RUBISCO Soil Moisture Working Group (SMWG) Mini-Workshop Report

September 11, 2024

Meeting Co-Chairs
Jiafu Mao, Forrest M. Hoffman, Yaoping Wang
Oak Ridge National Laboratory

Additional Contributors
Charles M. Bachmann, *Rochester Institute of Technology*; Jingyi Huang, *University of Wisconsin–Madison*; Yuefeng Hao, *University of Tennessee at Knoxville*; Lun Gao, *University of Minnesota*; Joshua B. Fisher, *Chapman University*; Trevor Keenan, *Berkeley National Laboratory*; Gabe Kooperman, *University of Georgia*; Gerbrand Koren, *Utrecht University*; Chuxuan Li, *University of California at Los Angeles*; Lingcheng Li, *Pacific Northwest National Laboratory*; Xiaoman Lu, *Oak Ridge National Laboratory*; Elias Massoud, *Oak Ridge National Laboratory*; Hernan Moreno, *University of Texas at El Paso*; Nayma Binte Nur, *Rochester Institute of Technology*; Stephanie Pennington, *Pacific Northwest National Laboratory*; Pushpendra Raghav, *University of Alabama*; Pallav Ray, *Florida Institute of Technology*; Vinit Sehgal, *Louisiana State University*; Mingjie Shi, *Pacific Northwest National Laboratory*; Sreedevi Vasu, *Florida Institute of Technology*; Huiqi Wang, *University of California, Berkeley*

1. Executive Summary

The RUBISCO Soil Moisture Working Group (SMWG), an initiative bolstered by the DOE RUBISCO Science Focus Area project, is committed to unraveling the intricacies of soil moisture and its profound impact on global hydrology, energy, and biogeochemical cycles. Confronted with the complex challenges in multi-scale SM data development, feedback analysis, and model benchmarking, the SMWG aims to catalyze advancements in the field through a synergistic approach that engages leaders in soil moisture research and encompasses diverse datasets and novel analytical techniques. Central to SMWG mission are three pivotal research questions:

- How can we assemble a global soil moisture database that not only integrates the latest remote sensing data and field measurements from initiatives like the DOE NGEEs, AmeriFlux, GEWEX, and ISMN but also applies robust quality control techniques, including machine learning (ML), to offer a completer and more reliable dataset than currently exists?
- In what ways can we advance the mechanistic understanding and predictive modeling of soil moisture responses and biogeochemical feedbacks?
- How can we effectively translate the nuanced statistical properties of soil moisture into evaluation metrics and functional relationships that will allow for the comprehensive benchmarking of Earth system models?

To foster collaboration and share recent progress, the SMWG held a virtual mini-workshop on September 11, 2024. The event brought together around 70 attendees and 15 speakers from diverse fields, including remote sensing, ML, hydrology, and climate science. The primary goal was to share ongoing research efforts, discuss recent advancements, and identify critical challenges in soil moisture estimation and its role in Earth system processes. Presentations covered a broad range of topics, from AI-enhanced soil moisture predictions using unoccupied aerial systems to the challenges of developing soil moisture datasets across critical zones. Interdisciplinary perspectives, such as the role of soil moisture in climate-driven hazards and ecosystem interactions, were explored, emphasizing the need for better data integration and coordination across disciplines. Attendees also discussed the importance of integrating process-based models with ML techniques, improving data quality and benchmarking, and scaling models across different ecosystems. The workshop concluded with a set of key discussions and action items aimed at guiding future research, collaboration, and advancements in soil moisture science. These outcomes will inform subsequent efforts in improving soil moisture estimation, data sharing, and model validation to better address the challenges posed by global climate change.

