Improving mesoscale soil moisture mapping with in situ networks and cosmic-ray neutron probes

by

Joaquin Arturo Peraza Rud

B.S., Universidad de la República, Uruguay, 2021

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Approved by:

Major Professor
Andres Patrignani

Copyright

Abstract

Soil moisture plays a vital role in the water and energy balance of rainfed agricultural and hydrological systems in the U.S. Great Plains. With the rise of mesoscale environmental monitoring networks that include soil moisture as part of the standard suite of measurements, this dissertation addresses a central question: *Can we upscale point-level soil moisture observations from sparse monitoring networks by integrating spatial model estimates and in situ observations through data assimilation*? The first chapter investigates a data-fusion method to leverage existing mesoscale rootzone soil moisture to create new, high-spatial resolution soil moisture maps for the state of Kansas. These maps are extensively validated using cross-validation and regional surveys conducted with a roving cosmic-ray neutron detector. The second chapter presents CRNPy, a Python library that compiles common correction methods for converting raw neutron counts into volumetric soil water content. The CRNPy library is then used to validate soil moisture estimates at watershed and state levels. The third chapter explores the use of a machine learning observation operator to translate soil moisture observations from stations of the Kansas Mesonet typically located in areas with perennial grassland vegetation to represent soil moisture conditions in nearby cropland fields. Collectively, these chapters offer new insights for leveraging *in situ* and proximal soil moisture data to advance our understanding of regional and mesoscale soil moisture, with a focus on harnessing the expanding infrastructure of mesoscale monitoring networks.

Table of Contents

List of Figures vi
List of Tables ix
Acknowledgements x
Chapter 1 - General Introduction 1
References 5
Chapter 2 - Mapping Mesoscale Rootzone Soil Moisture Using a Model-Data Fusion Approach 11
Abstract 11
Introduction 13
Methodology 16
In situ soil moisture observations 16
Soil water storage model 17
Data Assimilation 19
Validation of soil moisture maps 19
Results and discussion 22
Model-Data fusion cross-validation 22
Spatial validation using a CRNP rover 24
Daily soil water loss 25
Soil properties 26
Limitations of this study 26
Conclusions 28
References 29
Chapter 3 - CRNPy: An Open-Source Python Library for Cosmic-Ray Neutron Probe Data Processing 46
Abstract 46
Introduction 48
Library computation routines 50
Incoming neutron flux correction 50
Atmospheric pressure correction 51

Atmospheric humidity correction 52
Biomass correction 52
Road correction 53
Converting corrected counts into volumetric water content 54
Sensing depth 55
Outlier detection 55
Exponential filter 56
Library Features 57
Description of Use-case Scenarios 58
Use-case scenario 1 – Comparison of stationary CRNP corrections 58
Use-case scenario 2 – Temporal filter of neutron counts vs. soil moisture 59
Use-case scenario 3 – Spatial filtering of neutron counts vs. soil moisture 60
Results and discussion 62
Use-case scenario 1 – Comparison of stationary CRNP corrections 62
Use case scenario 2 – Temporal filter of neutron counts vs. soil moisture 63
Use-case scenario 3 – Spatial filtering of neutron counts vs. soil moisture 63
Conclusions 65
References 66
Chapter 4 - Using an Observation Operator to Translate Soil Moisture Conditions Across Land Covers 85
Abstract 85
Introduction 86
Materials and Methods 89
Soil moisture data 89
Vegetation data 89
Results and discussion 93
Conclusion 95
References 96
Chapter 5 - General conclusion 105
Appendix A - Reference ET vs. λ 108
Appendix B - CRNP devices 109

List of Figures

Figure 2-1. Spatial configuration of 74 stations composing the Kansas Mesonet 37
Figure 2-2. Soil water storage dynamics observed at the Gypsum Kansas Mesonet station 38
Figure 2-3. Plant available water (PAW) evolution in an 8-days period. 39
Figure 2-4. Median absolute error (MAE) of the soil water storage model (A), interpolated *in situ* observations (B), and the proposed model-data fusion approach (B) based on a leave-one-out cross-validation approach. 40
Figure 2-5. 1:1 plot of a randomly selected 5% sample of daily results from the cross-validation 41
Figure 2-6. Time series of the map soil water storage (0 - 50 cm) daily average for the Gypsum Area, the Gypsum Kansas Mesonet station top 50 cm soil water storage observations, and each rover transect average 42
Figure 2-7. Vapor pressure deficit (VPD) negative linear relationship with the soil water storage loss coefficient (λ). 43
Figure 2-8. Fitted linear models between sand content extracted from SoilGrids and the observed upper and lower limits of soil water storage at each location of the Kansas Mesonet. 44
Figure 2-9. Maps of soil water storage upper and lower limit for the state of Kansas created using the adjusted linear regression between sand content and soil water storage limits. 45
Figure 3-1. Example workflow for stationary CRNP, dashed lines represent optional steps. 77
Figure 3-2. Example workflow for roving CRNP, dashed lines represent optional steps. 78
Figure 3-3. Volumetric water content estimated from 1) neutron counts corrected by hourly pressure, humidity, and incoming neutron flux and 2) neutron counts corrected by hourly pressure and mean humidity for the site. 79
Figure 3-4. Raw and corrected neutron counts recorded with a stationary CRNP at the Konza Prairie Biological Station near Manhattan, Kansas. 80
Figure 3-5. A) Barometric pressure and air absolute humidity, B) incoming neutron flux recorded by the Irkutst Neutron Monitor part of the Neutron Monitor Database (NMDB), which has similar elevation and cutoff rigidity as the experimental site, and C) correction factors for the stationary CRNP within the Konza Prairie Biological Station near Manhattan, Kansas. 81

Figure 3-6. Volumetric water content obtained by smoothing using an 11-hour third-order Savitzky-Golay filter in the corrected neutron counts signal and smoothing the resulting volumetric water content. 82
Figure 3-7. Volumetric water content from the Kings Creek *in situ* soil moisture monitoring network, averaging 5 and 20 cm depth sensors with the standard error band, daily SMAP L4 surface soil moisture values for the pixel covering the watershed, and the average of each CRNP roving transects with the standard deviation of the corrected counts or volumetric water content estimations. 83
Figure 3-8. Comparison of pressure observations from the CRNP onboard barometer, and different stations of the Kansas Mesonet 84
Figure 4-1. Orthoimage illustrating the location of the Ashland Bottoms station (ASB) of the Kansas Mesonet, the adjacent plot experiment that was used for both calibration and validation of the observation operator, and the KONA cropland site of the National Ecological Observatory Network (NEON KONA) 99
Figure 4-2. Exponential model used to estimate EVI from canopy cover observations 100
Figure 4-3. Timeline showing the different land covers for each crop rotation of the replicated plot experiment (CR1 – CR5), the KONA site of the National Ecological Observatory Network (NEON), and the Ashland Bottoms station of the Kansas Mesonet. 101
Figure 4-4. A) Daily precipitation, observed soil water storage for the ASBK1 station of the Kansas Mesonet, observed cropland soil moisture for the CR2 validation rotation of the replicated plot experiment, and predicted soil water storage for the rotation site using the random forest observation operator. Depicted soil water storage considers the top 50 cm of the soil profile from October 2020 to August 2023. B) enhanced vegetation index (EVI) time series of observed values in the ASB station and CR2 validation rotation. 102
Figure 4-5 A) Daily precipitation, observed soil water storage for the ASB station of the Kansas Mesonet, observed cropland soil moisture for the KONA station of the National Ecological Observatory Network (NEON), and predicted soil water storage for the NEON KONA site using the random forest observation operator. Soil water storage is for the top 50 cm of the soil profile from September 2020 to August 2023. B) enhanced vegetation index (EVI) time series of observed values in the ASB station and NEON KONA validation rotation. 103

Figure 4-6. Timeline of the observed values of soil water storage (A) and Enhanced Vegetation Index (B) dynamics for each crop rotation of the replicated plot experiment (CR1 - CR5), the KONA site of the National Ecological Observatory Network (NEON), and the Ashland Bottoms station of the Kansas Mesonet from October 2020 to August 2023..... 104

List of Tables

Table 3-1. Average raw neutron counts per minute, atmospheric pressure, absolute humidity and reference incoming neutron flux obtained from the Dourbes (3.18 GV, 225 m a.s.l) neutron monitor in Belgium part of the NMDB network, for each CRNP rover transect collected in the Kings Creek watershed. 74

Table 3-2. Correction factors range and Mean Absolute Error (MAE) of three different approaches, using only the atmospheric pressure correction sourced from the collocated barometer, using the atmospheric pressure correction factor from the collocated barometer and a correction factor for absolute humidity computed from long term site average observed in the Ashland Bottoms Kansas Mesonet station, and using the atmospheric pressure, absolute humidity, correction factor from collocated sensors and the incoming neutron flux corrections. 75

Table 3-3. Average corrected neutron count and estimated volumetric water content for each rover transect of the Kings Creek watershed. *In situ* observations of sensors at 5 and 15 cm depth that are part of the NEON KONA site, and value of surface volumetric water content (0 – 5 cm) of a SMAP pixel that covers the watershed..... 76

Table 4-1. Random Forest (RF) Feature importance. Daily values of vegetation residuals from the 60 days window passed to the model were grouped by week for clarity purposes. DOY feature represents the day of the year (0-365)..... 98

Acknowledgements

I would like to acknowledge my family for their unwavering support throughout my career. My deepest gratitude goes to Dr. Andres Patrignani for providing me with the opportunity to pursue my M.S. degree and for the valuable guidance and hands-on experience in applied soil physics that I gained during my graduate studies. I also want to acknowledge the support and feedback that I received from my committee members, Dr. Ignacio Ciampitti and Dr. Eduardo Santos. Additionally, I am grateful to my fellow graduate students for their consistent support, both academically and personally.

