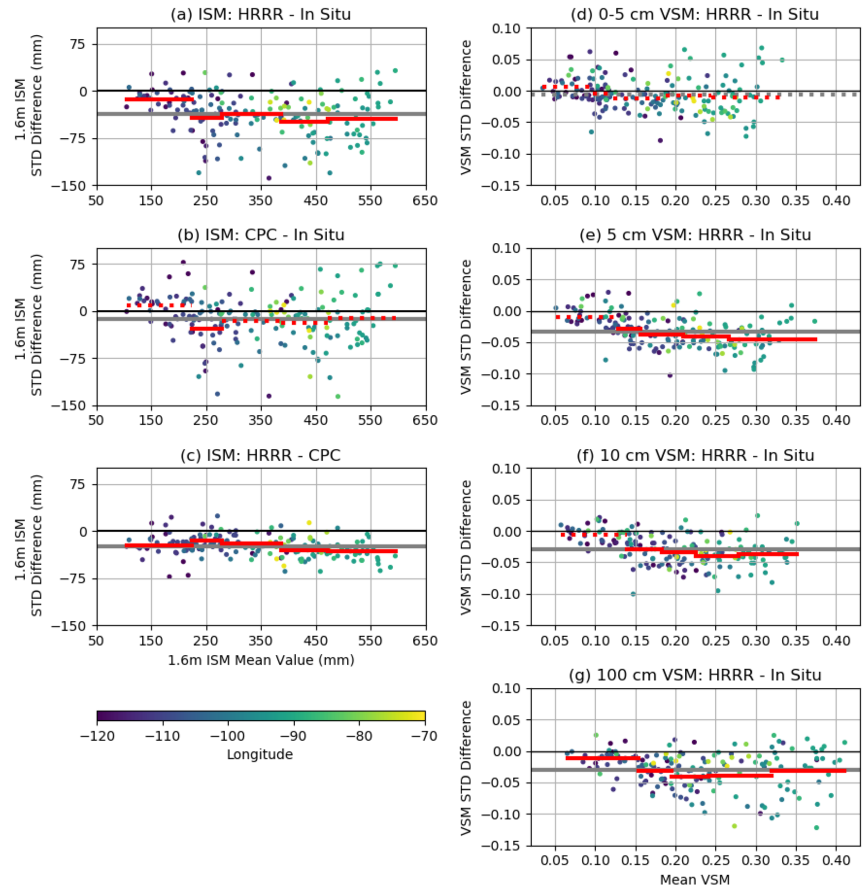
4.2 VSM Mean Comparisons at Varying Depths

To better understand the differences in ISM data, the VSM data at different depths are compared in the HRRR and in situ data (Figure 2d-g). Note, that the VSM data are also separated into quintiles from driest to wettest in a similar way to the ISM data, and this procedure is done for each vertical level. Therefore, locations may fall into different quintiles for different vertical levels. The same trend of the HRRR being wetter in $L_{00\text{-}20}$ (i.e., the driest regions) and dryer in $L_{80\text{-}100}$ (i.e., the wettest regions) is present at all depths. There are generally smaller differences in the middle quintiles. However, at the lowest depths (100 cm below ground; Fig 2g), the differences between the HRRR and in situ VSMs have larger magnitudes, especially for the driest and wettest regimes ($L_{00\text{-}20}$ difference of +0.08 and $L_{80\text{-}100}$difference of -0.18). Near the surface (Figure 2d), the driest 40% of the regions have relatively small differences ($\pm$ 0.02), although significant dry biases are present for the wetter regions (-0.11). The VSM data demonstrate that the deeper soil layers (i.e., 100 cm below ground) are the primary drivers of the ISM trends. Shallower soil layers have similar trends yet smaller differences as compared to those at deeper levels and have a smaller contribution to the ISM differences. It is important to note that the magnitude of both ISM and VSM mean differences vary throughout the year, and the evolution of these differences as a function of month and season are provided in the supporting information document.

5 Soil Moisture Variance

5.1 ISM Standard Deviation Comparisons

Understanding the temporal variability (e.g., standard deviations) in soil moisture allows for an assessment on whether models are accurately capturing the processes that result in soil moisture changes. Maps of the ISM standard deviations in the HRRR and CPC ISM data show similar patterns, with the highest variance occurring along the Pacific Northwest coast (Figure 1d-e). The lowest ISM standard deviations occur along the Intermountain West and High Plains regions. In general, the HRRR ISMs have lower variance than the CPC data across the US except for in the parts of the Rocky Mountains and in the Great Lakes and Northeast regions (Figure 1f). The HRRR produces larger spatial variability within the mountain and valley regions across the western US that cannot be resolved in the CPC data.



The in situ observations generally have the highest ISM variance for most quintiles followed by CPC and HRRR, which has the lowest variances of the three datasets regardless of the wetness regime (Figure 3a-c). The in situ and CPC quintile mean ISM standard deviations are not significantly different for all quintiles except $L_{20\text{-}40}$, while the HRRR quintile mean differences are significant for all quintiles when compared to both other datasets. Dirmeyer et al. (2016) showed that spatial scaling differences do not have a large impact (˜10%) on in situ observation standard deviations via conducting tests where many stations that are separated by several km to up to 100 km are averaged together. Therefore, the differences between the in situ and NOAA modeled standard deviations are likely due to other factors outside of dataset spatial scale differences (e.g., model processes representation or model input data). Furthermore, the large variability or scatter in the ISM standard deviation differences between the in situ data and both models (Figure 3a-b) suggests that the cause of these differences depends on the specific location. The differences between HRRR and CPC ISM standard deviations, however, demonstrate a more systematic bias between these two modeling frameworks (Figure 3c).

Figure 3. Same as Figure 2, except for comparisons of ISM standard deviations (a-c) and VSM standard deviations (d-g).

5.2 VSM Standard Deviation Comparisons at Varying Depths

VSM standard deviations at four different depths are compared between HRRR and in situ data (Figure 3d-g) to determine whether a certain depth is driving the differences in the ISM standard deviations (Figure 3a). Regardless of the soil moisture regime, the HRRR surface VSM standard deviations compare well to the near-surface in situ observations. When averaging over all locations, the mean near-surface percentage differences in HRRR soil moisture from the in situ value is only +4.2% (Figure 3d). This is likely related to improvements made in the HRRR's RUC LSM and the moderately coupled land data assimilation system that have been applied (Benjamin et al., 2022). However, at depths of 5 cm below ground and deeper (Figure 3e-g), most quintiles have statistically significant differences in the standard deviations. The mean percentage differences over all locations are -38.7%, -36.4% and -45.2% for the 5 cm, 10 cm and 100 cm levels, respectively, with the HRRR always having lower standard deviations than the in situ datasets for every quintile. These differences in VSM standard deviations are larger for wetter soil moisture regimes ($L_{20\text{-}100}$). While the differences for the different soil moisture regimes are generally similar for the 5 cm, 10 cm and 100 cm depths, there are more extreme differences for specific locations at the 100 cm level. To summarize, the lower standard deviations in the HRRR ISM are being driven by the lower VSM standard deviations occurring below the surface level. It is important to note that both ISM and VSM standard deviation differences vary throughout the year, and the evolution of these differences as a function of month and season are provided in the supporting information document.

6 Conclusions and Future Work

A comparison of 1.6 m ISM between three different NOAA soil moisture products is conducted. This analysis uniquely includes the HRRR model with its RUC LSM, the CPC leaky-bucket model and in situ observations from two national networks. These soil moisture estimates are used in many operational and research applications, including atmospheric forecasting, drought monitoring, and assessing flood and fire risks. Therefore, quantifying differences in these NOAA models to observational networks across CONUS is critical.

Several conclusions are drawn from these comparisons.

1) The HRRR and CPC ISMs are both larger (i.e., wetter) in the driest regions and smaller (i.e., drier) in the wettest regions as compared to in situ observations.

2) These differences in the HRRR and in situ ISM amounts are largely caused by deep soil levels (~100 cm below ground). Shallower layers have similar trends to the deeper layers but have smaller differences, and thus a weaker contribution to the ISM differences.

3) The in situ observations have the largest ISM standard deviations, followed by the CPC leaky-bucket model and the HRRR model.

4) The HRRR soil moisture standard deviations compare well with the in situ standard deviations near the surface, but large differences are present at 5 cm below the surface and deeper.

The soil moisture differences presented in this study can be caused by a variety of reasons. In terms of modeled soil moisture, biases in the input datasets (i.e., precipitation or radiation), whether they come from a coupled atmospheric model in the case of HRRR or external sources in the case of CPC, have been shown to lead to biases in land surface model calculations (e.g., Mitchell et al., 2004; Min et al., 2021). Choices in the land surface model structure, such as the number and thickness of soil layers, the representation of soil and vegetation, and other model parameters, can also lead to biases in soil moisture prediction (e.g., Mitchel et al., 2004; Xia et al., 2014; 2015b). Min et al. (2021) found that snowmelt, freezing/thawing, and/or biases in precipitation and evapotranspiration led to differences in HRRR soil moisture as compared to in situ observations in New York and that the most relevant processes causing these differences varied throughout the year. The results in this study demonstrate consistent, region-dependent biases in NOAA modeled soil moisture as compared to in situ observations across CONUS, and future research should focus on understanding the model processes that are causing these biases.

Our results also provide important context to the current users of these models and observations. For example, HRRR's land data assimilation system has recently undergone changes that primarily impact the near-surface soil state (Benjamin et al., 2022). The comparisons presented in this study do show better performance of HRRR soil moisture near the surface and thus may provide a first step towards understanding the impact of these model changes. Furthermore, these results can assist with the continued development and refinement of soil moisture models and products. The analyses presented here are currently being utilized for preparing training and validation data for a machine learning algorithm that uses data from the Advanced Baseline Imager on-board NOAA's Geostationary Operational Environmental Satellite to estimate the soil moisture state at very high resolution (i.e., on the order of ˜1 km). With a recent focus on land-atmosphere coupling and a continued shift towards higher-resolution models, such a product could be used as a supplementary input for strongly coupled land atmosphere data assimilation in the next generation of atmospheric models.

Acknowledgments

This work was supported by the NOAA FY21 High Performance Computing and Communications Program's Information Technology Incubator. We would also like to acknowledge helpful feedback on and interest in this work from Liaofan Lin, Tanya Smirnova, Stan Benjamin, Curtis Alexander and Eric James.

Open Research

Several in situ and model datasets are used in this study. The USCRN data (Palecki et al., 2013) and the SCAN data (SCAN, 2016) will be archived upon publication. This archiving process is underway and will be with the Mountain Scholar repository through Colorado State University. We have uploaded a copy of these data as Supporting Information for the review process. The CPC data was accessed via https://ftp.cpc.ncep.noaa.gov/wd51yf/us/w_daily/ through the U.S. Data download link on the NOAA CPC product webpage (https://www.cpc.ncep.noaa.gov/products/Soilmst_Monitoring/). The HRRR operational model data (HRRRv3) was stored and accessed via the NOAA Hera supercomputer and is publicly archived at either https://registry.opendata.aws/noaa-hrrr-pds/ or https://console.cloud.google.com/marketplace/product/noaa-public/hrrr. The analysis code used to generate the analyses and figures in this manuscript are available at https://github.com/pjmarinescu/CIRA_Soil_Moisture and will also be archived in the same Mountain Scholar repository as the data upon publication.