2. Key Presentations and Findings

- Dr. Moreno's presentation demonstrated the potential of AI, particularly random forest models, to enhance soil moisture estimation from multi-source data, including unoccupied aerial vehicles (UAVs). This research highlighted that ML models can create accurate high-resolution (100m) soil moisture maps. Despite the success of AI-based soil moisture predictions, challenges such as high computational costs, integration of multi-sensor data, and limited availability of high-resolution training data were acknowledged. The findings point to a strong future for UAV and AI integration in soil moisture estimation.
- Dr. Wang (Oak Ridge National Laboratory)'s discussion of soil moisture quality in boreal regions highlighted the need for customized quality control and ML approaches due to extreme temperature environments. Boreal regions exhibit unique challenges such as snow cover and freeze-thaw cycles, which were not adequately considered in current ML models. Future work will focus on testing convolutional neural network (CNN) and long short-term memory (LSTM) models to address these region-specific issues.
- The unified dataset, as presented by Dr. Shi, offers a comprehensive view of soil moisture across the U.S., combining data from multiple sources. By comparing soil moisture estimates from various models and remote sensing products, the study revealed discrepancies in the data, particularly with climate-zone dependent performance and remote sensing providing drier soil estimates than other sources. These findings emphasize the importance of cross-validation and the integration of multiple data sources in soil moisture modeling.
- Dr. Massoud's benchmarking of soil moisture in Earth system models uncovered notable inter-model disagreements in soil moisture, and biases in model predictions of soil moisture and of relationships between soil moisture and ecosystem variables, especially in boreal regions. These findings suggest a need for enhanced datasets and models to better simulate soil moisture-carbon dynamics.
- Dr. Huang's hybrid ML-data assimilation tool ("ML-HRSM2.0") for automatic generation of high-resolution (30-500m) soil moisture and associated uncertainty for cropland ecosystems outperformed conventional data assimilation when compared with existing satellite datasets and in-situ observation sites. This research highlighted the benefits of ML in improving soil moisture predictions for both surface and root-zone layers. However, scaling these predictions to broader ecosystems and incorporating irrigation and human activity data remain as challenges.
- Dr. Raghav's innovative approach to retrieve gap-free daily root zone soil moisture based on evapotranspiration theory was shown to be effective when compared to existing products like ALEXI, VIC, and SMAP L4 over the Oklahoma Mesonet. This method promises improved root zone soil moisture predictions with no need for parameter calibration and minimal data requirements. However, more field validations are needed to generalize these findings.
- Dr. Pennington presented the extensive coastal soil moisture data generated by the COMPASS-FME project. This near-real-time data serves as a valuable tool for model benchmarking and ecosystem monitoring. Addressing sensor calibration and managing large volumes of data are critical challenges that were identified.
- Dr. Bachmann's use of hyperspectral imagery and radiative transfer models proved to be an effective method for retrieving soil moisture in coastal zones. This innovative approach was validated against field measurements and showed promise for future coastal zone applications, though combining hyperspectral and microwave sensing remains a future goal.

- Dr. Li's review emphasized the critical role of soil moisture in climate hazards, such as droughts and floods. The review called for more research into how extreme soil moisture conditions contribute to compound hazards, highlighting gaps in high-quality data for such studies.
- Dr. Hao comprehensively reviewed the influence of soil moisture on greenhouse gas emissions and carbon sequestration across ecosystems. These findings point to the need for models that account for soil moisture dynamics in predicting future carbon cycling under climate change.
- Dr. Sehgal's research into flash droughts showed that using root zone soil moisture as a key metric can improve real-time drought tracking and prediction. Dr. Sehgal also presented a water-and-energy-coupling approach to retrieve root-zone soil moisture at remote sensing footprint scale. The method demonstrated strong correlations with model-based rootzone soil moisture estimates from SMAP level 4 product. However, key challenges to the proposed approach include strong inter-seasonal variability in the active rootzone depth, temporal variability in the terrestrial water energy coupling processes, and observability challenges related to coarse-scale remote-sensing observations of water and energy fluxes.
- Dr. Kooperman's work reviewed the mechanisms of interactions between soil moisture and different types of tropical convection systems, highlighting the increasing variability of soil moisture under climate change. Understanding these interactions is critical for predicting precipitation extremes, but current modeling capabilities are limited due to the lack of tropical observation data.
- Dr. Wang (Berkeley National Laboratory)'s research revisits the previously observed pattern of stronger spatial sensitivity compared to temporal sensitivity in the relationship between ecosystem production and water availability. The findings confirm that spatial-temporal discrepancies persist not only across different plant functional types but also within a single plant functional type. Notably, the spatial and temporal GPP sensitivities approximate each other more closely for soil moisture than for precipitation in water-limited regions. This underscores the crucial role of soil moisture in moderating the effects of rainfall variability on ecosystem function.
- Dr. Vasu's study revealed that soil moisture modulates Madden-Julian Oscillation (MJO) convection, and that MJO-induced precipitation variability affects soil moisture levels. These findings call for more research to improve models that capture the interaction between soil moisture and tropical climate systems.
- Dr. Fisher provided recent updates on the ECOSTRESS, SMAP, and Hydrosat satellite missions, emphasizing their role in monitoring agricultural droughts, forest stress, and food security. These space-based platforms offer critical data for high-resolution soil moisture analysis and will continue to play a pivotal role in global environmental monitoring.