Chapter 1 - General Introduction

Soil moisture is a fundamental component of the soil water balance and the surface energy balance, representing a key variable in meteorological, ecological, hydrological, and agricultural decision-making process, as well as in modeling and forecasting (Ochsner et al., 2013). For instance, soil moisture plays an important role in flash flooding predictions by partitioning rainfall into infiltration and runoff (Crow et al., 2018; Hino et al., 1988; Jacobs et al., 2003; Ruggenthaler et al., 2016), in agricultural drought monitoring by modulating the amount of available water for plant growth and yield production (Denmead & Shaw, 1960; Ford et al., 2015; Lollato et al., 2017), in wildfire preparedness by controlling the amount of fuel biomass (Cooke et al., 2012; Krueger et al., 2017, 2022; Rigden et al., 2020), and in mesoscale weather dynamics by regulating land-atmosphere feedback (Koster et al., 2003).

The relevance and need for accurate *in situ* soil moisture information over large spatial scales has been recognized since the 1950s, when sparse observations were collected using the gravimetric method and neutron probes as part of the RUSWET-GRASS and RUSWET-AGRO networks across the former Soviet Union (Dorigo et al., 2011; Robock et al., 2000). With the advent of electronic dataloggers and new soil moisture sensors based on heat pulse probes (Bristow et al., 1993; Campbell et al., 1991) and time domain reflectometry sensors (Dalton & Genuchten, 1986; Topp & Reynolds, 1998) that enabled automated soil moisture measurements (Baker & Allmaras, 1990), existing environmental monitoring networks and agricultural weather networks started to adopt soil moisture as a standard measurement variable. One of the first mesoscale networks to integrate automated soil moisture observations was the Oklahoma Mesonet (Brock et al., 1995; McPherson et al., 2007), which started in 1996 and currently has more than 110 soil temperature and moisture monitoring sites (McPherson et al., 2007),

constituting one of the most comprehensive datasets of *in situ* soil moisture globally. Following the Oklahoma Mesonet, other statewide mesoscale networks, including the West Texas Mesonet (Schroeder et al., 2005), the Nebraska Mesonet (Shulski et al., 2018), the Kentucky Mesonet (Mahmood et al., 2019), the New York Mesonet (Brotzge et al., 2020), the Kansas Mesonet (Patrignani et al., 2020), and the North Carolina ECONet (Saia et al., 2023) started to adopt rootzone soil moisture measurements. Similarly, nationwide mesoscale monitoring such as the U.S. Climate Reference Network (USCRN) (Palecki et al., 2013) and the Soil Climate Analysis Network (SCAN) (Schaefer et al., 2007) adopted soil moisture observations to provide accurate soil moisture observations at high-temporal (i.e., minute, hourly, daily) resolution for advancing climate research. Yet, the sparse distribution of these networks and the small sensing footprint (<1,000 cm^3) of conventional point-level soil moisture sensors (Patrignani et al., 2022; Vaz et al., 2013) hinder a continuous representation of soil moisture conditions across the spatial domain of the *in situ* networks.

To upscale *in situ* observations and effectively bridge the gap between point-level data from individual monitoring stations and the coarser (i.e., several kilometers) spatial scope of land-surface model predictions and remote sensing products, researchers have employed data assimilation frameworks (Ni-Meister, 2008). These frameworks enable the consolidation of both magnitude and uncertainty data from multiple sources of soil moisture information (De Lannoy et al., 2019). For instance, some studies used a Kalman filter to combine sparse *in situ* measurements of soil moisture with a parsimonious model to generate a spatially continuous representation of soil moisture conditions (Gruber et al., 2015). In this context, *in situ* soil moisture observations from mesoscale networks represent an accurate, but discrete, source of soil moisture information, which may be useful to remove model uncertainties. Thus, in this study we

propose to use a model-data assimilation approach to develop 250-m spatial resolution daily maps of rootzone (0-50 cm depth) soil moisture across Kansas by leveraging the existing infrastructure of the Kansas Mesonet, which currently have 54 stations across Kansas equipped with research-grade soil moisture sensors. However, a critical aspect of newly developed soil moisture maps is the validation of the soil moisture estimates across large spatial extents and multiple land covers. The validation process is particularly challenging because mesoscale networks are often deployed for long-term monitoring in undisturbed land and tend to have a bias towards grassland soil moisture conditions (Brown et al., 2023; Patrignani & Ochsner, 2018), which make common validation methods, such as the leave-one-out cross-validation method, ineffective. To circumvent this limitation, we propose to use of a roving cosmic-ray neutron detector capable of monitoring soil moisture conditions on-the-go across landscapes with intermixed land covers to strengthen the validation of the resulting soil moisture maps.

The cosmic-ray neutron sensing technique is an emerging method that bridges the measurement gap between point-level soil moisture sensors and remote sensing observations of soil moisture, which have coarse spatial resolutions and mostly represent the soil's skin layer (i.e., 0-5 cm). Cosmic-ray neutron probes (CRNPs) operate on the principle of measuring the intensity of epithermal neutrons generated when cosmic rays interact with Earth's atmosphere (Zreda et al., 2008). These epithermal neutrons have energies ranging from 0.5 (electronvolts, eV) to 100 keV and exhibit high sensitivity to energy loss (i.e., thermalization) due to elastic collisions with hydrogen atoms because of their relatively low mass (Köhli et al., 2015). Thus, when an epithermal neutron collides with a hydrogen atom, it transfers a significant portion of its kinetic energy to the hydrogen atom. As a result, the intensity of epithermal neutrons close to the Earth's surface is inversely correlated with the hydrogen content within the upper decimeters of

the soil. This hydrogen pool is typically dominated by the soil water content, making the technique suitable for non-invasive measurements of soil moisture at the field scale (i.e., 10-30 hectares) (Desilets et al., 2010). However, converting raw neutron counts into volumetric soil water content requires several steps and corrections that are spread across multiple peer-reviewed studies. Thus, there is need for developing a comprehensive code library that encompass the most common correction routines required to convert raw neutron counts into soil moisture estimates. The second chapter of this thesis introduces CRNPy, a new Python library specifically designed to simplify and standardize the processing of CRNP data by compiling common methods documented across multiple peer-reviewed manuscripts spanning the past 15 years, which approximately represents the period since the inception of this technique (Zreda et al., 2008) to the present time.

Another challenge when upscaling *in situ* soil moisture information across landscapes with intermixed land covers is the variability in soil moisture conditions driven by vegetation cover (Patrignani & Ochsner, 2018). In the final chapter of this thesis, we investigate the correlation between rootzone soil moisture measurements at stations of the Kansas Mesonet, which are often recorded under warm-season grassland vegetation, and rootzone soil moisture conditions in adjacent cropland fields. The goal was to explore the possible use of a machine learning observation operator to translate and leverage readily available soil moisture observations collected by the Kansas Mesonet into cropland soil moisture conditions.

References

Baker, J. M., & Allmaras, R. R. (1990). System for Automating and Multiplexing Soil Moisture Measurement by Time-Domain Reflectometry. Soil Science Society of America Journal, 54(1), 1–6. https://doi.org/10.2136/sssaj1990.03615995005400010001x

Bristow, K. L., Campbell, G. S., & Calissendorff, K. (1993). Test of a heat-pulse probe for measuring changes in soil water content. Soil Science Society of America Journal, 57(4), 930–934. https://doi.org/10.2136/sssaj1993.03615995005700040008x

Brock, F. V., Crawford, K. C., Elliott, R. L., Cuperus, G. W., Stadler, S. J., Johnson, H. L., & Eilts, M. D. (1995). The Oklahoma Mesonet: A technical overview. Journal of Atmospheric and Oceanic Technology, 12(1), 5–19. https://doi.org/10.1175/1520-0426(1995)012%3C0005:TOMATO%3E2.0.CO;2

Brotzge, J. A., Wang, J., Thorncroft, C. D., Joseph, E., Bain, N., Bassill, N., Farruggio, N., Freedman, J. M., Hemker, K., Johnston, D., Kane, E., McKim, S., Miller, S. D., Minder, J. R., Naple, P., Perez, S., Schwab, J. J., Schwab, M. J., & Sicker, J. (2020). A Technical Overview of the New York State Mesonet Standard Network. Journal of Atmospheric and Oceanic Technology, 37(10), 1827–1845. https://doi.org/10.1175/JTECH-D-19-0220.1

Brown, W. G., Cosh, M. H., Dong, J., & Ochsner, T. E. (2023). Upscaling soil moisture from point scale to field scale: Toward a general model. Vadose Zone Journal, 22(2), e20244. https://doi.org/10.1002/vzj2.20244

Campbell, G. S., Calissendorff, C., & Williams, J. H. (1991). Probe for Measuring Soil Specific Heat Using A Heat-Pulse Method. Soil Science Society of America Journal, 55(1), 291–293. https://doi.org/10.2136/sssaj1991.03615995005500010052x

Cooke, W. H., Mostovoy, G. V., Anantharaj, V. G., & Jolly, W. M. (2012). Wildfire Potential Mapping over the State of Mississippi: A Land Surface Modeling Approach. GIScience & Remote Sensing, 49(4), 492–509. https://doi.org/10.2747/1548-1603.49.4.492

Crow, W. T., Chen, F., Reichle, R. H., Xia, Y., & Liu, Q. (2018). Exploiting Soil Moisture, Precipitation, and Streamflow Observations to Evaluate Soil Moisture/Runoff Coupling in Land Surface Models. Geophysical Research Letters, 45(10), 4869–4878. https://doi.org/10.1029/2018GL077193

Dalton, F. N., & Genuchten, M. T. V. (1986). The time-domain reflectometry method for measuring soil water content and salinity. Geoderma, 38(1), 237–250. https://doi.org/10.1016/0016-7061(86)90018-2

De Lannoy, G. J., Rosnay, P., & Reichle, R. H. (2019). Soil moisture data assimilation. Springer Nature.