References

Arevalo, J., Welty, J., Fan, Y., & Zeng, X. (2021). Implementation of snowpack treatment in the cpc water balance model and its impact on drought assessment. *Journal of Hydrometeorology* , *22* (5), 1235–1247. https://doi.org/10.1175/JHM-D-20-0201.1

Atiah, W. A., Amekudzi, L. K., Akum, R. A., Quansah, E., Antwi-Agyei, P., & Danuor, S. K. (2022). Climate variability and impacts on maize (Zea mays) yield in Ghana, West Africa. *Quarterly Journal of the Royal Meteorological Society* , *148* (742), 185–198. https://doi.org/10.1002/qj.4199

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., et al. (2013). U.S. climate reference network soil moisture and temperature observations. *Journal of Hydrometeorology* , *14* (3), 977–988. https://doi.org/10.1175/JHM-D-12-0146.1

Benjamin, S. G., Smirnova, T. G., James, E. P., Lin, L.-F., Hu, M., Turner, D. D., & He, S. (2022). Land-snow data assimilation including a moderately coupled initialization method applied to NWP. *Journal of Hydrometeorology* . https://doi.org/10.1175/jhm-d-21-0198.1

Carrera, M. L., Bilodeau, B., Bélair, S., Abrahamowicz, M., Russell, A., & Wang, X. (2019). Assimilation of passive L-band microwave brightness temperatures in the Canadian Land data assimilation system: Impacts on short-range warm season numerical weather prediction. *Journal of Hydrometeorology* , *20* (6), 1053–1079. https://doi.org/10.1175/JHM-D-18-0133.1

Delworth, T. L., & Manabe, S. (1988). The Influence of Potential Evaporation on the Variabilities of Simulated Soil Wetness and Climate.*Journal of Climate* , *1* (5), 523–547. https://doi.org/10.1175/1520-0442(1988)001<0523:TIOPEO>2.0.CO;2

Dirmeyer, P. A., Wu, J., Norton, H. E., Dorigo, W. A., Quiring, S. M., Ford, T. W., et al. (2016). Confronting weather and climate models with observational data from soil moisture networks over the United

States.*Journal of Hydrometeorology* , *17* (4), 1049–1067. https://doi.org/10.1175/JHM-D-15-0196.1

Dowell, D. C., Alexander, C. R., James, E. P., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., et al. (2022). The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part 1: Motivation and System Description. *Wea. Forecasting* . https://doi.org/10.1175/WAF-D-21-0151.1

Ek, M. B., & Holtslag, A. A. M. (2004). Influence of Soil Moisture on Boundary Layer Cloud Development. *Journal of Hydrometeorology* ,*5* (1), 86–99. https://doi.org/10.1175/1525-7541(2004)005<0086:IOSMOB>2.0.CO;2

Fan, Y., & van den Dool, H. (2004). Climate Prediction Center global monthly soil moisture data set at 0.5° resolution for 1948 to present.*Journal of Geophysical Research D: Atmospheres* , *109* (10), 1–8. https://doi.org/10.1029/2003JD004345

Ford, T. W., & Quiring, S. M. (2019). Comparison of Contemporary In Situ, Model, and Satellite Remote Sensing Soil Moisture With a Focus on Drought Monitoring. *Water Resources Research* , 1565–1582. https://doi.org/10.1029/2018WR024039

Guttman, N. B., & Quayle, R. G. (1996). A historical perspective of U.S. climate divisions. *Bulletin of the American Meteorological Society* , *77* (2), 293–303. https://doi.org/10.1175/1520-0477(1996)077<0293:AHPOUC>2.0.CO;2

Huang, J., Van Den Dool, H. M., & Georgakakos, K. P. (1996). Analysis of model-calculated soil moisture over the United States (1931-1993) and applications to long-range temperature forecasts. *Journal of Climate* . https://doi.org/10.1175/1520-0442(1996)009<1350:AOMCSM>2.0.CO;2

James, E. P., Alexander, C. R., Dowell, D. C., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., et al. (2022). The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part 2: Forecast Performance. *Wea. Forecasting.*https://doi.org/10.1175/WAF-D-21-0130.1

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., et al. (2004). Regions of Strong Coupling Between Soil Moisture and Precipitation. *Science* , 305(5687), 1138–1140. https://doi.org/10.1126/science.1100217

Lin, L. F., & Pu, Z. (2020). Improving near-surface short-range weather forecasts using strongly coupled land-atmosphere data assimilation with gsi-enkf. *Monthly Weather Review* , 148(7), 2863–2888. https://doi.org/10.1175/MWR-D-19-0370.1

Liu, J., Zhan, X., Hain, C., Yin, J., Fang, L., Li, Z., & Zhao, L. (2016). NOAA Soil Moisture Operational Product System (SMOPS) and its validations. In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3477–3480). IEEE. https://doi.org/10.1109/IGARSS.2016.7729899

Madadgar, S., AghaKouchak, A., Farahmand, A., & Davis, S. J. (2017). Probabilistic estimates of drought impacts on agricultural production.*Geophysical Research Letters* , *44* (15), 7799–7807. https://doi.org/10.1002/2017GL073606

Min, L., Fitzjarrald, D. R., Du, Y., Rose, B. E. J., Hong, J., & Min, Q. (2021). Exploring Sources of Surface Bias in HRRR Using New York State Mesonet. *Journal of Geophysical Research: Atmospheres* ,*126* (20), 1–18. https://doi.org/10.1029/2021JD034989

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., et al. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system.*Journal of Geophysical Research: Atmospheres* , 109(7), 1–32. https://doi.org/10.1029/2003jd003823

Muñoz-Sabater, J., Lawrence, H., Albergel, C., Rosnay, P., Isaksen, L., Mecklenburg, S., et al. (2019). Assimilation of SMOS brightness temperatures in the ECMWF Integrated Forecasting System. *Quarterly Journal* *of the Royal Meteorological Society* , *145* (723), 2524–2548. https://doi.org/10.1002/qj.3577

Palecki, M. A., Lawrimore, J.H., Leeper, R. D., Bell, J. E., Embler, S., Casey, N. (2013). U.S. Climate Reference Network Products, Daily. [Dataset]. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V5H13007. https://www.ncei.noaa.gov/pub/data/uscrn/products/daily01/. Accessed 11 Nov. 2021.

Pan, M., Cai, X., Chaney, N. W., Entekhabi, D., & Wood, E. F. (2016). An initial assessment of SMAP soil moisture retrievals using high-resolution model simulations and in situ observations. *Geophysical Research Letters* , 43(18), 9662–9668. https://doi.org/10.1002/2016GL069964

Quiring, S. M., Ford, T. W., Wang, J. K., Khong, A., Harris, E., Lindgren, T., et al. (2016). The North American Soil Moisture Database: Development and Applications. *Bulletin of the American Meteorological Society* , 97(8), 1441–1459. https://doi.org/10.1175/BAMS-D-13-00263.1

Rigden, A. J., Powell, R. S., Trevino, A., McColl, K. A., & Huybers, P. (2020). Microwave Retrievals of Soil Moisture Improve Grassland Wildfire Predictions. *Geophysical Research Letters* , *47* (23), 1–8. https://doi.org/10.1029/2020GL091410

Robock, A., Luo, L., Wood, E. F., Wen, F., Mitchell, K. E., Houser, P. R., et al. (2003). Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season. *Journal of Geophysical Research: Atmospheres* ,*108* (22). https://doi.org/10.1029/2002jd003245

Robock, A., Vinnikov, K. Y., Schlosser, C. A., Speranskaya, N. A., & Xue, Y. (1995). Use of Midlatitude Soil Moisture and Meteorological Observations to Validate Soil Moisture Simulations with Biosphere and Bucket Models. *Journal of Climate* , *8* (1), 15–35. https://doi.org/10.1175/1520-0442(1995)008<0015:UOMSMA>2.0.CO;2

Soil Climate Analysis Network (SCAN) (2016). SCAN Daily Historic Provisional Data. [Dataset]. USDA National Resources Conservation Service National Water and Climate Center. www.wcc.nrcs.usda.gov/scan. Access date 16 Feb. 2022.

Schaefer, G. L., Cosh, M. H., & Jackson, T. J. (2007). The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *Journal of Atmospheric and Oceanic Technology* , *24* (12), 2073–2077. https://doi.org/10.1175/2007JTECHA930.1

Shellito, P. J., Small, E. E., Colliander, A., Bindlish, R., Cosh, M. H., Berg, A. A., et al. (2016). SMAP soil moisture drying more rapid than observed in situ following rainfall events. *Geophysical Research Letters* , 43(15), 8068–8075. https://doi.org/10.1002/2016GL069946

Smirnova, T. G., Brown, J. M., & Benjamin, S. G. (1997). Performance of different soil model configurations in simulating ground surface temperature and surface fluxes. *Monthly Weather Review* ,*125* (8), 1870–1884. https://doi.org/10.1175/1520-0493(1997)125<1870:PODSMC>2.0.CO;2

Smirnova, T. G., Brown, J. M., Benjamin, S. G., & Kim, D. (2000). Parameterization of cold-season processes in the MAPS land-surface scheme. *Journal of Geophysical Research: Atmospheres* , 105(D3), 4077–4086. https://doi.org/10.1029/1999JD901047

Smirnova, T. G., Brown, J. M., Benjamin, S. G., & Kenyon, J. S. (2016). Modifications to the Rapid Update Cycle Land Surface Model (RUC LSM) Available in the Weather Research and Forecasting (WRF) Model. *Monthly Weather Review* , 144(5), 1851–1865. https://doi.org/10.1175/MWR-D-15-0198.1

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., et al. (2002). THE DROUGHT MONITOR. *Bulletin of the American Meteorological Society* , *83* (8), 1181–1190. https://doi.org/10.1175/1520-0477-83.8.1181

Taylor, C. M., Gounou, A., Guichard, F., Harris, P. P., Ellis, R. J., Couvreux, F., & De Kauwe, M. (2011). Frequency of sahelian storm initiation enhanced over mesoscale soil-moisture patterns. Nature Geoscience, 4(7), 430–433. https://doi.org/10.1038/ngeo1173

van den Dool, H., Huang, J., & Fan, Y. (2003). Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981-2001. *Journal of Geophysical Research: Atmospheres* ,*108* (16), 1–16. https://doi.org/10.1029/2002jd003114

Vinnikov, K. Y., & Yeserkepova, I. B. (1991). Soil Moisture: Empirical Data and Model Results. *Journal of Climate* , *4* (1), 66–79. https://doi.org/10.1175/1520-0442(1991)004<0066:SMEDAM>2.0.CO;2

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. Journal of Geophysical Research Atmospheres, 117(3). https://doi.org/10.1029/2011JD016048

Xia, Y., Sheffield, J., Ek, M. B., Dong, J., Chaney, N., Wei, H., et al. (2014). Evaluation of multi-model simulated soil moisture in NLDAS-2. Journal of Hydrology, 512, 107–125. https://doi.org/10.1016/j.jhydrol.2014.02.027

Xia, Y., Ek, M. B., Wu, Y., Ford, T., & Quiring, S. M. (2015a). Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: Comparison and analysis. *Journal of Hydrometeorology* ,*16* (5), 1962–1980. https://doi.org/10.1175/JHM-D-14-0096.1

Xia, Y., Ek, M. B., Wu, Y., Ford, T., & Quiring, S. M. (2015b). Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part II: Impact of soil texture classification and vegetation type mismatches. *Journal of Hydrometeorology* , 16(5), 1981–2000. https://doi.org/10.1175/JHM-D-14-0097.1

Yao, Y., Ciais, P., Viovy, N., Li, W., Cresto-Aleina, F., Yang, H., et al. (2021). A Data-Driven Global Soil Heterotrophic Respiration Dataset and the Drivers of Its Inter-Annual Variability. *Global Biogeochemical Cycles* , *35* (8), 1–23. https://doi.org/10.1029/2020GB006918

Peter J. Marinescu[1], Kyle Hilburn[1], Daniel Abdi[2], Isidora Jankov[3]

[1] Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado.