3. Key Discussions

Discussion One: Addressing Challenges and Opportunities in Soil Moisture Estimation

The discussion session began with an exploration of the grand challenges in advancing soil moisture research, particularly focusing on the use of integrative approaches such as ML and ecohydrological modeling. Participants highlighted that sparse data observations remain a fundamental barrier. There are considerable gaps in soil moisture data for many regions (e.g., in boreal zones and croplands) and beyond the surface layer, especially for the root zone. Although satellite observations provide valuable data, they also pose challenges in terms of coverage and resolution. The session also brought attention to the complexity of managing data, the computational infrastructure required, and the intricacies of modeling such complex systems. Uncertainty persists regarding the most appropriate ML techniques, such as LSTM or transformers, and how best to incorporate them into existing models. Additionally, the reliability and explainability of extrapolating ML results for future predictions in complex Earth systems remain a challenge. Participants also discussed significant gaps in research coordination, infrastructure, and data integration. Balancing the use of process-based models with data-driven ML techniques is particularly challenging, especially when scaling point measurements to larger areas. Combining soil moisture data from multiple sources, such as remote sensing and in-situ observations, further complicates the development of reliable and comprehensive soil moisture estimates. One key challenge posed was how to seamlessly integrate these diverse datasets. Several opportunities were discussed as potential solutions to overcome these challenges. Advancements in ML, the increasing availability of satellite data, and the declining cost of sensors were seen as promising developments. Additionally, perspective papers and collaborative research efforts can help bridge the gaps across disciplines and scales. One notable suggestion was to validate new ML models and approaches, using, for example, NASA's SMAP campaigns for ground-truthing efforts. Combining process-based models with ML techniques, such as radiative transfer models, was also highlighted as a promising opportunity for improving soil moisture estimation. Improved coordination between ML experts, hydrologists, and Earth system modelers was identified as a critical need to advance soil moisture research. Participants also highlighted the need to train the next generation of scientists in interdisciplinary methods, encouraging collaboration across different fields. Additionally, organizations and funding agencies like the DOE were recognized as essential in linking young scientists with communities that need soil moisture data, improving data-sharing practices, and encouraging interdisciplinary research efforts. Finally, participants discussed the role that agencies like the DOE could play in furthering scientific advancement. Calls were made for expanding access to DOE tools and infrastructure, such as ILAMB and E3SM, particularly for university researchers. Providing these tools through accessible platforms like Docker was considered a key step. Moreover, agencies were urged to promote better data-sharing systems (e.g., the DOE ESS-DIVE) and to develop the necessary infrastructure for broader tool development and model integration to support future advancements in soil moisture science.