Denmead, O. T., & Shaw, R. H. (1960). The Effects of Soil Moisture Stress at Different Stages of Growth on the Development and Yield of Corn 1. Agronomy Journal, 52(5), 272–274. https://doi.org/10.2134/agronj1960.00021962005200050010x

Desilets, D., Zreda, M., & Ferré, T. P. A. (2010). Nature's neutron probe: Land surface hydrology at an elusive scale with cosmic rays: Nature's Neutron Probe. Water Resources Research, 46(11). https://doi.org/10.1029/2009WR008726

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, A., & Jackson, T. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. Hydrology and Earth System Sciences, 15(5), 1675–1698. https://doi.org/10.5194/hess-15-1675-2011

Ford, T. W., McRoberts, D. B., Quiring, S. M., & Hall, R. E. (2015). On the utility of in situ soil moisture observations for flash drought early warning in Oklahoma, USA. Geophysical Research Letters, 42(22), 9790–9798. https://doi.org/10.1002/2015GL066600

Gruber, A., Crow, W., Dorigo, W., & Wagner, W. (2015). The potential of 2D Kalman filtering for soil moisture data assimilation. Remote Sensing of Environment, 171, 137–148. https://doi.org/10.1016/j.rse.2015.10.019

Hino, M., Odaka, Y., Nadaoka, K., & Sato, A. (1988). Effect of initial soil moisture content on the vertical infiltration process—A guide to the problem of runoff-ratio and loss. Journal of Hydrology, 102(1), 267–284. https://doi.org/10.1016/0022-1694(88)90102-3

Jacobs, J. M., Myers, D. A., & Whitfield, B. M. (2003). Improved rainfall-runoff estimates using remotely sensed soil moisture. Journal of the American Water Resources Association, 39(2), 313–324. https://doi.org/10.1111/j.1752-1688.2003.tb04386.x

Köhli, M., Schrön, M., Zreda, M., Schmidt, U., Dietrich, P., & Zacharias, S. (2015). Footprint characteristics revised for field-scale soil moisture monitoring with cosmic-ray neutrons. Water Resources Research, 51(7), 5772–5790. https://doi.org/10.1002/2015WR017169

Koster, R. D., Suarez, M. J., Higgins, R. W., & Van Den Dool, H. M. (2003). Observational evidence that soil moisture variations affect precipitation. Geophysical Research Letters, 30(5), n/a-n/a. https://doi.org/10.1029/2002GL016571

Krueger, E. S., Levi, M. R., Achieng, K. O., Bolten, J. D., Carlson, J. D., Coops, N. C., Holden, Z. A., Magi, B. I., Rigden, A. J., & Ochsner, T. E. (2022). Using soil moisture information to better understand and predict wildfire danger: A review of recent developments and outstanding questions. International Journal of Wildland Fire, 32(2), 111–132. https://doi.org/10.1071/WF22056

Krueger, E. S., Ochsner, T. E., Quiring, S. M., Engle, D. M., Carlson, J. D., Twidwell, D., & Fuhlendorf, S. D. (2017). Measured Soil Moisture is a Better Predictor of Large Growing-Season Wildfires than the Keetch–Byram Drought Index. Soil Science Society of America Journal, 81(3), 490–502. https://doi.org/10.2136/sssaj2017.01.0003

Lollato, R. P., Edwards, J. T., & Ochsner, T. E. (2017). Meteorological limits to winter wheat productivity in the U.S. southern Great Plains. Field Crops Research, 203, 212–226. https://doi.org/10.1016/j.fcr.2016.12.014

Mahmood, R., Schargorodski, M., Foster, S., & Quilligan, A. (2019). A Technical Overview of the Kentucky Mesonet. Journal of Atmospheric and Oceanic Technology, 36(9), 1753–1771. https://doi.org/10.1175/JTECH-D-18-0198.1

McPherson, R. A., Fiebrich, C. A., Crawford, K. C., Kilby, J. R., Grimsley, D. L., Martinez, J. E., Basara, J. B., Illston, B. G., Morris, D. A., Kloesel, K. A., Melvin, A. D., Shrivastava, H., Wolfinbarger, J. M., Bostic, J. P., Demko, D. B., Elliott, R. L., Stadler, S. J., Carlson, J. D., & Sutherland, A. J. (2007). Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet. Journal of Atmospheric and Oceanic Technology, 24(3), 301–321. https://doi.org/10.1175/JTECH1976.1

Ochsner, T. E., Cosh, M. H., Cuenca, R. H., Dorigo, W. A., Draper, C. S., Hagimoto, Y., Kerr, Y. H., Larson, K. M., Njoku, E. G., Small, E. E., & Zreda, M. (2013). State of the Art in Large-Scale Soil Moisture Monitoring. Soil Science Society of America Journal, 77(6), 1888–1919. https://doi.org/10.2136/sssaj2013.03.0093

Palecki, M. A., Lawrimore, J. H., Leeper, R. D., Bell, J. E., Embler, S., & Casey, N. (2013). U.S. Climate Reference Network Products. https://doi.org/10.7289/V5H13007

Patrignani, A., Knapp, M., Redmond, C., & Santos, E. (2020). Technical Overview of the Kansas Mesonet. Journal of Atmospheric and Oceanic Technology, 37(12), 2167–2183. https://doi.org/10.1175/JTECH-D-19-0214.1

Patrignani, A., & Ochsner, T. E. (2018). Modeling transient soil moisture dichotomies in landscapes with intermixed land covers. Journal of Hydrology, 566, 783–794. https://doi.org/10.1016/j.jhydrol.2018.09.049

Patrignani, A., Ochsner, T. E., Feng, L., Dyer, D., & Rossini, P. R. (2022). Calibration and validation of soil water reflectometers. Vadose Zone Journal, 21(3). https://doi.org/10.1002/vzj2.20190

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), e2020GL091410. https://doi.org/10.1029/2020GL091410

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., Liu, S., & Namkhai, A. (2000). The Global Soil Moisture Data Bank. Bulletin of the American Meteorological Society, 81(6), 1281–1299. https://doi.org/10.1175/1520-0477(2000)081<1281:TGSMDB>2.3.CO;2

Ruggenthaler, R., Meißl, G., Geitner, C., Leitinger, G., Endstrasser, N., & Schöberl, F. (2016). Investigating the impact of initial soil moisture conditions on total infiltration by using an adapted double-ring infiltrometer. Hydrological Sciences Journal, 1–17. https://doi.org/10.1080/02626667.2015.1031758

Saia, S. M., Heuser, S. P., Neill, M. D., Iv, W. A. L., Mcguire, J. A., & Dello, K. D. (2023). A Technical Overview of the North Carolina ECONet. JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY, 40.

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

Schroeder, J. L., Burgett, W. S., Haynie, K. B., Sonmez, I., Skwira, G. D., Doggett, A. L., & Lipe, J. W. (2005). The West Texas Mesonet: A Technical Overview. Journal of Atmospheric and Oceanic Technology, 22(2), 211–222. https://doi.org/10.1175/JTECH-1690.1

Topp, G. C., & Reynolds, W. D. (1998). Time domain reflectometry: A seminal technique for measuring mass and energy in soil. Soil & Tillage Research, 47(1), 125–132. https://doi.org/10.1016/S0167-1987(98)00083-X

Vaz, C. M. P., Jones, S., Meding, M., & Tuller, M. (2013). Evaluation of Standard Calibration Functions for Eight Electromagnetic Soil Moisture Sensors. Vadose Zone Journal, 12(2), 1–16. https://doi.org/10.2136/vzj2012.0160

Zreda, M., Desilets, D., Ferré, T. P. A., & Scott, R. L. (2008). Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. Geophysical Research Letters, 35(21), L21402. https://doi.org/10.1029/2008GL035655

Chapter 2 - Mapping Mesoscale Rootzone Soil Moisture Using a Model-Data Fusion Approach

Joaquin Peraza[1], Andres Patrignani[1]
Department of Agronomy, Kansas State University, Manhattan, KS 66506