[2] Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado.

[3] Global Systems Laboratory, NOAA/OAR, Boulder, Colorado.

Corresponding author: Peter J. Marinescu (peter.marinescu@colostate.edu)

Key Points:

- NOAA modeled soil moisture is drier in wet regions and wetter in dry regions as compared to in situ observations.
- NOAA modeled soil moisture has lower variance than in situ observations.
- The differences between NOAA modeled soil moisture and in situ data are larger at deeper soil depths as compared to near the surface.

Abstract

Three estimates of soil moisture from National Oceanic and Atmospheric Administration (NOAA) programs are compared. The estimates are from a high-resolution atmospheric model with a land surface model, a hydrologic model and in situ observations. Both models demonstrate wetter soil moisture in dry regions and drier soil moistures in wet regions, as compared to the in situ observations. These soil moisture differences occur at most soil depths but are larger at the deeper depths below the surface (100 cm). In terms of soil moisture variance, both models generally have lower standard deviations as compared to the in situ observations, except for near the surface where the in situ and high-resolution, land surface model compare well. These NOAA soil moisture estimates are used for a variety of forecasting and societal applications, and understanding their differences provides important context for their applications and can lead to model improvements.

Plain Language Summary

Soil moisture is an essential variable coupling the land surface to the atmosphere. Accurate estimates of soil moisture are important for predicting where clouds will form, assessing drought and fire weather risks, and assisting with decisions for agricultural production. There are multiple estimates of soil moisture available, and in this study, we compare three different types of soil moisture estimates from the National Oceanic and Atmospheric Administration (NOAA), including two types of soil moisture models and direct observations, as well as direct observations from the United States Department of Agriculture. Both models have soil moisture values that are too wet in dry regions of the United States and too dry in wet regions. The modeled soil moisture also does not change as rapidly from day-to-day as they do in the direct observations of soil moisture. Further analysis at different soil depths shows which depths have the

largest differences between these three soil moisture estimates. Understanding these NOAA soil moisture estimate differences is important for forecasters and decision makers and can be useful for the further development of atmospheric and land surface models.

1 Introduction

Knowledge of soil moisture is essential for many Earth system applications, such as forecasting cloud formation (e.g., Ek and Holtslag, 2004), monitoring drought, flood and fire risks (e.g., Svoboda et al., 2002; Rigden et al., 2020), and providing instrumental information for agricultural production (e.g., Madadgar et al., 2017). As such, several advancements in the estimation and utilization of soil moisture have recently transpired. For example, in efforts to improve the accuracy of numerical weather prediction (NWP) and climate models, model developers have focused on increasing the coupling between the land surface and atmosphere components of their data assimilation systems to eliminate persistent atmospheric prediction biases (e.g., Benjamin et al., 2022). Furthermore, NWP models are also beginning to explore the direct assimilation of new soil moisture observations (e.g., Carrera et al., 2019; Muñoz-Sabater et al., 2019; Lin and Pu, 2020).

For these reasons, it is critical to understand the differences in the available soil moisture estimates from models or observations that are used in science, forecasting or agricultural applications. Different estimates of soil moisture exist across the continental United States (CONUS), each with its own benefits and shortfalls. For example, in situ observations are often considered to be the most accurate and are therefore used as a benchmark. However, they have limited spatial coverage (e.g., Quiring et al., 2016). Other products, such as those from low Earth orbiting satellites, have lower temporal and spatial resolution (e.g., Liu et al., 2016). Soil moisture estimates from models depend on many assumptions and reflect the influence of observation-based data to different degrees (e.g., Smirnova et al., 1997; Huang et al., 1996; Mitchell et al., 2004). Several recent studies have made significant progress in comparing many of the available soil moisture estimates. Some studies have compared soil moisture temporal variability and memory between many large-scale land surface models (LSMs) and in situ soil monitoring networks across the CONUS, noting certain biases and uncertainties in the various estimates (e.g., Robock et al., 2003; Xia et al., 2014; 2015a; Dirmeyer et al., 2016). Recent studies have extended soil moisture comparisons to include soil moisture retrievals from new satellite platforms (e.g., Shellito et al., 2016; Pan et al., 2016; Ford and Quiring, 2019).

While various LSMs have been compared to in situ observations for several decades, these prior studies have primarily focused on models with horizontal resolutions on the scales of   degree or larger (i.e., the North American Land Data Assimilation System models; Mitchell et al., 2004; Xia et al., 2012). Given the local, mesoscale variability of soil moisture process and the subsequent impacts of soil moisture on atmospheric prediction (e.g., Koster et al., 2004; Taylor et al., 2011), high-resolution models should also be tested. Recently, Min et al. (2021)

compared near-surface atmospheric and soil variables from the High-Resolution Rapid Refresh (HRRR) model (Dowell et al., 2022; James et al., 2022), which has horizontal grid spacing of 3 km, to observations from the New York State Mesonet. They found that soil moisture was underestimated in HRRR, which contributed to warm and dry biases in atmospheric forecasts. However, how does soil moisture from the HRRR model (NOAA's current operational, convection-allowing model) compare to soil moisture estimates across all of CONUS?

In this study, we focus on a comparison of soil moisture estimates from NOAA. In particular, we uniquely include soil moisture estimates across CONUS from the NOAA operational HRRR model, which utilizes the RUC land surface model (RUC LSM, Smirnova et al., 1997). We also include the NOAA Climate Prediction Center (CPC) leaky-bucket hydrological model (Huang et al., 1996; van den Dool et al., 2003) and in situ observations from two nationwide networks: the NOAA/NCEI United States (US) Climate Reference Network (USCRN; Bell et al., 2013) and the US Department of Agriculture Soil Climate Analysis Network (SCAN; Schaefer et al., 2007). This work provides an assessment of the similarities and differences of soil moisture amounts and variance across three different products, which are all used in various operational and research applications.

2 Soil Moisture Data

2.1 In Situ Observations

The study uses in situ soil moisture observations from two nationwide networks. USCRN provides climate monitoring measurements of atmospheric and soil properties. To increase the coverage of the in situ observations, SCAN is also included. SCAN uses similar sensors to USCRN (i.e., Hydra Probe sensors) and typically have volumetric soil moisture (VSM; ${m_{water}}^3$ / ${m_{soil}}^3$) measurements at the same soil depths (~5, ~10, ~20, ~50, and ~100 cm) as USCRN. Only data from these five levels are used for consistency. Dirmeyer et al. (2016) also found that these two networks have similar error variances. The observations represent point measurements of the soil moisture at specific sites across the US. Daily data are used in this study and represent the average VSM of the entire 24-hour period based on local standard time.

2.2 HRRR Model

The HRRR model is NOAA's operational, convection-allowing model, which has 3 km horizontal grid spacing and covers CONUS with a one-hour temporal refresh rate (Dowell et al., 2022; James et al., 2022). In this study, we use HRRRv3, which was operational between 12 July 12 2018 and 2 December 2020. The HRRR model utilizes a one-dimensional land surface model (RUC LSM; Smirnova et al., 1997), which predicts heat and moisture transfer vertically throughout the soil column. The RUC LSM has undergone several enhancements over the years, including increasing its resolution and incorporating new features, such as snow and ice models (Smirnova et al., 2000; 2016). The current version predicts VSM at nine vertical levels (0, 1, 4, 10, 30, 60, 100, 160 and 300 cm) and utilizes cycling of soil conditions over several years to better capture the soil moisture state. The HRRR utilizes moderately coupled land data assimilation, meaning that near-surface atmospheric data assimilation increments are used to adjust the soil analysis (e.g., Benjamin et al., 2022). Given the recent and continued development of the HRRR data assimilation system and land surface model, it is critical for assessments of HRRR's soil moisture to other estimates. This study provides a benchmark for HRRR soil moisture estimates in support of the continued development of NOAA's land surface prediction capabilities in the Unified Forecast System. The focus of this study is on the analyzed soil moisture field, rather than forecast fields, and thus our results are most directly relevant to the LSM and data assimilation system development.

2.3 CPC Leaky-Bucket Model

The CPC soil moisture product utilizes a leaky-bucket model that solves the time tendency equation in soil moisture over a region from several inputs: precipitation minus evapotranspiration, net streamflow divergence and net groundwater loss (Huang et al., 1996; van den Dool et al., 2003). These inputs to the time tendency equation for soil moisture have been improved over the years with new observations and parameterizations (Fan and van den Dool, 2004; Arevalo et al., 2021). The CPC model provides 1.6 m deep integrated soil moisture (ISM, mm), and these estimates are provided daily for each of the NOAA climate divisions across the United States (Guttman and Quayle, 1996). There are typically about 7-10 climate divisions per state, although there are fewer for states with smaller geographical areas, like those in the Northeast United States. The CPC soil moisture data are used as an input to the United States Drought Monitor (Svoboda et al., 2002) and continues to be used as a reference data set in various soil moisture application studies, from assessing soil moisture impacts on carbon fluxes (e.g., Yao et al., 2021) to understanding climate impacts on agricultural production (e.g., Atiah et al., 2022).

3 Methods

3.1 ISM and VSM Comparisons

Since the CPC product only provides 1.6 m ISM, VSM values from the in situ and HRRR data are vertically integrated in order to compare 1.6 m ISM in all three datasets. The VSM values are assumed to represent the mean value over a depth between the midpoints of the specified levels, as has been done in other studies (e.g., Dirmeyer et al., 2016; Ford and Quiring, 2019). For example, the 10 cm VSM observation in the in situ data is assumed to represent the average VSM for the layer between 7.5 cm (i.e., the midpoint between 5 and 10 cm) and 15 cm (i.e., the midpoint between 10 and 20 cm). The 100 cm VSM in the in situ data is also assumed to be constant to the depth of 160 cm. The HRRR ISM calculations are better constrained than the in situ ISM calculations, since the HRRR VSM data span 3.0 m below ground using 9 levels. VSM values are also compared between the HRRR and in situ data to glean whether certain vertical levels are driving differences between these two datasets.

An understanding of soil moisture differences at varying depths is also critical since soil moisture's role in Earth system processes is depth dependent. In terms of spatial comparisons, the HRRR and CPC data are linearly interpolated to the locations of the in situ stations. The analysis is completed over a ~2.4 year period from 12 July 2018 through 2 December 2020, which is the timeframe that HRRRv3 was operational. By confining the analyses to this time frame, uncertainties associated with model version changes are avoided.