Discussion Two: Addressing Challenges and Opportunities for Soil Moisture in Climate Hazards and Ecosystem Interactions

The second discussion session concentrated on the challenges and opportunities regarding soil moisture's role in climate hazards, ecosystem interactions, and human impacts. A key issue raised was the lack of sufficient data availability and quality standards, particularly for extreme weather events. Participants emphasized the difficulty in selecting appropriate soil moisture data products due to the wide variation in their quality and coverage. Additionally, process-based models were noted to present significant difficulties in terms of benchmarking, calibration, and overall reliability, particularly in the context of climate extremes like flash droughts. Several critical gaps in current knowledge and research infrastructure were highlighted during the discussion. One of the major challenges was constraining model structures using empirical relationships, which are often derived from limited or highly uncertain data. Hazards such as droughts or floods tend to occur on small spatial scales, but many models and remote sensing data are available only at much coarser resolutions, leading to significant scale compatibility issues. Addressing these challenges will require more comprehensive datasets and improved model calibration. Participants also pointed to interdisciplinary collaboration as an opportunity to bridge these knowledge gaps. Specifically, linking soil moisture dynamics with broader fields such as climate science, ecology, and human systems offers promise. Remote sensing technologies were identified as particularly valuable for overcoming some of the data limitations in soil moisture research. Feedback studies within models could further refine the understanding of soil moisture's role across different ecosystems and in response to climate hazards. Additionally, citizen science initiatives were mentioned as a potential method to enhance data collection and improve understanding of soil moisture-related hazards. The importance of collaboration within the soil moisture community, as well as across other disciplines like meteorology, climate science, and agriculture, was strongly emphasized. Over the past decade, collaborations have become more prevalent, with climate and hydrological models being used together to examine extreme weather events. Furthermore, crop science has spurred the development of high-resolution soil moisture products to better inform agricultural irrigation practices, showing the value of cross-disciplinary efforts. Finally, the role of funding agencies, such as the DOE and USDA, was discussed. Both agencies have been pivotal in promoting interdisciplinary research, with the USDA excelling in data product development and the DOE focusing on ecosystem feedbacks and hydrological research. However, a key challenge remains in preserving data records beyond the life cycle of a particular project. Participants called for more support for data creation and tool development to ensure that field campaign data can be scaled and made accessible to the broader scientific community.

4. Action Items

To address the critical challenges highlighted during the workshop, several key action items have been proposed. First, increasing data availability in boreal regions and croplands through new in-situ observation sites and remote sensing initiatives is essential. Uniform standards for soil moisture data collection and metadata documentation must be established to ensure consistency across datasets and models. A comprehensive integration framework for remote sensing, UAV data, in-situ measurements, and ML products should be developed to enable better soil moisture modeling for climate and ecosystem analysis. Advanced ML models such as CNNs, LSTMs, and transformers should be tested and validated against field measurements and multi-source datasets, while efforts to combine process-based models with ML techniques, like radiative transfer models, need further exploration. Emphasis on creating physics-informed or knowledge-guided models through the integration of domain knowledge with ML approaches will enhance the representation of soil moisture dynamics. There is also a need for increased investment in computational power, data management, and cloud-based systems for large-scale data analysis, and broader access to DOE tools, such as ILAMB and E3SM, for universities and researchers should be facilitated through cloud platforms. Strong partnerships between ML experts, hydrologists, ecosystem scientists, and Earth system modelers are crucial, alongside programs to train the next generation of interdisciplinary researchers. Citizen science initiatives can be explored to improve soil moisture data collection in underrepresented regions. Continuing efforts to refine model-data integration and benchmarking methods will improve predictions of soil moisture and its interactions with variables such as gross primary productivity and carbon sequestration. Research on soil moisture's role in climate hazards, buffering rainfall variability, and improving flash drought prediction will be key in addressing climate resilience. Finally, engaging agencies like DOE, USDA, WMO, FAO and others will enhance interdisciplinary research, data sharing, and long-term preservation of research data beyond individual project lifecycles.