This chapter will be submitted to Soil Science Society of America Journal

Abstract

Mesoscale environmental monitoring networks provide accurate *in situ* soil moisture data, but the sparse distribution of stations and the small sensing footprint of conventional point-level soil moisture sensors limit the spatial representation of soil moisture conditions over large areas. The objective of this study was to develop a 250-m spatial resolution map of rootzone (0-50 cm depth) soil moisture across Kansas by merging sparse *in situ* soil moisture observations from the Kansas Mesonet with spatially continuous estimates from a simple drydown model using a data-fusion approach. The drydown model employs a lumped loss coefficient (λ) that accounts for all soil water losses and is modeled as a function of the daily vapor pressure deficit (VPD) obtained from the Kansas Mesonet. Upper and lower bounds in rootzone soil water storage were empirically determined using available soil texture and soil moisture information for the Kansas Mesonet. Soil moisture and VPD observations were obtained from 54 Kansas Mesonet stations from 1 January 2018 to 31 October 2022. Daily precipitation data was obtained from the National Weather Service multi-sensor gridded product at 4-km spatial resolution. Model predictions and *in situ* observations were assimilated using a conditional merging approach. The resulting maps were validated using a leave-one-out cross-validation approach and a series of regional surveys conducted with a roving cosmic-ray neutron probe (CRNP). The

relationship between daily VPD and λ resulted in a statistically significant negative linear correlation (r = -0.64). The leave-one-out cross-validation resulted in a median absolute error (MedAE) of 14.2 mm for the model alone, 16.2 mm for the data alone, while the data-fusion approach presented an MedAE to 11.2 mm. The spatial validation using the CRNP rover resulted in an MedAE of 16.1 mm. Combining in situ soil moisture observations with a parsimonious model using a data-fusion approach improved soil moisture estimates and seems to be a viable alternative for upscaling soil moisture observations from mesoscale *in situ* soil moisture networks.

Introduction

Soil moisture is an ecosystem variable that synthesizes the complex interactions between weather, soil, and vegetation processes (Grayson & Western, 1998; Rodriguez-Iturbe et al., 2001; Western et al., 1998). Soil moisture can be measured using a wide range of methods with different spatio-temporal scales and accuracy levels. Perhaps, the most common method for automated measurements of *in situ* soil moisture is the use of electromagnetic sensors (Evett et al., 2008; Evett & Parkin, 2005; Topp et al., 1980), which can easily interface with electronic dataloggers and data telemetry systems. Electromagnetic soil moisture sensors have been recently adopted by multiple statewide and nationwide mesoscale environmental monitoring networks including the West Texas Mesonet (Schroeder et al., 2005), the Nebraska Mesonet (Shulski et al., 2018), the New York Mesonet (Brotzge et al., 2020), the Kansas Mesonet (Patrignani et al., 2020), the U.S. Climate Reference Network (USCRN) (Palecki et al., 2013), and the Soil Climate Analysis Network (SCAN) (Schaefer et al., 2007). However, as a result of the small sensing footprint of point-level sensors and the sparse nature of weather and environmental monitoring networks, there is a need to estimate soil moisture conditions in the large spatial extents between unmonitored locations.

To fill this gap, scientists have used a wide range of tools including spatial interpolation techniques, the use of physically-based models, remote sensing soil moisture products, and model-data assimilation techniques that can consolidate multiple sources of soil moisture information into a single product. The simplest and most straight forward approach is to spatially interpolate point-level soil moisture observations across the entire spatial domain of the network (Xie et al., 2020; Yuan & Quiring, 2017). This approach can be an effective solution for spatially-dense mesoscale networks over small spatial domains (e.g., New York Mesonet) and

somewhat homogeneous landscapes, but interpolation methods alone, like the inverse distance weight and ordinary kriging, have been shown to inadequately represent soil moisture spatial patterns in areas with different land covers, sharp changes in soil type (Ochsner et al., 2019), and complex terrain that exhibits poor soil moisture spatial autocorrelation (Yao et al., 2013). Mechanistic models can circumvent some of the limitations of interpolation methods by computing the soil water balance at small grid cells and can incorporate available gridded soil and precipitation products (Abatzoglou et al., 2018). However, mechanistic models often require a large number of parameters and the need for calibration and *in situ* validation, particularly when applied across large spatial domains (e.g., large watershed, state level). Remote sensing of soil moisture like the Soil Moisture Active-Passive (SMAP) satellite mission that consists of advanced microwave radiometers and synthetic aperture radars have been designed to provide soil moisture on a global scale and represent a viable alternative to estimate soil moisture conditions between monitoring stations. The main limitation of current L-band microwave radiometers onboard orbiting satellites is the coarse (i.e., several kilometers) spatial scale and the shallow sensing depth, which is typically confined to the top few centimeters (i.e., 0-5 cm) of the soil profile (Ochsner et al., 2013; Reichle & Liu, 2020). One alternative that has demonstrated potential to leverage accurate, but sparse, observations from *in situ* networks with spatially-continuous soil moisture data are model-data fusion methods, that enable researchers to consolidate different sources of information into a single product.

Previous studies have shown that the integration of sparse observations with simple models can be effective to represent soil moisture conditions. For instance, a study in the US utilized a Kalman filter to merge sparse *in situ* measurements from USCRN and SCAN nationwide mesoscale networks with a parsimonious antecedent precipitation index model,

revealing that approximately one station every 7,000 km^2 is necessary for a significant increase in the accuracy using model-data assimilation of soil moisture data (Gruber et al., 2015). The conditional merging is a simpler algorithm than the Kalman filter, that was initially developed to improve radar precipitation estimates by assimilating *in situ* rain gauge observations (Hambali et al., 2019; Sinclair & Pegram, 2005). A recent study in South Korea used the conditional merging method to integrate *in situ* soil moisture observations with remote sensing products, finding that the conditional merging method improved accuracy in locations with a high spatial density of *in situ* observations compared to interpolation techniques like ordinary kriging, outperforming conditional merging in terms of accuracy (Kim et al., 2016). Another study in the same region, found an accuracy increase of about 20-30% when employing conditional merging to assimilate *in situ* soil moisture observations with a simple model based on the normalized vegetation difference index (NDVI) and precipitation data (Jung et al., 2017). These advancements underscore the potential of combining *in situ* measurements from mesoscale with parsimonious models for precise estimations of root zone soil water storage.

Our hypothesis is that a model-data assimilation approach that combines *in situ* soil moisture observations from a mesoscale environmental monitoring network and a parsimonious model that simulates first-order soil moisture dynamics will provide more accurate estimations of rootzone soil water storage than either using model estimates or interpolated observations alone. The objectives of this study were to: 1) develop high-spatial resolution maps of rootzone soil moisture for the state of Kansas by integrating *in situ* observations from the Kansas Mesonet and a parsimonious drydown model using a model-data fusion approach; and 2) validate the resulting maps of rootzone soil water storage.

Methodology

In situ soil moisture observations

In situ soil moisture information was obtained from the Kansas Mesonet, which is a mesoscale environmental monitoring network with 73 stations across the state of Kansas, 54 of which are equipped with research-grade soil moisture sensors (model CS655, Campbell Scientific, Inc.) at 5, 10, 20 and 50 cm depths (Patrignani et al., 2020) (Figure 2-1). Daily observations of volumetric water content (VWC) at each sensing depth were used to calculate rootzone soil water storage in the top 50 cm of the soil profile using a trapezoidal integration rule (Eq. 3) (Gao et al., 2019; Nachabe et al., 2004; Parker & Patrignani, 2021):

$$S_t = \theta_{1,t}\, Z_{1,t} + \sum_{i=2}^{n} \frac{\theta_{i-1,t} + \theta_{1,t}}{2} \left(Z_{i,t} - Z_{i-1,t}\right) \quad (1)$$

where S_t is the soil water storage, t is time in days, Z_i (mm) is the depth of the ith sensor, θ_i is the volumetric water content of the ith sensor, and n is the total number of sensors in the soil profile. The soil moisture from the CS655 water reflectometers was computed using the following equation proposed by Kargas & Soulis (2019):

$$\theta = a + b\sqrt{K_a} + cEC \quad (2)$$

where θ represents the volumetric water content, K_a denotes the bulk relative permittivity, EC stands for the bulk electrical conductivity, a, b, and c are fitting parameters with respective values of a = -0.115, b = 0.0989, and c = -0.0572 found in a previous laboratory calibration study (Patrignani et al. 2022).

Soil water storage model

Soil moisture dynamics in the top 50 cm of the soil profile were simulated using a modified antecedent precipitation index model that represents soil moisture as a sequence of exponentially decaying drydown events following the occurrence of precipitation (Kurc & Small, 2004; Raoult et al., 2021; Shellito et al., 2016). Additionally, the soil water storage oscillates between a constrained interval determined by an upper limit and a lower limit. Different soil-plant-atmosphere attributes define those limits (Reynolds et al., 2000; Ritchie, 1981). Rootzone soil water storage was simulated using a modified antecedent precipitation index model (e.g., Kurc & Small, 2004):

$$S_t = \begin{cases} (S_{t-1} - S_{LL})\,\lambda_t + S_{LL} + P_t, & S_t \leq S_{UL} \\ S_{UL}, & S_t > S_{UL} \end{cases} \quad (3)$$

where S_t (mm) is the soil water storage in the rootzone (0 to 50 cm depth) on day t, S_{t-1} (mm) is the soil water storage on the previous day, λ_t (unitless) is a lumped loss coefficient representing all losses in the soil water balance, P_t (mm) is daily precipitation, S_{LL} (mm) is the lower limit of soil water storge, and S_{UL} (mm) is the maximum limit of soil water storage.