3.2 Quality Control

In situ data provide the most direct physical estimate of soil moisture, but it is important to ensure that the in situ data are of the highest quality. As such, a variety of quality control procedures are undertaken. First, in situ data are only included if they have VSM values available at all five vertical levels (~5cm, ~10cm, ~20cm, ~50cm, and ~100cm), since missing data could lead to larger uncertainties in the ISM calculation. Second, in situ stations directly along the coast are removed due to unphysical spatial interpolations from the CPC and HRRR data. From the remaining in situ data, we estimate the ratio of error variance to ISM variance using the method defined in Robock et al. (1995) and used in more recent soil moisture comparisons studies (e.g., Dirmeyer et al., 2016). In essence, soil moisture can be well approximated by a red-noise process (i.e., first-order Markov process; e.g., Delworth and Manabe, 1988; Vinnikov and Yeserkepova, 1991) with the natural logarithm of the soil moisture autocorrelation (r) decreasing linearly with increased lag times ( ). For this study, autocorrelations are computed for of 1-30 days for each station's ISM daily anomalies. A linear fit is applied to the ln(r) versus data and is extrapolated to = 0. Deviations from 1 at = 0 can be used to solve for to the ratio of error variance to ISM variance (e.g., Robock et al., 1995). For this study, stations where this ratio is greater than 0.08 are removed. This error ratio threshold is in-line with estimates of the mean error ratio of the USCRN and SCAN networks (Dirmeyer et al., 2016). While this error variance ratio threshold results in the removal of 63 (~27%) of the 235 in situ stations, it provides more confidence that only the highest quality in situ observations are being used in the analyses. Even with the significant reduction of the in situ data, the stations span the entirety of the CONUS (Figure 1, dots). Different thresholds, autocorrelation lengths and dataset lengths were tested and did not qualitatively impact the results.

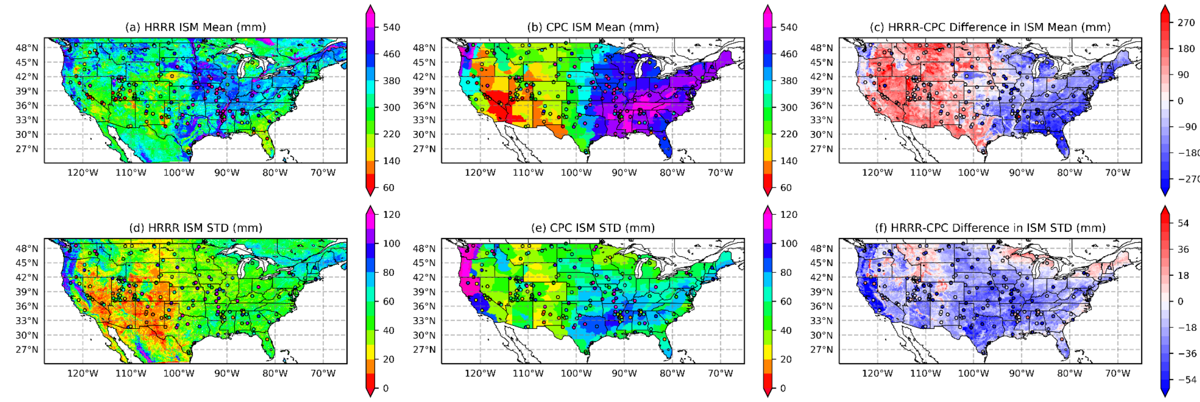


Fig 1: 1.6 m ISM values and standard deviations temporally averaged over the study period. (a) represents the HRRR model, (b) represents the CPC model, and (c) represents their difference. (d-f) are the same as (a-c), except for the ISM standard deviations. Filled circles represent locations of the 172 in situ observations from the USCRN and SCAN networks that pass quality control checks, as described in the text. In (c) and (f), the filled circles represent the HRRR minus in situ differences, while the filled circles represent the in situ station ISM mean and standard deviations, respectively, for the other panels.

3.3 Quintile Analysis

To determine whether differences in soil moisture estimates vary in different soil moisture regimes (i.e., wetter versus drier conditions), we composite the comparisons over locations with similar soil moisture amounts. The ISM or VSM values for each dataset are averaged temporally for each location and are then averaged among the available datasets (i.e., all three datasets for ISM and the in situ and HRRR datasets for VSM). Using this mean, the locations are separated into five quintiles. For example, locations with a mean ISM estimate that is greater than or equal to the 0^{th} percentile ISM and less than the 20^{th} percentile ISM are placed in the lowest quintile of soil moisture amounts (i.e., driest locations). These locations are termed $L_{00\text{-}20}$. Similarly, locations that fall within the 20^{th}-40^{th}, 40^{th}-60^{th}, 60^{th}-80^{th} and 80^{th}-100^{th} percentiles are termed $L_{20\text{-}40}$, $L_{40\text{-}60}$, $L_{60\text{-}80}$ and $L_{80\text{-}100}$, respectively and represent dry to wet soil moisture regimes. There are either 34 or 35 locations that are included in each of these quintiles.

4 Soil Moisture Amounts

4.1 ISM Mean Comparisons

Stark, regionally-dependent differences are apparent in the ISM estimates (Figure 1c,f). The HRRR ISM values (Figure 1a) have a more muted range than the CPC ISM values (Figure 1b). The HRRR ISMs are larger (i.e., wetter) than the CPC ISMs across the drier regions of CONUS (most of the western CONUS; Figure 1c) and are smaller (i.e., drier) than the CPC ISMs in the wetter regions of CONUS (eastern half of CONUS and the coastal regions of the Pacific Northwest). This regional dependence is also clearly associated with varying soil moisture amounts.

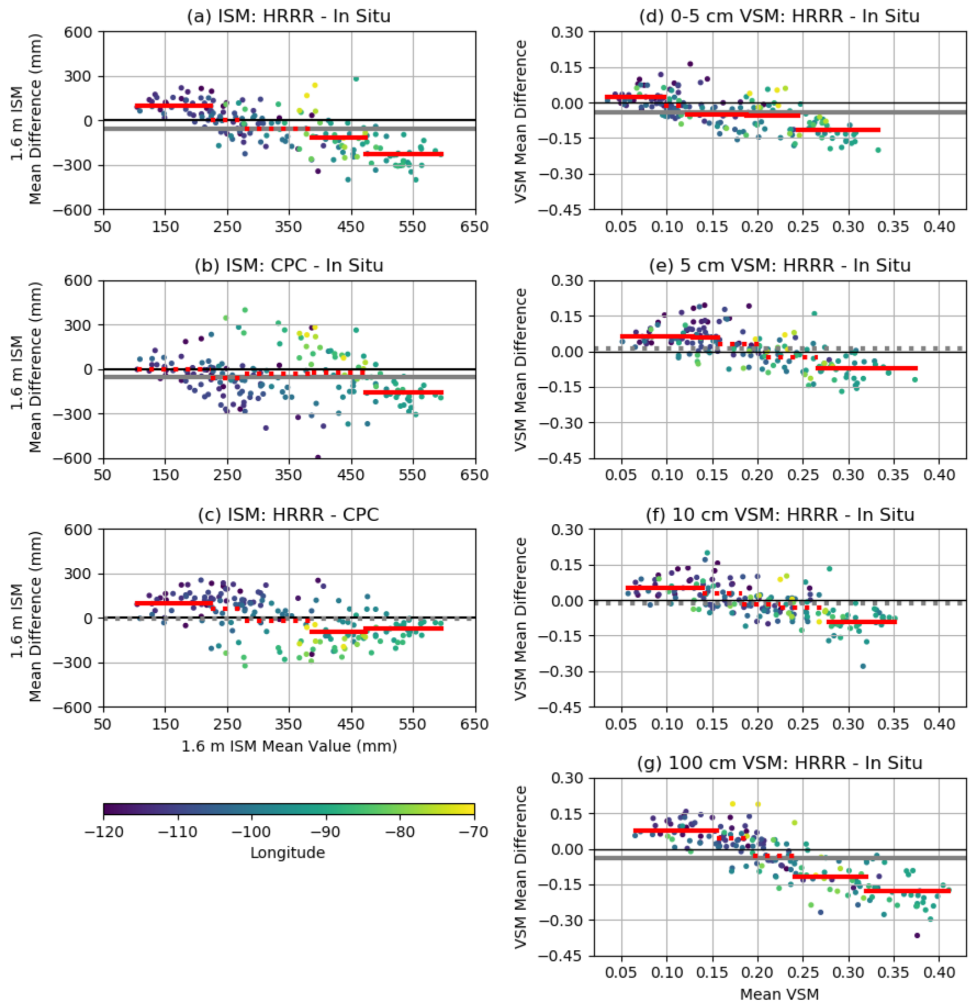


Figure 2: Comparisons of the 1.6 m ISM means between (a) HRRR and in situ values, (b) CPC and in situ values, and (c) HRRR and CPC values. Note, (c) only shows results from the same locations where in situ data are available for consistency. (d-g) shows comparisons of VSM between the HRRR and in situ data for 4 different level combinations: (d) surface in HRRR to 5 cm in the in situ data, (e) 5 cm for both, (f) 10 cm for both, and (g) 100 cm for both. The gray, horizontal line represents the mean difference for the entire dataset, while red horizontal lines represent the mean difference over the five quintiles. Horizontal gray and red lines are solid if their associated data differences pass statistical significance using a paired Student's t-test at the 99.9th percentile.

Figure 2 includes the in situ ISM estimates within the comparisons in order to better quantify the differences between the three datasets. When compared to both in situ ISM (Figure 2a) and CPC ISM (Figure 2c), HRRR is wetter in $L_{00\text{-}20}$ (+101% and +97%, respectively, when taking the mean percentage difference for all locations in $L_{00\text{-}20}$) and drier in $L_{80\text{-}100}$ (-34% and -13%, respectively). Even though there are fewer moisture quintiles with statistically significant differences between the CPC and in situ ISMs, the longitude locations show large regional differences between the CPC and in situ datasets, even within a given quintile range. For example, $L_{40\text{-}60}$ and $L_{60\text{-}80}$ show minimal differences between the mean CPC and in situ ISMs (Figure 2b, red lines), but at eastern locations the CPC ISMs are generally larger (Figure 2b., yellow and green dots) and at western locations the CPC ISMs are generally smaller (Figure 2b, blue dots). As was also shown in Figure 1, Figure 2 demonstrates clear regional relationships to these ISM differences between the datasets.

Despite larger, systematic, moisture-regime-dependent differences between HRRR and in situ observations (Figure 2a), HRRR aligns more closely to the in situ observation than does the CPC for some locations. These occurrences are typically associated with local soil characteristics or topography, which are not captured in the coarser climate division regions used in the CPC product. For example, based on comparisons with the in situ ISM, the HRRR produces similarly dry soil moisture conditions within the Sand Hills region of Nebraska (e.g., -101.4 W, 42.1 N in Figure 1). The temporally averaged mean ISM for this location is 138 mm and 165 mm for the in situ and HRRR data, respectively, as compared to 372 mm in the CPC data (2-3x larger).