5. Conclusion

The RUBISCO SMWG Mini-Workshop highlighted significant advancements in soil moisture research, particularly the integration of ML, remote sensing, and ecohydrological models. Presentations underscored the central role soil moisture plays in influencing global hydrology, energy fluxes, and biogeochemical cycles. A key takeaway was the urgent need for improved datasets and more refined models capable of accurately capturing the complexity of soil moisture dynamics across varied landscapes, such as boreal regions, croplands, and coastal zones. The workshop also identified several grand challenges, including the scarcity of high-resolution soil moisture data, difficulties in integrating multi-source datasets, and issues related to the scalability of ML techniques. Interdisciplinary collaboration remains crucial, as integrating the expertise of hydrologists, ML specialists, and Earth system modelers is essential to address these challenges. Action items emerged to tackle these issues through better data standardization, enhanced model benchmarking, and the development of a robust data-sharing infrastructure. Looking ahead, the RUBISCO SMWG is poised to take a leading role in advancing soil moisture science by fostering global collaborations, developing comprehensive datasets, and refining models to enhance predictive capabilities. By leveraging the growing potential of ML and integrating diverse data sources, the group is well-positioned to contribute to more accurate soil moisture monitoring, improved climate resilience strategies, and sustainable land management practices.

Appendix: RUBISCO Soil Moisture Working Group Virtual Mini-Workshop Agenda (U.S. EDT)
September 11, 2024
(*8-minute presentation followed by a 2-minute Q&A period*)

11:00am Jiafu Mao/Forrest Hoffman/Yaoping Wang: Welcome, goals of meeting, and RUBISCO SMWG status update

[Section One Moderator: Jiafu]

Topic #1: Integrative Approaches in Soil Moisture: From Machine Learning-Enhanced Products to Global Ecohydrological Modeling and Benchmarking

11:10am Hernan Moreno - AI for soil moisture estimation from multi-source data including unmanned aerial systems [Current T1.1]

11:20am Yaoping Wang/Jiafu Mao - Soil moisture quality check and machine learning-based prediction in boreal (>45ºN) regions [Current T1.2]

11:30am Mingjie Shi/Lingcheng Li - Unified Ensemble Soil Moisture Dataset Across the Continental United States [Current T1.3]

11:40am Elias Massoud - Benchmarking Soil Moisture and Ecohydrologic Interactions in Earth System Models [Current T1.4]

11:50am Jingyi Huang, University of Wisconsin-Madison - Machine-Learning High-Resolution Soil Moisture (ML-HRSM 2.0): performance in US cropland [Current T1.5]

12:00pm Pushpendra Raghav - Retrieving Gap-Free Daily Root Zone Soil Moisture Using Surface Flux Equilibrium Theory [New T1.6?]

12:10pm Discussion One [Moderator: Forrest]

12:40pm Break

[Section Two Moderator: Yaoping]

1:10pm Stephanie Pennington - 50 million observations and counting: open availability of near-real-time COMPASS-FME soil moisture data [New T1.7?]

1:20pm Charles M. Bachmann/Nayma Binte Nur - Radiative transfer models of Hyperspectral Imagery for Retrieval of Soil Moisture and Related Biophysical Parameters in the Coastal Zone [New T1.8?]

Topic #2: Interdisciplinary Perspectives on Soil Moisture Dynamics: From Climate Hazards to Ecosystem Interactions and Human Impacts

1:30pm Chuxuan Li/Gerbrand Koren- The role of soil moisture in climate-driven compound hazards: a review [Current T2.1]

1:40pm Yuefeng Hao/Jiafu Mao - Soil Moisture Controls over Greenhouse Gas Emissions and Carbon Sequestration: A Review and Future Directions [Current T2.2]

1:50pm Vinit Sehgal/Josh Fisher - Drought-related applications of multiscale soil moisture-with a specific focus on flash droughts [Current T2.3]

2:00pm Gabe Kooperman/Pallav Ray - The Roles of Soil Moisture–Convection Interactions in the Tropics: A Review [Current T2.4]

2:10pm Huiqi Wang/Trevor Keenan - Soil Buffers the Impact of Rainfall Variability on Ecosystem Function [New T2.5?]

2:20pm Sreedevi Vasu/Pallav Ray - On the relationship between soil moisture and the Madden-Julian oscillation in the Maritime continent [New T2.6?]

2:30pm Josh Fisher - Quick updates on ECOSTRESS, SMAP, and Hydrosat

2:40pm Discussion Two, Wrap-up and Close-out [Moderators: Jiafu and Forrest]