The S_{LL} and S_{UL} were estimated based on the sand content in the top 50 cm of the soil profile (Van Den Berg et al., 1997). Data for the area of study were obtained from the SoilGrids gridded soil product at 250 m spatial resolution (Hengl et al., 2017) for the entire region. The SoilGrids dataset was generated using a database of 1,284 *in situ* soil samples collected in the state of Kansas by the USDA Natural Resources Conservation Service and state-of-the-art machine learning techniques. The soil texture for the top 50 cm was calculated using a weighted average and a linear model was adjusted using soil texture information for 54 Kansas Mesonet locations, in conjunction with S_{LL} and S_{UL} calculated for each station by the extraction of the five per cent of the lowest and uppermost soil water storage observations. Daily precipitation

estimates were downloaded from the National Weather Service (NWS, 2021) at 4-km spatial resolution and then linearly downscaled to match with the 250-m spatial grid.

Drydown events are periods without measurable precipitation that are characterized by an exponential decrease in soil moisture. This decrease in soil moisture is mostly attributed to evapotranspiration and drainage when the soil moisture is at or near saturation conditions. Drydown events were extracted from rootzone soil moisture data obtained from 54 Kansas Mesonet stations equipped with soil moisture sensors from January 2017 to December 2021. The dataset of resulting drydown periods was filtered to only retain drydowns that had >7 days of duration without precipitations to ensure that drydown periods were well defined and exhibited the typical exponential decay shape (Figure 2-2). Each drydown was fitted to individual exponential decays using the ordinary least squares method. For prediction of the water decay magnitude λ, the atmospheric vapor pressure deficit (VPD) observed during each drydown, was tested as an easy-to-obtain indicator of the atmospheric water demand. Parameter λ was estimated for each drydown, and VPD was averaged for the temporal window covered by each of the fitted drydown events. A linear model was fitted between parameter λ and VPD using a robust linear fitting routine with outlier weighting implemented in the *statsmodels* (Seabold & Perktold, 2010) Python library, allowing to determine the water storage loss in a daily basis as represented in Eq. 3. Additionally, reference evapotranspiration (ETo) for a short grass was computed using the Penman-Monteith method as described in the FAO Irrigation and drainage paper 56 (Allen et al., 1998), and tested as an alternative predictor instead of VPD. Vegetation data was also tested as additional input to predict the magnitude of soil water depletion using MODIS Enhanced Vegetation Index (EVI) pixel values for each station, linearly interpolated to a daily time scale (Didan, 2021).

Data Assimilation

A conditional merging algorithm was implemented (Sinclair & Pegram, 2005) to assimilate daily observations of rootzone soil moisture obtained from the Kansas Mesonet and daily soil water storage predicted with the drydown model. The data-assimilation approach consisted of the following steps: 1) we first simulated soil moisture for each 250-m pixel of the grid applying the drydown model using gridded versions of fixed and daily parameters, 2) then, we calculated the soil moisture residuals between the drydown model and *in situ* observations at the stations of the Kansas Mesonet, 3) then, we interpolated the residuals across the entire grid using universal kriging, and 4) we corrected model predictions by adding the interpolated residuals to the initial map. These four steps resulted in the final map representing spatially continuous soil moisture across Kansas.

A grid with a spatial resolution of 250 m was defined to address the main goal of obtaining a daily map of soil water storage, representing the conditions across the state in a spatially continuous approach. This resulted in a gridded representation of 1,284 rows by 3,210 columns to cover the entire state of Kansas. The soil water storage model, the daily *in situ* soil moisture data request from the Kansas Mesonet, and the data-fusion procedure were implemented in the Python programming language version 3.8, so that the approach could be deployed in a server to automatically generate daily maps of soil moisture operationally.

Validation of soil moisture maps

The performance of the model-data fusion approach was validated by calculating the MedAE (Median Absolute Error) for the entire year 2022 observations in a leave-one-out cross-

validation routine, comparing the impact of model-data fusion for different stations. Each cross-validation iteration differed from the final model in two aspects to make this evaluation more robust, 1) the evaluated station was not included in the conditional merging computation, and 2) the saved output did not include the model-data fusion correction, allowing the accumulation of biases in the recursive chain of daily computations. The data fusion performance was measured as the difference between MedAE of the model with conditional merging using neighbor stations, and the MedAE using only the model output in 54 Kansas Mesonet locations with rootzone soil moisture data during year 2022. Additionally, the error of estimating the soil water storage by interpolating the values at neighbor stations was measured during cross-validation.

Soil moisture maps generated using the model-data fusion approach assimilating soil moisture information from the Kansas Mesonet were validated using a cosmic-ray neutron probe rover. Raw neutron counts were corrected for the effect of atmospheric pressure (Zreda et al., 2012), humidity (Rosolem et al., 2013), and incoming neutron flux using the CRNPy Python library. Corrected neutron counts were converted into volumetric water content (θ) using the following equation described in Desilets et al. (2010):

$$\theta = \left(\frac{a_0}{\frac{N_{corr}}{N_0} - a_1} - a_2 - w_L \right) \rho_b \qquad (4)$$

where N_0 is an instrument specific parameter that represents the corrected neutron counts on dry soil and is often obtained by conducting a field calibration, N_{corr} are the corrected neutron counts, w_L is the soil lattice water, ρ_b is the soil bulk density, and parameters a_0, a_1 and a_2 are fitting parameters that usually have the following values: a_0=0.0808, a_1=0.372 , a_2 =0.115 (Desilets et al., 2010).

The 29 CRNP rover transects collected across a 9 squared kilometers area located within Gypsum, KS (38.706, -97.428, 370 m a.s.l.) between 31 July 2017 and 07 May 2018 covered a wide range of land covers including cropland and grassland. Each transect was processed using the Python CRNPy library, obtaining the surface volumetric water content and top 50 cm soil water storage using the library built-in exponential filter.

The accuracy of soil water storage estimations was evaluated using the Median Absolute Error (MedAE). In Eq. 5, S_n represents the observed soil water storage for the i_{th} data point, and $\hat{S}_n$ is the map estimated soil water storage for the same point, over n total observations:

$$\text{MedAE} = \text{median}\left(\left|S_1 - \hat{S}_1\right|, \left|S_2 - \hat{S}_2\right|, \ldots, \left|S_n - \hat{S}_n\right|\right) \quad (5)$$

Results and discussion

Model-Data fusion cross-validation

The implemented model-data fusion algorithm using the conditional merging technique, successfully estimated the soil water storage in the top 50 cm. The MedAE obtained in the leave-one-out cross-validation was 11.2 mm. This study revealed that using a model data fusion approach combining accurate *in situ* observations can decrease the error, improving the estimations produced by either the model alone (MedAE = 14.2 mm) or the *in situ* data interpolated (MedAE = 16.2 mm).

In the cross-validation routine, the model results without data assimilation presented a Median Absolute Error (MedAE) of 14.2 mm for the entire state, this represents approximately 6% of the soil porosity and 21% of the average difference between the observed soil water storage upper and lower limits. Only four stations out of the 54 cross validated, presented MedAE higher than 25 mm when comparing the model output to the *in situ* observations (Figure 2-4A). These results are comparable to an evaluation of different spatial models conducted in Oklahoma, where land surface models Noah, Mosaic, SAC and VIC where compared to the Oklahoma Mesonet stations, yielding a relative MAE of 15.9%, 20.6% ,38.4% and 11.9% respectively when considering the top 40 cm of the soil profile (Xia et al., 2014).

When validating the use of interpolated *in situ* observations of soil water storage from the Kansas Mesonet, the leave-one-out resulted in a MedAE of 16.2 mm. Different error spatial distributions were observed within the state, where 11 locations of the Kansas Mesonet showed a MedAE higher than 25 mm and 13 locations a MedAE of less than 10 mm (Figure 2-4B). A study in Oklahoma reported similar results when testing three interpolation techniques (reduced optimal interpolation, inverse distance weighting, and Cokriging) to estimate soil moisture

values in the top meter using *in situ* observations using 65 stations from the Oklahoma Mesonet for interpolation, and 39 stations for validation, finding MAE values between 10 and 15 mm (Yuan & Quiring, 2017). Recently, using the same Oklahoma Mesonet network, Ochsner et al. (2019) performed a cross-validation routine comparing Barnes Objective Analysis (BOA), Ordinary kriging (OK), and regression kriging (RK) interpolation methods in soil moisture at different depths. The leave-one-out cross-validation routine was tested in 120 stations over a period of 2 years, reporting an MAE for the top 60 cm between 0.0576 and 0.0715 $m^3\ m^{-3}$, roughly representing between 34 and 43 mm of soil water storage.