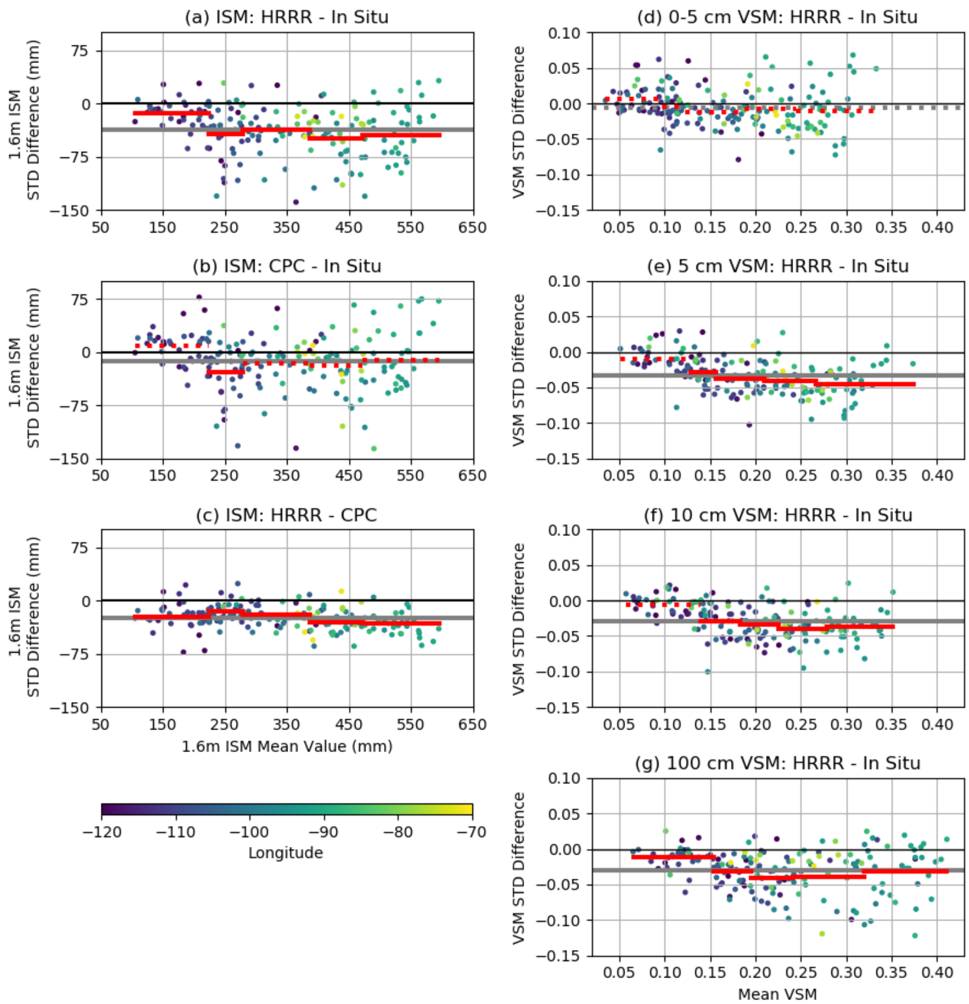
4.2 VSM Mean Comparisons at Varying Depths

To better understand the differences in ISM data, the VSM data at different depths are compared in the HRRR and in situ data (Figure 2d-g). Note, that the VSM data are also separated into quintiles from driest to wettest in a similar way to the ISM data, and this procedure is done for each vertical level. Therefore, locations may fall into different quintiles for different vertical levels. The same trend of the HRRR being wetter in $L_{00\text{-}20}$ (i.e., the driest regions) and dryer in $L_{80\text{-}100}$ (i.e., the wettest regions) is present at all depths. There are generally smaller differences in the middle quintiles. However, at the lowest depths (100 cm below ground; Fig 2g), the differences between the HRRR and in situ VSMs have larger magnitudes, especially for the driest and wettest regimes ($L_{00\text{-}20}$ difference of +0.08 and $L_{80\text{-}100}$ difference of -0.18). Near the surface (Figure 2d), the driest 40% of the regions have relatively small differences ($\pm$ 0.02), although significant dry biases are present for the wetter regions (-0.11). The VSM data demonstrate that the deeper soil layers (i.e., 100 cm below ground) are the primary drivers of the ISM trends. Shallower soil layers have similar trends yet smaller differences as compared to those at deeper levels and have a smaller contribution to the ISM differences. It is important to note that the magnitude of both ISM and VSM mean differences vary throughout the year, and the evolution of these differences as a function of month and season are provided in the supporting information document.

5 Soil Moisture Variance

5.1 ISM Standard Deviation Comparisons

Understanding the temporal variability (e.g., standard deviations) in soil moisture allows for an assessment on whether models are accurately capturing the processes that result in soil moisture changes. Maps of the ISM standard deviations in the HRRR and CPC ISM data show similar patterns, with the highest variance occurring along the Pacific Northwest coast (Figure 1d-e). The lowest ISM standard deviations occur along the Intermountain West and High Plains regions. In general, the HRRR ISMs have lower variance than the CPC data across the US except for in the parts of the Rocky Mountains and in the Great Lakes and Northeast regions (Figure 1f). The HRRR produces larger spatial variability within the mountain and valley regions across the western US that cannot be resolved in the CPC data.



The in situ observations generally have the highest ISM variance for most quintiles followed by CPC and HRRR, which has the lowest variances of the three datasets regardless of the wetness regime (Figure 3a-c). The in situ and CPC quintile mean ISM standard deviations are not significantly different for all quintiles except $L_{20\text{-}40}$, while the HRRR quintile mean differences are significant for all quintiles when compared to both other datasets. Dirmeyer et al. (2016) showed that spatial scaling differences do not have a large impact (~10%) on in situ observation standard deviations via conducting tests where many stations that are separated by several km to up to 100 km are averaged together. Therefore, the differences between the in situ and NOAA modeled standard deviations are likely due to other factors outside of dataset spatial scale differences (e.g., model processes representation or model input data). Furthermore, the large variability or scatter in the ISM standard deviation differences between the in situ data and both models (Figure 3a-b) suggests that the cause of these differences depends on the specific location. The differences between HRRR and CPC ISM standard deviations, however, demonstrate a more systematic bias between these two modeling frameworks (Figure 3c).

Figure 3. Same as Figure 2, except for comparisons of ISM standard deviations (a-c) and VSM standard deviations (d-g).

5.2 VSM Standard Deviation Comparisons at Varying Depths

VSM standard deviations at four different depths are compared between HRRR and in situ data (Figure 3d-g) to determine whether a certain depth is driving the differences in the ISM standard deviations (Figure 3a). Regardless of the soil moisture regime, the HRRR surface VSM standard deviations compare well to the near-surface in situ observations. When averaging over all locations, the mean near-surface percentage differences in HRRR soil moisture from the in situ value is only +4.2% (Figure 3d). This is likely related to improvements made in the HRRR's RUC LSM and the moderately coupled land data assimilation system that have been applied (Benjamin et al., 2022). However, at depths of 5 cm below ground and deeper (Figure 3e-g), most quintiles have statistically significant differences in the standard deviations. The mean percentage differences over all locations are -38.7%, -36.4% and -45.2% for the 5 cm, 10 cm and 100 cm levels, respectively, with the HRRR always having lower standard deviations than the in situ datasets for every quintile. These differences in VSM standard deviations are larger for wetter soil moisture regimes ($L_{20\text{-}100}$). While the differences for the different soil moisture regimes are generally similar for the 5 cm, 10 cm and 100 cm depths, there are more extreme differences for specific locations at the 100 cm level. To summarize, the lower standard deviations in the HRRR ISM are being driven by the lower VSM standard deviations occurring below the surface level. It is important to note that both ISM and VSM standard deviation differences vary throughout the year, and the evolution of these differences as a function of month and season are provided in the supporting information document.

6 Conclusions and Future Work

A comparison of 1.6 m ISM between three different NOAA soil moisture products is conducted. This analysis uniquely includes the HRRR model with its RUC LSM, the CPC leaky-bucket model and in situ observations from two national networks. These soil moisture estimates are used in many operational and research applications, including atmospheric forecasting, drought monitoring, and assessing flood and fire risks. Therefore, quantifying differences in these NOAA models to observational networks across CONUS is critical.

Several conclusions are drawn from these comparisons.

1) The HRRR and CPC ISMs are both larger (i.e., wetter) in the driest regions and smaller (i.e., drier) in the wettest regions as compared to in situ observations.

2) These differences in the HRRR and in situ ISM amounts are largely caused by deep soil levels (~100 cm below ground). Shallower layers have similar trends to the deeper layers but have smaller differences, and thus a weaker contribution to the ISM differences.

3) The in situ observations have the largest ISM standard deviations, followed by the CPC leaky-bucket model and the HRRR model.

4) The HRRR soil moisture standard deviations compare well with the in situ standard deviations near the surface, but large differences are present at 5 cm below the surface and deeper.

The soil moisture differences presented in this study can be caused by a variety of reasons. In terms of modeled soil moisture, biases in the input datasets (i.e., precipitation or radiation), whether they come from a coupled atmospheric model in the case of HRRR or external sources in the case of CPC, have been shown to lead to biases in land surface model calculations (e.g., Mitchell et al., 2004; Min et al., 2021). Choices in the land surface model structure, such as the number and thickness of soil layers, the representation of soil and vegetation, and other model parameters, can also lead to biases in soil moisture prediction (e.g., Mitchel et al., 2004; Xia et al., 2014; 2015b). Min et al. (2021) found that snowmelt, freezing/thawing, and/or biases in precipitation and evapotranspiration led to differences in HRRR soil moisture as compared to in situ observations in New York and that the most relevant processes causing these differences varied throughout the year. The results in this study demonstrate consistent, region-dependent biases in NOAA modeled soil moisture as compared to in situ observations across CONUS, and future research should focus on understanding the model processes that are causing these biases.

Our results also provide important context to the current users of these models and observations. For example, HRRR's land data assimilation system has recently undergone changes that primarily impact the near-surface soil state (Benjamin et al., 2022). The comparisons presented in this study do show better performance of HRRR soil moisture near the surface and thus may provide a first step towards understanding the impact of these model changes. Furthermore, these results can assist with the continued development and refinement of soil moisture models and products. The analyses presented here are currently being utilized for preparing training and validation data for a machine learning algorithm that uses data from the Advanced Baseline Imager on-board NOAA's Geostationary Operational Environmental Satellite to estimate the soil moisture state at very high resolution (i.e., on the order of ~1 km). With a recent focus on land-atmosphere coupling and a continued shift towards higher-resolution models, such a product could be used as a supplementary input for strongly coupled land atmosphere data assimilation in the next generation of atmospheric models.

Acknowledgments

This work was supported by the NOAA FY21 High Performance Computing and Communications Program's Information Technology Incubator. We would also like to acknowledge helpful feedback on and interest in this work from Liaofan Lin, Tanya Smirnova, Stan Benjamin, Curtis Alexander and Eric James.

Open Research

Several in situ and model datasets are used in this study. The USCRN data (Palecki et al., 2013) and the SCAN data (SCAN, 2016) will be archived upon publication. This archiving process is underway and will be with the Mountain Scholar repository through Colorado State University. We have uploaded a copy of these data as Supporting Information for the review process. The CPC data was accessed via https://ftp.cpc.ncep.noaa.gov/wd51yf/us/w_daily/ through the U.S. Data download link on the NOAA CPC product webpage (https://www.cpc.ncep.noaa.gov/products/Soilmst_Monitoring/). The HRRR operational model data (HRRRv3) was stored and accessed via the NOAA Hera supercomputer and is publicly archived at either https://registry.opendata.aws/noaa-hrrr-pds/ or https://console.cloud.google.com/marketplace/product/noaa-public/hrrr. The analysis code used to generate the analyses and figures in this manuscript are available at https://github.com/pjmarinescu/CIRA_Soil_Moisture and will also be archived in the same Mountain Scholar repository as the data upon publication.