When analyzing the performance of the model-data fusion approach at a station-level, all the sites presented an improvement, measured as a 27% and 47% decrease of MedAE, compared to model output and *in situ* data interpolation respectively (Figure 2-4C). The model-data fusion estimates were better than the neighbor stations interpolation in 89% of the evaluated sites (Figure 2-5). The proposed conditional merging technique has been widely used to assimilate rain gauge observations with radar precipitation estimates (Hambali et al., 2019; Sinclair & Pegram, 2005), but research using soil moisture is limited. Two studies in South Korea tested the use of the conditional merging technique with *in situ* soil moisture sensor observations at 10 cm depth from the Rural Development Administration network and remote sensing products of soil moisture. Kim et al. (2016) analyzed the improvement with a leave-one-out cross-validation comparing *in situ* observations with the soil moisture estimated by the Advanced Microwave Scanning Radiometer 2 sensor. The results revealed that the conditional merging method increased the accuracy of the estimations in most of the cases, however in locations with high spatial density of *in situ* observations interpolation methods as ordinary Kriging alone outperforms the conditional merging accuracy. Jung et al. (2017) reported that the conditional

merging technique increased the accuracy by about 20-30% when correcting soil moisture estimates from a linear model based on MODIS LST, NDVI, and precipitation data, with *in situ* observations. Similarly, a study in the United States tested the use of a Kalman filter for correcting soil moisture estimations from an antecedent precipitation index (API) model with *in situ* observations from USCRN and SCAN networks, showing that current nationwide networks density is not suitable to increase the spatial accuracy of API model estimates to overcome the accuracy of remote sensing products like Advanced SCATterometer (ASCAT) soil moisture estimates. The authors suggested that a density of one station every 7000 km^2 is required to produce a substantial increase in the soil moisture estimation accuracy when assimilating *in situ* observations (Gruber et al., 2018), Aligned with this, the Kansas Mesonet features a station density approximately double that recommendation.

Spatial validation using a CRNP rover

The map successfully estimated soil water storage for the validation area in Gypsum, KS over the whole measured period. The agreement between the model-data fusion product and the average soil water storage reported by each CRNP transect resulted in a MedAE of 16.1 mm, this represents 12% of the soil water storage capacity for this location. The error in this validation is comparable to findings of a study in Oklahoma that reported approximately 15% of relative error when comparing cropland SMAP estimates vs *in situ* observations (Wyatt et al., 2021).

The transects contained between 148 to 192 one-minute observations, with an average corrected neutron count of 351 counts per minute (cpm) ranging between 330 to 370 cpm. The model captured the increase in soil water storage after precipitation events, and the posterior drydown events, however, due to the distance to the Gypsum station of the Kansas Mesonet, the

conditional merging is highly influenced by *in situ* observations. During late 2017 there was a disagreement between the map and the CRNP rover transects (Figure 2-6). The bias could be due to the different land covers between the Kansas Mesonet station located in a grassland area, and the high amount of cropland areas, of mostly winter wheat (*Triticum aestivum* L.), scouted by the CRNP rover. Previous studies have reported similar differences between permanent soil moisture sensors located in grass-covered buffer areas and adjacent agricultural fields (Han et al., 2012; Heathman et al., 2012).

Daily soil water loss

There was a significant negative linear relationship (r = -0.55, P<0.001) between λ and VPD when analyzing a total of 485 soil moisture drydowns obtained from 54 stations of the Kansas Mesonet. (Figure 2-7). The linear model between the VPD and the daily soil moisture presented a negative Pearson's correlation coefficient (r = -0.55). In our pursuit to refine the model, grassland reference evapotranspiration (ETo) was assessed as a potential alternative to using VPD. Interestingly, evapotranspiration exhibited a similar correlation pattern with an r value of -0.53. These findings underscore that both VPD and evapotranspiration offer analogous insights to estimate the soil moisture drydown dynamics. Yet, when incorporating vegetation data through EVI into a multiple linear model, its influence on estimating λ was found to be insignificant. This is substantiated by the 95% confidence interval of the EVI coefficient, which lies between -0.0034 and 0.00028, indicating a negligible impact. However, it is essential to note that the scope of this study was limited to the predominant warm-season grasses at the Mesonet sites, which could present similar growing dynamics, potentially introducing high autocorrelations between sites masking the effect of vegetation. Future research should

encompass a broader range of land covers to offer a comprehensive understanding. For instance, past studies found vegetation differences to be significative and bridged the gap between crop field measurements and data from permanent sensors in native grasses through a Cumulative Distribution Function (CDF) matching technique (Han et al., 2012; Heathman et al., 2012). Furthermore, in Oklahoma an artificial neural network has been used to upscale mesoscale soil moisture observations from a mesoscale network situated predominantly in warm-season grasses to estimate winter wheat soil moisture (Patrignani & Ochsner, 2018).

Soil properties

With the extracted soil texture information from SoilGrids for each station of the Kansas Mesonet and the respective historic soil water storage observations, a linear model was adjusted. The obtained inverse linear relationships (Figure 2-8) agree with similar studies in the same region comparing soil texture with field capacity and permanent wilting point presenting a similar range of coefficients when comparing the values in terms of volumetric water content (Parker et al., 2022). Soil water storage limits for the complete area of study were obtained by applying the linear models to the SoilGrids weighted average spatial layer for the top 50 cm at a 250 m resolution (Figure 2-9). In locations where data was available, texture information from SoilGrids was substituted with *in situ* data from Parker et al. (2022).

Limitations of this study

One of the limitations of this study is that the proposed approach only accounts for the top 50 cm of the soil profile, which may impose a limitation to accurately represent rootzone soil moisture conditions in some deep rooted crops and pastures. For instance, in the U.S. Great

Plains, the root system of winter wheat can reach a depth of 90 cm by the middle of the growing season (Weaver et al., 1924), summer crops like grain sorghum and suflowers can reach root depths between two to three meters depth (Stone et al., 2002), and perennial pastures commonly have root system reaching more than one meter depth (DuPont et al., 2014). Another limitation of the proposed soil water storage model is the uncertainty in the computation of the lumped loss coefficient, particularly at greater VPD values, which could be attributed by errors in the gridded soil products, uncertainties in the computation of the daily VPD, and uncertainties in the estimation of the loss coefficient caused by drydown events that do not follow the typical exponential decay pattern.

Conclusions

- The implemented simple model successfully estimated the soil water storage for the top 50 cm, having a MedAE of 14.2 mm in the leave one out cross validation using the *in situ* Kansas Mesonet observations.
- The data assimilation between the model and the *in situ* observations using a conditional merging algorithm improved the soil water storage estimations by 2 mm, being better than the simple model or the interpolated *in situ* observations alone.
- The validation using CRNP had a MedAE of 16.1 mm, however it allowed to capture the model's limitations when estimating soil water storage in cropland areas, since the model calibration, and data assimilation routines, are based on the Kansas Mesonet *in situ* observations, which are located in grassland areas. There is potential to explore solutions, such as the use of observation operators, to upscale the model-data fusion estimates to cropland soil moisture.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Scientific Data, 5(1), 1–12. https://doi.org/10.1038/sdata.2017.191

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., & others. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao, Rome, 300(9), D05109.

Brotzge, J. A., Wang, J., Thorncroft, C. D., Joseph, E., Bain, N., Bassill, N., Farruggio, N., Freedman, J. M., Hemker, K., Johnston, D., Kane, E., McKim, S., Miller, S. D., Minder, J. R., Naple, P., Perez, S., Schwab, J. J., Schwab, M. J., & Sicker, J. (2020). A Technical Overview of the New York State Mesonet Standard Network. Journal of Atmospheric and Oceanic Technology, 37(10), 1827–1845. https://doi.org/10.1175/JTECH-D-19-0220.1

Desilets, D., Zreda, M., & Ferré, T. P. A. (2010). Nature's neutron probe: Land surface hydrology at an elusive scale with cosmic rays: Nature's Neutron Probe. Water Resources Research, 46(11). https://doi.org/10.1029/2009WR008726

Didan, K. (2021). MODIS/Terra Vegetation Indices 16-Day L3 Global 500m SIN Grid V061. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD13A1.061.

DuPont, S. T., Beniston, J., Glover, J. D., Hodson, A., Culman, S. W., Lal, R., & Ferris, H. (2014). Root traits and soil properties in harvested perennial grassland, annual wheat, and never-tilled annual wheat. Plant and Soil, 381(1–2), 405–420. https://doi.org/10.1007/s11104-014-2145-2

Evett, S. R., Heng, L. K., Moutonnet, P., & Nguyen, M. (2008). Field estimation of soil water content: A practical guide to methods, instrumentation, and sensor technology. IAEA: Vienna.

Evett, Steven. R., & Parkin, G. W. (2005). Advances in Soil Water Content Sensing: The Continuing Maturation of Technology and Theory. Vadose Zone Journal, 4(4), 986–991. https://doi.org/10.2136/vzj2005.0099

Gao, X., Zhao, X., Brocca, L., Pan, D., & Wu, P. (2019). Testing of observation operators designed to estimate profile soil moisture from surface measurements. Hydrological Processes, 33(4), 575–584. https://doi.org/10.1002/hyp.13344

Grayson, R. B., & Western, A. W. (1998). Towards area estimation of soil water content from point measurements: Time and space stability of mean response. https://doi.org/10.1016/S0022-1694(98)00096-1

Gruber, A., Crow, W., Dorigo, W., & Wagner, W. (2015). The potential of 2D Kalman filtering for soil moisture data assimilation. Remote Sensing of Environment, 171, 137–148. https://doi.org/10.1016/j.rse.2015.10.019

Gruber, A., Crow, W. T., & Dorigo, W. A. (2018). Assimilation of Spatially Sparse In Situ Soil Moisture Networks into a Continuous Model Domain. Water Resources Research, 54(2), 1353–1367. https://doi.org/10.1002/2017WR021277

Hambali, R., Legono, D., Jayadi, R., Oishi, S., Department of Civil and Environmental Engineering, Universitas Gadjah Mada Jl. Grafika No.2 Bulaksumur, Yogyakarta 55281, Indonesia, & Research Center for Urban Safety and Security, Kobe University, Hyogo, Japan. (2019). Improving Spatial Rainfall Estimates at Mt. Merapi Area Using Radar-

Rain Gauge Conditional Merging. Journal of Disaster Research, 14(1), 69–79. https://doi.org/10.20965/jdr.2019.p0069

Han, E., Heathman, G. C., Merwade, V., & Cosh, M. H. (2012). Application of observation operators for field scale soil moisture averages and variances in agricultural landscapes. Journal of Hydrology, 444, 34–50. https://doi.org/10.1016/j.jhydrol.2012.03.035

Heathman, G. C., Cosh, M. H., Merwade, V., & Han, E. (2012). Multi-scale temporal stability analysis of surface and subsurface soil moisture within the Upper Cedar Creek Watershed, Indiana. Catena, 95, 91–103. https://doi.org/10.1016/j.catena.2012.03.008

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. PLOS ONE, 12(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Jung, C., Lee, Y., Cho, Y., & Kim, S. (2017). A Study of Spatial Soil Moisture Estimation Using a Multiple Linear Regression Model and MODIS Land Surface Temperature Data Corrected by Conditional Merging. Remote Sensing, 9(8), 870. https://doi.org/10.3390/rs9080870

Kargas, G., & Soulis, K. X. (2019). Performance evaluation of a recently developed soil water content, dielectric permittivity, and bulk electrical conductivity electromagnetic sensor. Agricultural Water Management, 213, 568–579. https://doi.org/10.1016/j.agwat.2018.11.002

Keables, M. J. (2010). A SOIL WATER CLIMATOLOGY FOR KANSAS. Great Plains Research, 20(2).

Kim, D., Lee, J., Kim, H., & Choi, M. (2016). Spatial composition of AMSR2 soil moisture products by conditional merging technique with ground soil moisture data. Stochastic Environmental Research and Risk Assessment, 30(8), 2109–2126. https://doi.org/10.1007/s00477-016-1300-0

Kurc, S. A., & Small, E. E. (2004). Dynamics of evapotranspiration in semiarid grassland and shrubland ecosystems during the summer monsoon season, central New Mexico. Water Resources Research, 40(9). https://doi.org/10.1029/2004WR003068

Nachabe, M., Masek, C., & Obeysekera, J. (2004). Observations and Modeling of Profile Soil Water Storage above a Shallow Water Table. Soil Science Society of America Journal, 68(3), 719–724. https://doi.org/10.2136/sssaj2004.7190

NWS. (2021). National Weather Service—Advanced Hydrologic Prediction Service. National Weather Service - Advanced Hydrologic Prediction Service. https://water.weather.gov/precip/index.php

Ochsner, T. E., Cosh, M. H., Cuenca, R. H., Dorigo, W. A., Draper, C. S., Hagimoto, Y., Kerr, Y. H., Larson, K. M., Njoku, E. G., Small, E. E., & Zreda, M. (2013). State of the Art in Large-Scale Soil Moisture Monitoring. Soil Science Society of America Journal, 77(6), 1888–1919. https://doi.org/10.2136/sssaj2013.03.0093

Ochsner, T. E., Linde, E., Haffner, M., & Dong, J. (2019). Mesoscale Soil Moisture Patterns Revealed Using a Sparse In Situ Network and Regression Kriging. Water Resources Research, 55(6), 4785–4800. https://doi.org/10.1029/2018WR024535

Palecki, M. A., Lawrimore, J. H., Leeper, R. D., Bell, J. E., Embler, S., & Casey, N. (2013). U.S. Climate Reference Network Products. https://doi.org/10.7289/V5H13007

Parker, N., Kluitenberg, G., Redmond, C., & Patrignani, A. (2022). A database of soil physical properties for the Kansas mesonet. Soil Science Society of America Journal, saj2.20465. https://doi.org/10.1002/saj2.20465

Parker, N., & Patrignani, A. (2021). Reconstructing Precipitation Events Using Collocated Soil Moisture Information. Journal of Hydrometeorology, 22(12), 3275–3290. https://doi.org/10.1175/JHM-D-21-0168.1

Patrignani, A., Knapp, M., Redmond, C., & Santos, E. (2020). Technical Overview of the Kansas Mesonet. Journal of Atmospheric and Oceanic Technology, 37(12), 2167–2183. https://doi.org/10.1175/JTECH-D-19-0214.1

Patrignani, A., Ochsner, T. E., Feng, L., Dyer, D., & Rossini, P. R. (2022). Calibration and validation of soil water reflectometers. Vadose Zone Journal, 21(3). https://doi.org/10.1002/vzj2.20190

Raoult, N., Ottlé, C., Peylin, P., Bastrikov, V., & Maugis, P. (2021). Evaluating and Optimizing Surface Soil Moisture Drydowns in the ORCHIDEE Land Surface Model at In Situ Locations. Journal of Hydrometeorology, 22(4), 1025–1043. https://doi.org/10.1175/JHM-D-20-0115.1

Reichle, R. ,. G. De Lannoy, R. D. Koster, W. T. Crow, J. S. Kimball, & Liu, Q. (2020). SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Analysis Update, Version 5. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/0D8JT6S27BS9

Reitz, M., Sanford, W. E., Senay, G. B., & Cazenas, J. (2017). Annual Estimates of Recharge, Quick-Flow Runoff, and Evapotranspiration for the Contiguous U.S. Using Empirical Regression Equations. JAWRA Journal of the American Water Resources Association, 53(4), 961–983. https://doi.org/10.1111/1752-1688.12546

Reynolds, C. A., Jackson, T. J., & Rawls, W. J. (2000). Estimating soil water-holding capacities by linking the Food and Agriculture Organization Soil map of the world with global pedon databases and continuous pedotransfer functions. Water Resources Research, 36(12), 3653–3662. https://doi.org/10.1029/2000WR900130

Ritchie, J. T. (1981). Soil water availability. Plant and Soil, 58(1–3), 327–338. https://doi.org/10.1007/BF02180061

Rodriguez-Iturbe, I., Porporato, A., Laio, F., & Ridol, L. (2001). Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress I. Scope and general outline. Advances in Water Resources, 11.

Rosolem, R., Shuttleworth, W. J., Zreda, M., Franz, T. E., Zeng, X., & Kurc, S. A. (2013). The Effect of Atmospheric Water Vapor on Neutron Count in the Cosmic-Ray Soil Moisture Observing System. Journal of Hydrometeorology, 14(5), 1659–1671. https://doi.org/10.1175/JHM-D-12-0120.1

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

Schroeder, J. L., Burgett, W. S., Haynie, K. B., Sonmez, I., Skwira, G. D., Doggett, A. L., & Lipe, J. W. (2005). The West Texas Mesonet: A Technical Overview. Journal of

Atmospheric and Oceanic Technology, 22(2), 211–222. https://doi.org/10.1175/JTECH-1690.1

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. 9th Python in Science Conference.

Shellito, P. J., Small, E. E., Colliander, A., Bindlish, R., Cosh, M. H., Berg, A. A., Bosch, D. D., Caldwell, T. G., Goodrich, D. C., McNairn, H., Prueger, J. H., Starks, P. J., van der Velde, R., & Walker, J. P. (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

Sinclair, S., & Pegram, G. (2005). Combining radar and rain gauge rainfall estimates using conditional merging. Atmospheric Science Letters, 6(1), 19–22. https://doi.org/10.1002/asl.85

Stone, L. R., Goodrum, D. E., Schlegel, A. J., Jaafar, M. N., & Khan, A. H. (2002). Water Depletion Depth of Grain Sorghum and Sunflower in the Central High Plains. Agronomy Journal, 94(4), 936–943. https://doi.org/10.2134/agronj2002.9360

Topp, G. C., Davis, J. L., & Annan, A. P. (1980). Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. Water Resources Research, 16(3), 574–582. https://doi.org/10.1029/WR016i003p00574

Van Den Berg, M., Klamt, E., Van Reeuwijk, L. P., & Sombroek, W. G. (1997). Pedotransfer functions for the estimation of moisture retention characteristics of Ferralsols and related soils. Geoderma, 78(3–4), 161–180. https://doi.org/10.1016/S0016-7061(97)00045-1

Weaver, J. E., Kramer, J., & Reed, M. (1924). Development of Root and Shoot of Winter Wheat Under Field Environment. Ecology, 5(1), 26–50. https://doi.org/10.2307/1929162

Western, A. W., Blöschl, G., & Grayson, R. B. (1998). Geostatistical characterisation of soil moisture patterns in the Tarrawarra catchment. Journal of Hydrology, 205(1–2), 20–37. https://doi.org/10.1016/S0022-1694(97)00142-X