References

Arevalo, J., Welty, J., Fan, Y., & Zeng, X. (2021). Implementation of snowpack treatment in the cpc water balance model and its impact on drought assessment. *Journal of Hydrometeorology*, *22*(5), 1235–1247. https://doi.org/10.1175/JHM-D-20-0201.1

Atiah, W. A., Amekudzi, L. K., Akum, R. A., Quansah, E., Antwi-Agyei, P., & Danuor, S. K. (2022). Climate variability and impacts on maize (Zea mays) yield in Ghana, West Africa. *Quarterly Journal of the Royal Meteorological Society*, *148*(742), 185–198. https://doi.org/10.1002/qj.4199

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., et al. (2013). U.S. climate reference network soil moisture and temperature observations. *Journal of Hydrometeorology*, *14*(3), 977–988. https://doi.org/10.1175/JHM-D-12-0146.1

Benjamin, S. G., Smirnova, T. G., James, E. P., Lin, L.-F., Hu, M., Turner, D. D., & He, S. (2022). Land-snow data assimilation including a moderately coupled initialization method applied to NWP. *Journal of Hydrometeorology*. https://doi.org/10.1175/jhm-d-21-0198.1

Carrera, M. L., Bilodeau, B., Bélair, S., Abrahamowicz, M., Russell, A., & Wang, X. (2019). Assimilation of passive L-band microwave brightness temperatures in the Canadian Land data assimilation system: Impacts on short-range warm season numerical weather prediction. *Journal of Hydrometeorology*, *20*(6), 1053–1079. https://doi.org/10.1175/JHM-D-18-0133.1

Delworth, T. L., & Manabe, S. (1988). The Influence of Potential Evaporation on the Variabilities of Simulated Soil Wetness and Climate. *Journal of Climate*, *1*(5), 523–547. https://doi.org/10.1175/1520-0442(1988)001<0523:TIOPEO>2.0.CO;2

Dirmeyer, P. A., Wu, J., Norton, H. E., Dorigo, W. A., Quiring, S. M., Ford, T. W., et al. (2016). Confronting weather and climate models with observational data from soil moisture networks over the United States. *Journal of Hydrometeorology*, *17*(4), 1049–1067. https://doi.org/10.1175/JHM-D-15-0196.1

Dowell, D. C., Alexander, C. R., James, E. P., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., et al. (2022). The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part 1: Motivation and System Description. *Wea. Forecasting*. https://doi.org/10.1175/WAF-D-21-0151.1

Ek, M. B., & Holtslag, A. A. M. (2004). Influence of Soil Moisture on Boundary Layer Cloud Development. *Journal of Hydrometeorology*, *5*(1), 86–99. https://doi.org/10.1175/1525-7541(2004)005<0086:IOSMOB>2.0.CO;2

Fan, Y., & van den Dool, H. (2004). Climate Prediction Center global monthly soil moisture data set at 0.5° resolution for 1948 to present. *Journal of Geophysical Research D: Atmospheres*, *109*(10), 1–8. https://doi.org/10.1029/2003JD004345

Ford, T. W., & Quiring, S. M. (2019). Comparison of Contemporary In Situ, Model, and Satellite Remote Sensing Soil Moisture With a Focus on Drought Monitoring. *Water Resources Research*, 1565–1582. https://doi.org/10.1029/2018WR024039

Guttman, N. B., & Quayle, R. G. (1996). A historical perspective of U.S. climate divisions. *Bulletin of the American Meteorological Society*, *77*(2), 293–303. https://doi.org/10.1175/1520-0477(1996)077<0293:AHPOUC>2.0.CO;2

Huang, J., Van Den Dool, H. M., & Georgakakos, K. P. (1996). Analysis of model-calculated soil moisture over the United States (1931-1993) and applications to long-range temperature forecasts. *Journal of Climate*. https://doi.org/10.1175/1520-0442(1996)009<1350:AOMCSM>2.0.CO;2

James, E. P., Alexander, C. R., Dowell, D. C., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., et al. (2022). The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part 2: Forecast Performance. *Wea. Forecasting*. https://doi.org/10.1175/WAF-D-21-0130.1

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., et al. (2004). Regions of Strong Coupling Between Soil Moisture and Precipitation. *Science*, 305(5687), 1138–1140. https://doi.org/10.1126/science.1100217

Lin, L. F., & Pu, Z. (2020). Improving near-surface short-range weather forecasts using strongly coupled land-atmosphere data assimilation with gsi-enkf. *Monthly Weather Review*, 148(7), 2863–2888. https://doi.org/10.1175/MWR-D-19-0370.1

Liu, J., Zhan, X., Hain, C., Yin, J., Fang, L., Li, Z., & Zhao, L. (2016). NOAA Soil Moisture Operational Product System (SMOPS) and its validations. In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3477–3480). IEEE. https://doi.org/10.1109/IGARSS.2016.7729899

Madadgar, S., AghaKouchak, A., Farahmand, A., & Davis, S. J. (2017). Probabilistic estimates of drought impacts on agricultural production. *Geophysical Research Letters*, *44*(15), 7799–7807. https://doi.org/10.1002/2017GL073606

Min, L., Fitzjarrald, D. R., Du, Y., Rose, B. E. J., Hong, J., & Min, Q. (2021). Exploring Sources of Surface Bias in HRRR Using New York State Mesonet. *Journal of Geophysical Research: Atmospheres*, *126*(20), 1–18. https://doi.org/10.1029/2021JD034989

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., et al. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research: Atmospheres*, 109(7), 1–32. https://doi.org/10.1029/2003jd003823

Muñoz-Sabater, J., Lawrence, H., Albergel, C., Rosnay, P., Isaksen, L., Mecklenburg, S., et al. (2019). Assimilation of SMOS brightness temperatures in the ECMWF Integrated Forecasting System. *Quarterly Journal of the Royal Meteorological Society*, *145*(723), 2524–2548. https://doi.org/10.1002/qj.3577

Palecki, M. A., Lawrimore, J.H., Leeper, R. D., Bell, J. E., Embler, S., Casey, N. (2013). U.S. Climate Reference Network Products, Daily. [Dataset]. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V5H13007. https://www.ncei.noaa.gov/pub/data/uscrn/products/daily01/. Accessed 11 Nov. 2021.

Pan, M., Cai, X., Chaney, N. W., Entekhabi, D., & Wood, E. F. (2016). An initial assessment of SMAP soil moisture retrievals using high-resolution model simulations and in situ observations. *Geophysical Research Letters*, 43(18), 9662–9668. https://doi.org/10.1002/2016GL069964

Quiring, S. M., Ford, T. W., Wang, J. K., Khong, A., Harris, E., Lindgren, T., et al. (2016). The North American Soil Moisture Database: Development and Applications. *Bulletin of the American Meteorological Society*, 97(8), 1441–1459. https://doi.org/10.1175/BAMS-D-13-00263.1

Rigden, A. J., Powell, R. S., Trevino, A., McColl, K. A., & Huybers, P. (2020). Microwave Retrievals of Soil Moisture Improve Grassland Wildfire Predictions. *Geophysical Research Letters*, *47*(23), 1–8. https://doi.org/10.1029/2020GL091410

Robock, A., Luo, L., Wood, E. F., Wen, F., Mitchell, K. E., Houser, P. R., et al. (2003). Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season. *Journal of Geophysical Research: Atmospheres*, *108*(22). https://doi.org/10.1029/2002jd003245

Robock, A., Vinnikov, K. Y., Schlosser, C. A., Speranskaya, N. A., & Xue, Y. (1995). Use of Midlatitude Soil Moisture and Meteorological Observations to Validate Soil Moisture Simulations with Biosphere and Bucket Models. *Journal of Climate*, *8*(1), 15–35. https://doi.org/10.1175/1520-0442(1995)008<0015:UOMSMA>2.0.CO;2

Soil Climate Analysis Network (SCAN) (2016). SCAN Daily Historic Provisional Data. [Dataset]. USDA National Resources Conservation Service National Water and Climate Center. www.wcc.nrcs.usda.gov/scan. Access date 16 Feb. 2022.

Schaefer, G. L., Cosh, M. H., & Jackson, T. J. (2007). The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *Journal of Atmospheric and Oceanic Technology*, *24*(12), 2073–2077. https://doi.org/10.1175/2007JTECHA930.1

Shellito, P. J., Small, E. E., Colliander, A., Bindlish, R., Cosh, M. H., Berg, A. A., et al. (2016). SMAP soil moisture drying more rapid than observed in situ following rainfall events. *Geophysical Research Letters*, 43(15), 8068–8075. https://doi.org/10.1002/2016GL069946

Smirnova, T. G., Brown, J. M., & Benjamin, S. G. (1997). Performance of different soil model configurations in simulating ground surface temperature and surface fluxes. *Monthly Weather Review*, *125*(8), 1870–1884. https://doi.org/10.1175/1520-0493(1997)125<1870:PODSMC>2.0.CO;2

Smirnova, T. G., Brown, J. M., Benjamin, S. G., & Kim, D. (2000). Parameterization of cold-season processes in the MAPS land-surface scheme. *Journal of Geophysical Research: Atmospheres*, 105(D3), 4077–4086. https://doi.org/10.1029/1999JD901047

Smirnova, T. G., Brown, J. M., Benjamin, S. G., & Kenyon, J. S. (2016). Modifications to the Rapid Update Cycle Land Surface Model (RUC LSM) Available in the Weather Research and Forecasting (WRF) Model. *Monthly Weather Review*, 144(5), 1851–1865. https://doi.org/10.1175/MWR-D-15-0198.1

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., et al. (2002). THE DROUGHT MONITOR. *Bulletin of the American Meteorological Society*, *83*(8), 1181–1190. https://doi.org/10.1175/1520-0477-83.8.1181

Taylor, C. M., Gounou, A., Guichard, F., Harris, P. P., Ellis, R. J., Couvreux, F., & De Kauwe, M. (2011). Frequency of sahelian storm initiation enhanced over mesoscale soil-moisture patterns. Nature Geoscience, 4(7), 430–433. https://doi.org/10.1038/ngeo1173

van den Dool, H., Huang, J., & Fan, Y. (2003). Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981-2001. *Journal of Geophysical Research: Atmospheres*, *108*(16), 1–16. https://doi.org/10.1029/2002jd003114

Vinnikov, K. Y., & Yeserkepova, I. B. (1991). Soil Moisture: Empirical Data and Model Results. *Journal of Climate*, *4*(1), 66–79. https://doi.org/10.1175/1520-0442(1991)004<0066:SMEDAM>2.0.CO;2

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. Journal of Geophysical Research Atmospheres, 117(3). https://doi.org/10.1029/2011JD016048

Xia, Y., Sheffield, J., Ek, M. B., Dong, J., Chaney, N., Wei, H., et al. (2014). Evaluation of multi-model simulated soil moisture in NLDAS-2. Journal of Hydrology, 512, 107–125. https://doi.org/10.1016/j.jhydrol.2014.02.027

Xia, Y., Ek, M. B., Wu, Y., Ford, T., & Quiring, S. M. (2015a). Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: Comparison and analysis. *Journal of Hydrometeorology*, *16*(5), 1962–1980. https://doi.org/10.1175/JHM-D-14-0096.1

Xia, Y., Ek, M. B., Wu, Y., Ford, T., & Quiring, S. M. (2015b). Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part II: Impact of soil texture classification and vegetation type mismatches. *Journal of Hydrometeorology*, 16(5), 1981–2000. https://doi.org/10.1175/JHM-D-14-0097.1

Yao, Y., Ciais, P., Viovy, N., Li, W., Cresto-Aleina, F., Yang, H., et al. (2021). A Data-Driven Global Soil Heterotrophic Respiration Dataset and the Drivers of Its Inter-Annual Variability. *Global Biogeochemical Cycles*, *35*(8), 1–23. https://doi.org/10.1029/2020GB006918

Geophysical Research Letters

Supporting Information for

A Comparison of NOAA Modeled and In Situ Soil Moisture Estimates Across the Continental United States

Peter J. Marinescu[1], Kyle Hilburn[1], Daniel Abdi[2], Isidora Jankov[3]

1: Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado
2: Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado
3: Global Systems Laboratory, NOAA/OAR, Boulder, Colorado

Contents of this file

Introduction

In this supporting information document, we provided extended analyses of comparisons of soil moisture mean and standard deviations between our three datasets as a function of month and season. While this seasonal variability was not a focus of our study, we believe that this information may be useful to the broader scientific community. We further note that because the analyses include ~2.4 years of data, the monthly and seasonal analyses have only 2-3 years of data and therefore, outlier results from one year could possibly skew these results. The analyses of quintiles based on mean soil moisture amounts (Figs. S1-S4) follow the same methods as described in the manuscript. Figures (S5-S8) are similar to Figure 1 in the manuscript and represent maps of temporally averaged integrated soil moisture (ISM) means and standard deviations for the different seasons: winter (December-February), spring (March-May), summer (June-August) and autumn (September-November).

Text S1.

It is important to note that the differences in the data's 1.6 m integrated soil moisture (ISM) and volumetric soil moisture (VSM) means and standard deviations vary throughout the year (e.g., Fig. S1-S8). For most locations, the sign of the differences between the datasets are constant throughout the year, although the magnitudes of these differences can change significantly throughout the year. Also, regionally, there are instances where the sign of the ISM difference changes during different times of the years. For example, in the coastal Pacific Northwest region, the ISM mean difference changes from negative in DJF and MAM (Fig. S5-6) to positive in JJA and SON (Fig. S7-8). Similarly, standard deviation differences vary significantly at different times of the year across the Southeast U.S. (Fig. S5-8). These monthly and seasonal changes are likely associated with the changing precipitation patterns and/or soil moisture processes (i.e., evapotranspiration, snow melt and runoff) that vary regionally and seasonally across CONUS.

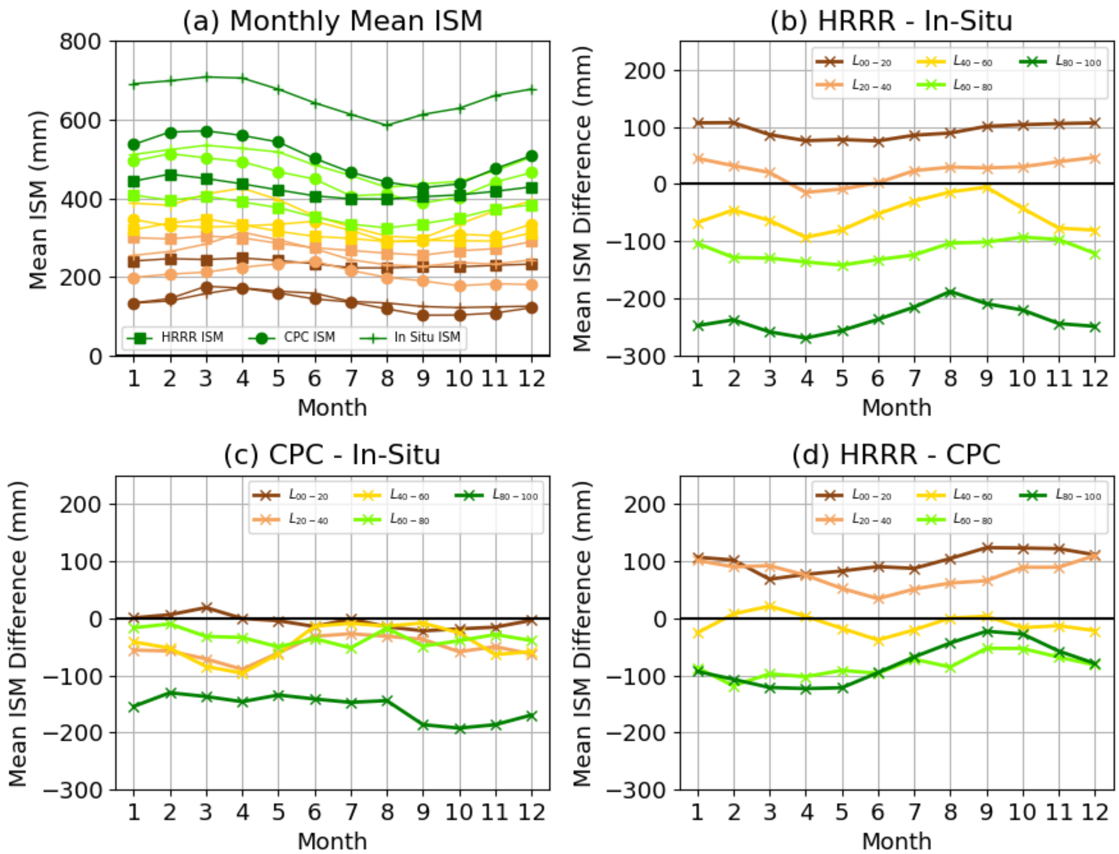


Figure S1. (a) Monthly 1.6 m deep ISM means for HRRR, CPC and in situ data. Data shown is the mean for the five quintiles based on ISM amounts from the driest 20% (brown) to the wettest 20% (dark green). (b-d) shows the mean of the ISM mean differences between all three datasets. Locations are placed into quintiles based on the average ISM for all three datasets.

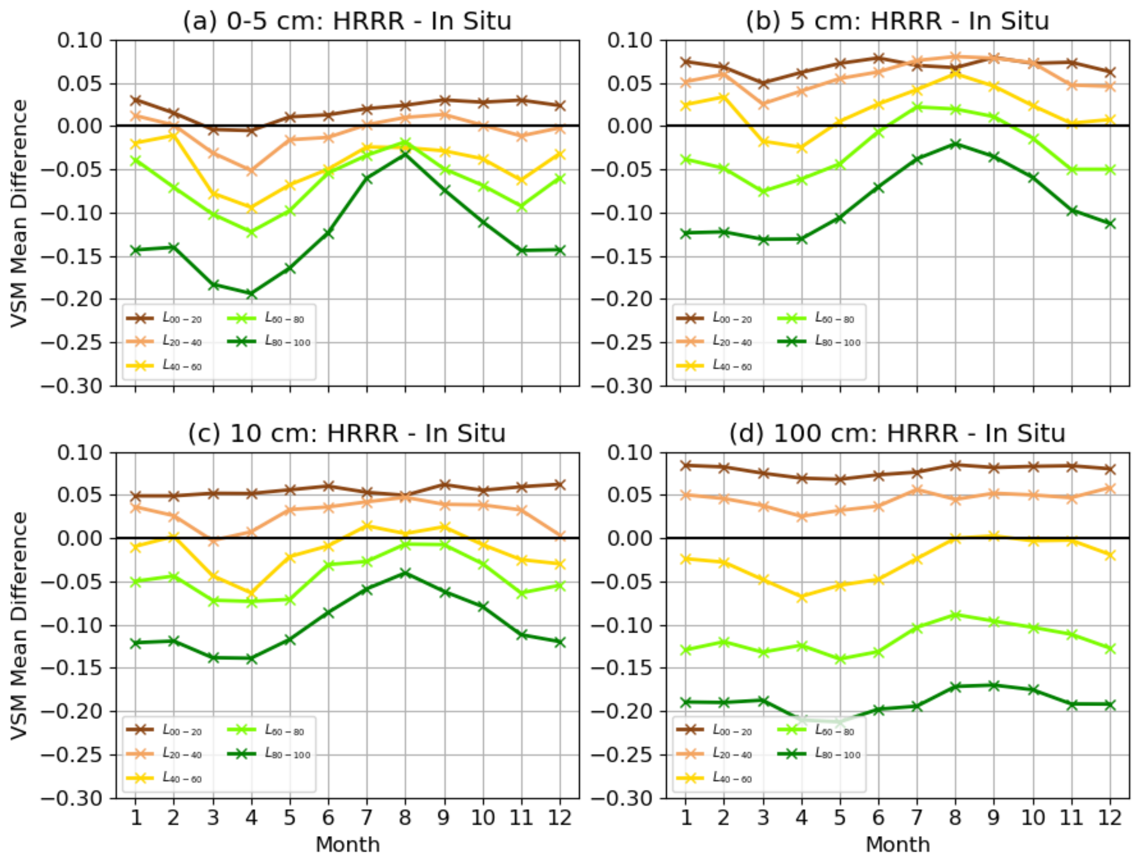


Figure S2. Monthly differences of HRRR and in situ VSM means at four different soil depths: (a) 0 cm in HRRR and 5 cm in the in situ observations; (b) 5 cm, (c) 10 cm, and (d) 100 cm. Data shown is the mean difference for the data separated into five quintiles based on VSM amounts from the driest 20% (brown) to the wettest 20% (dark green). Locations are placed into quintiles based on the average VSM for both HRRR and in situ data and are calculated separately for each soil depth.

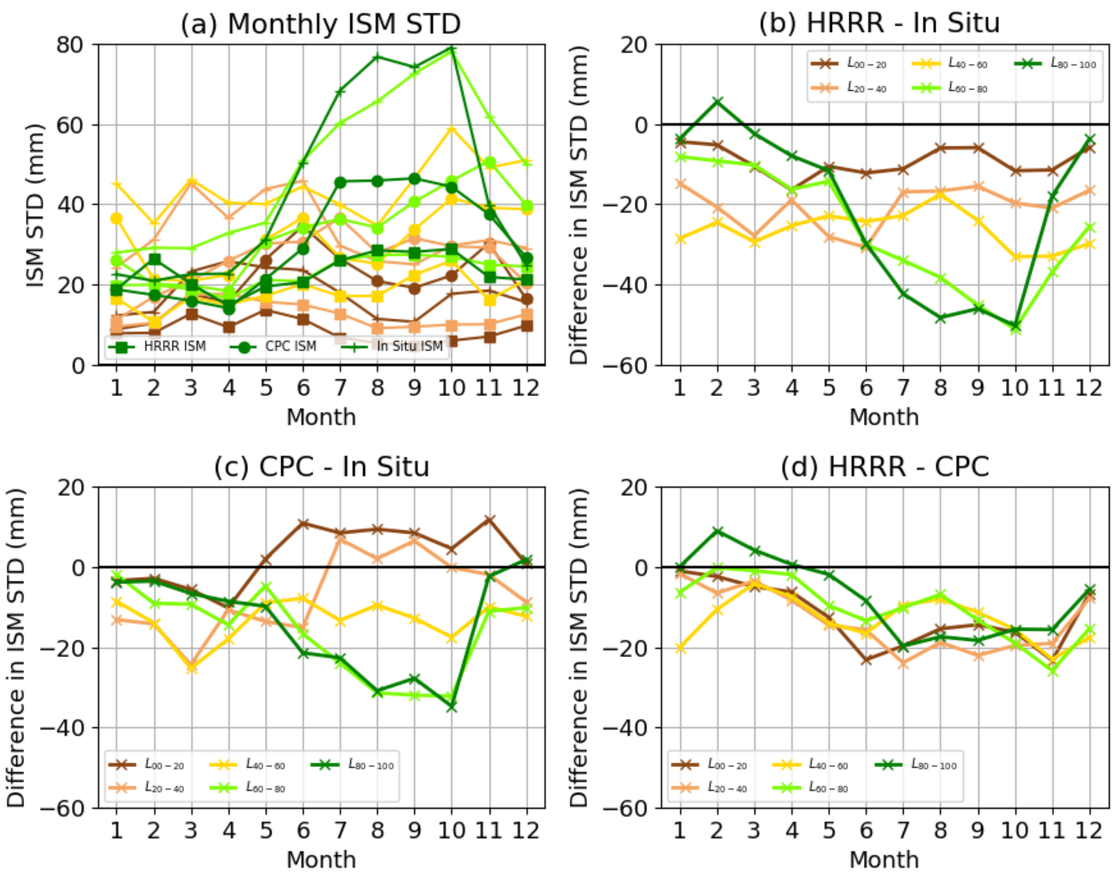


Figure S3. (a) Monthly 1.6 m deep ISM standard deviation for HRRR, CPC, and in situ data. Data shown as the mean of the quintiles from the driest 20% (brown) to the wettest 20% (dark green). (b-d) shows the mean differences of the ISM standard deviation comparisons between all three datasets.

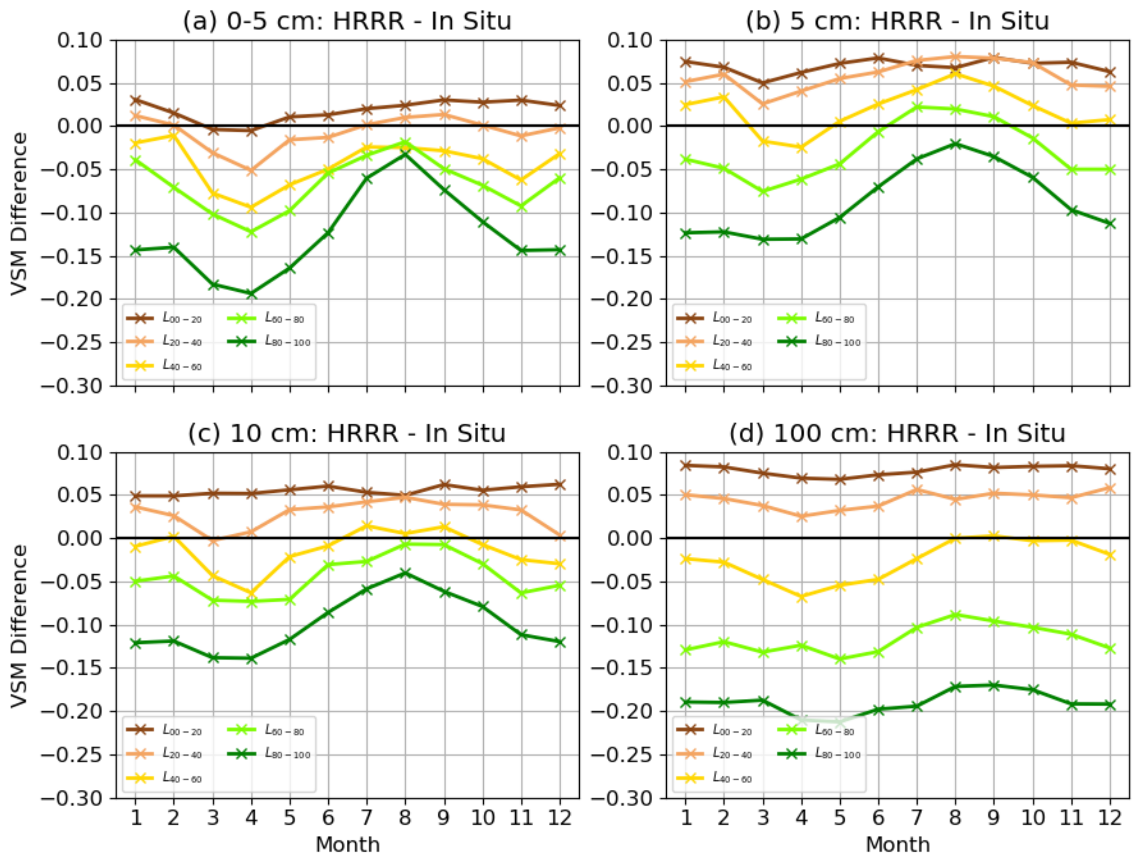


Figure S4. Monthly differences of HRRR and in situ VSM standard deviations at four different soil depths: (a) 0 cm in HRRR and 5 cm in the in situ observations; (b) 5 cm, (c) 10 cm, and (d) 100 cm. Data shown as the mean of the five quintiles of data from the driest 20% (brown) to the wettest 20% (dark green). Quintiles are based on the average VSM for both HRRR and in situ data and are calculated separately for each soil depth.

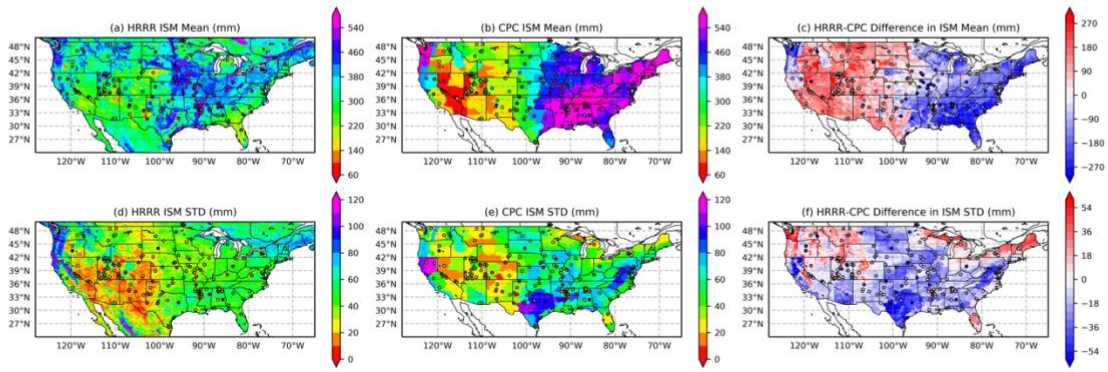


Figure S5. 1.6 m deep ISM values and standard deviations temporally averaged over the study period for the winter months of December, January and February. (a) represents the HRRR model, (b) represents the CPC model, and (c) represents their difference. (d-f) are the same as (a-c), except for the ISM standard deviations. Filled circles represent locations of the 172 in situ observations from the USCRN and SCAN networks that pass quality control checks, as described in the text. In (c) and (f), the filled circles represent the difference HRRR minus in situ observation, while the filled circles represent the in situ station ISM mean and standard deviations, respectively, for the other panels.

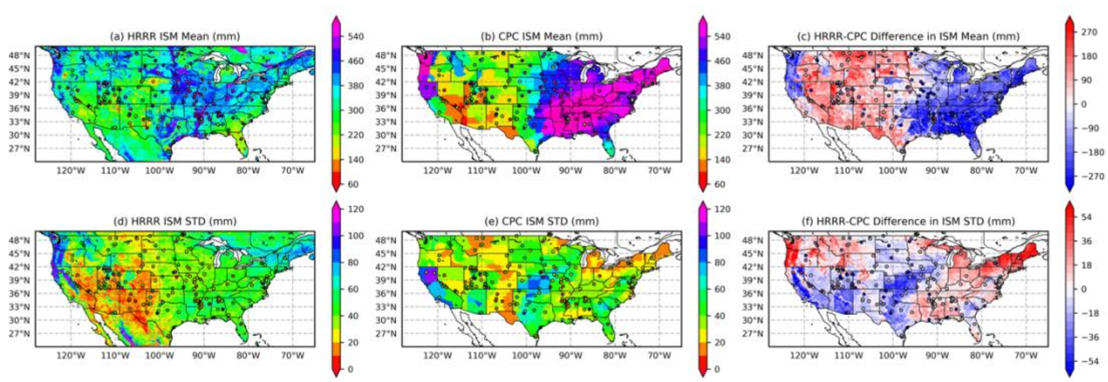


Figure S6. Same as Figure S5, but for the spring months of March, April and May.

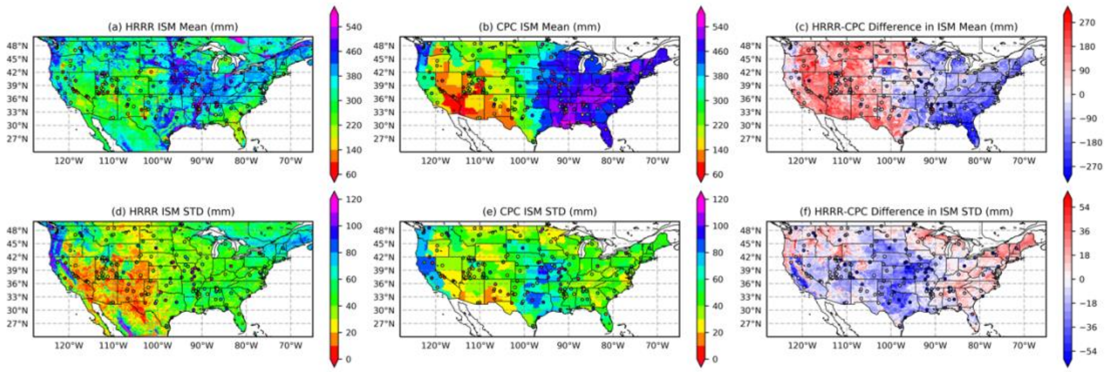


Figure S7. Same as Figure S5, but for the summer months of June, July and August.

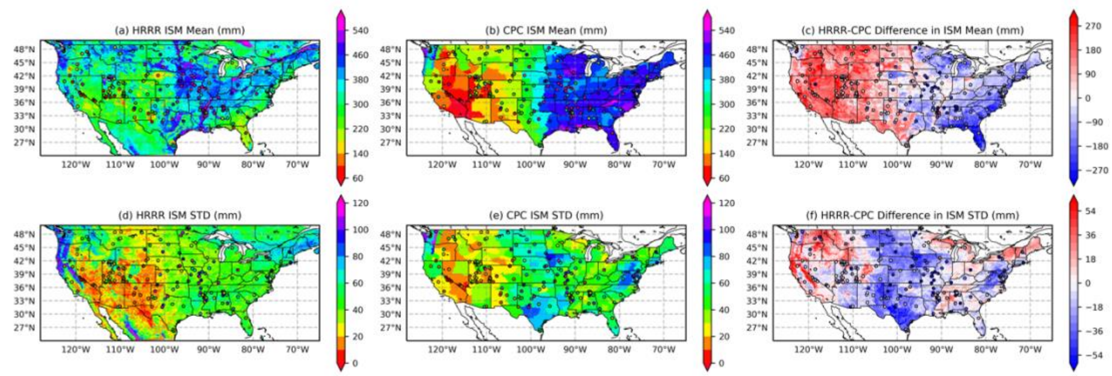


Figure S8. Same as Figure S5, but for the autumn months of September, October and November.