Wyatt, B. M., Ochsner, T. E., & Zou, C. B. (2021). Estimating root zone soil moisture across diverse land cover types by integrating in-situ and remotely sensed data. Agricultural and Forest Meteorology, 307, 108471. https://doi.org/10.1016/j.agrformet.2021.108471

Xia, Y., Sheffield, J., Ek, M. B., Dong, J., Chaney, N., Wei, H., Meng, J., & Wood, E. F. (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

Xie, B., Jia, X., Qin, Z., Zhao, C., & Shao, M. (2020). Comparison of interpolation methods for soil moisture prediction on China's Loess Plateau. Vadose Zone Journal, 19(1), e20025. https://doi.org/10.1002/vzj2.20025

Yao, X., Fu, B., Lü, Y., Sun, F., Wang, S., & Liu, M. (2013). Comparison of Four Spatial Interpolation Methods for Estimating Soil Moisture in a Complex Terrain Catchment. PLoS ONE, 8(1), e54660. https://doi.org/10.1371/journal.pone.0054660

Yuan, S., & Quiring, S. M. (2017). Comparison of three methods of interpolating soil moisture in Oklahoma. International Journal of Climatology, 37(2), 987–997. https://doi.org/10.1002/joc.4754

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., & Rosolem, R. (2012). COSMOS: The COsmic-ray Soil Moisture Observing System. Hydrology and Earth System Sciences, 16(11), 4079–4099. https://doi.org/10.5194/hess-16-4079-2012

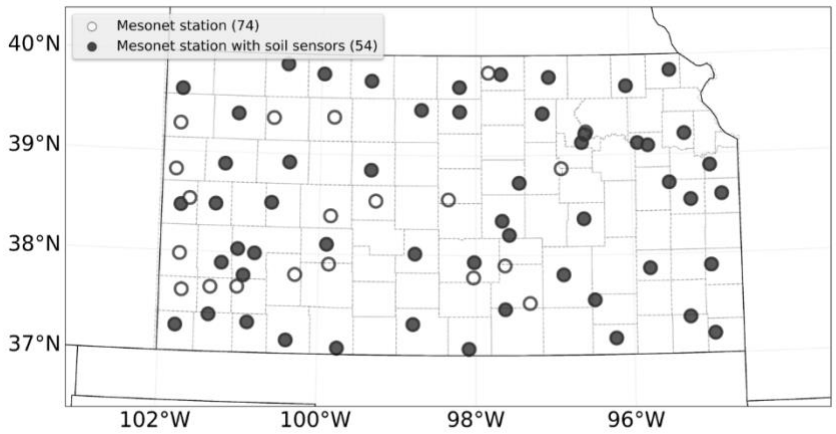


Figure 2-1. Map showing the 54 out of 74 stations of the Kansas Mesonet that are equipped with research-grade soil moisture sensors at 5, 10, 20 and 50 cm depth as of September 2023,

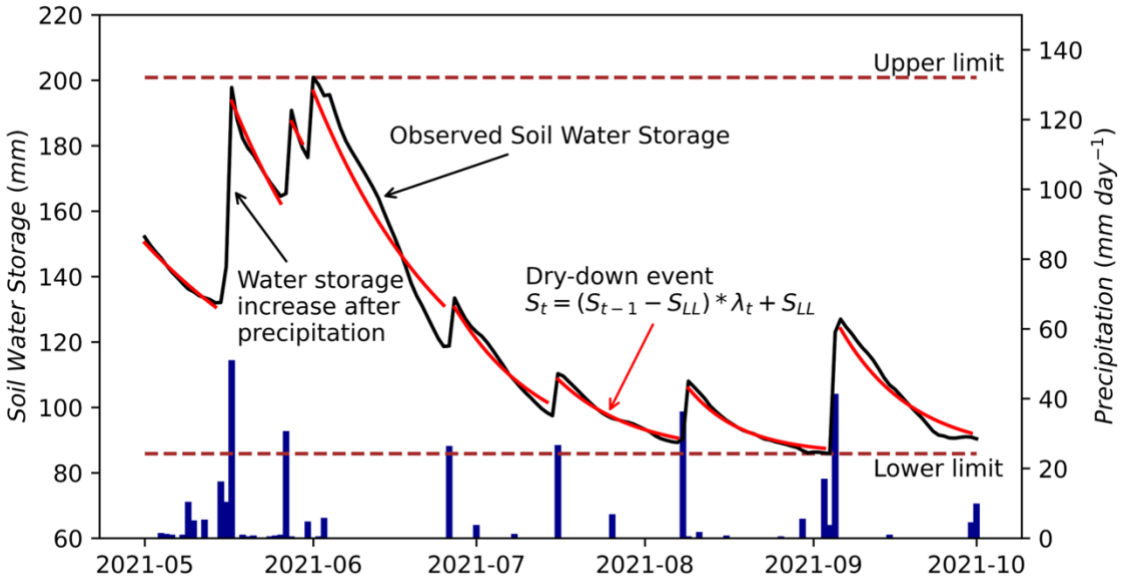


Figure 2-2. Example of soil water storage dynamics observed at the Gypsum station of the Kansas Mesonet (black line) and the corresponding soil moisture drydowns fitted using the simple model used in this study. The upper and lower limits represent the maximum and minimum soil water storage of the model.

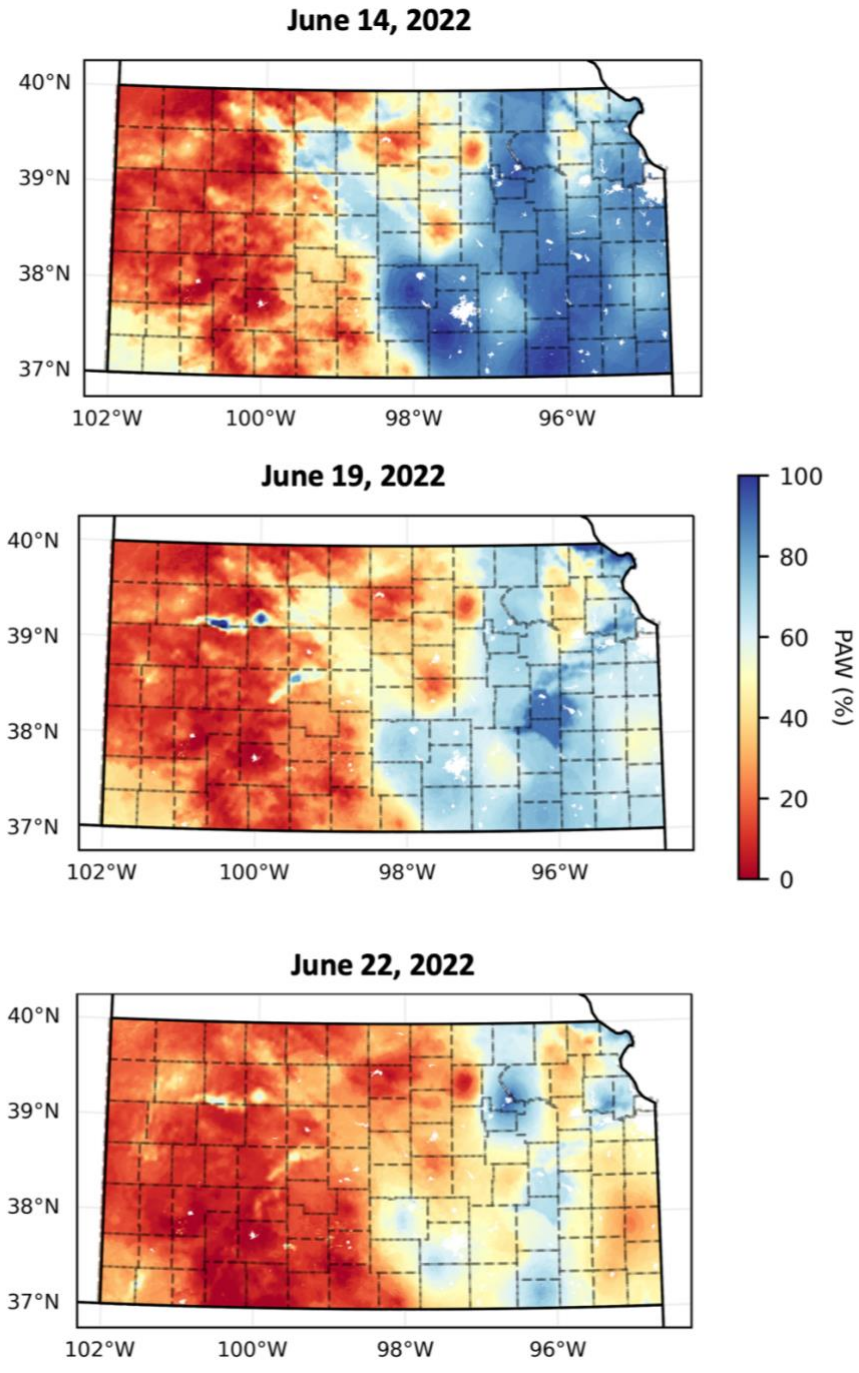


Figure 2-3. Temporal progression of plant available water (PAW) for an 8-day period. The rapid decrease in soil moisture during short periods shows the importance of high-temporal resolution monitoring to develop flash drought warning systems.

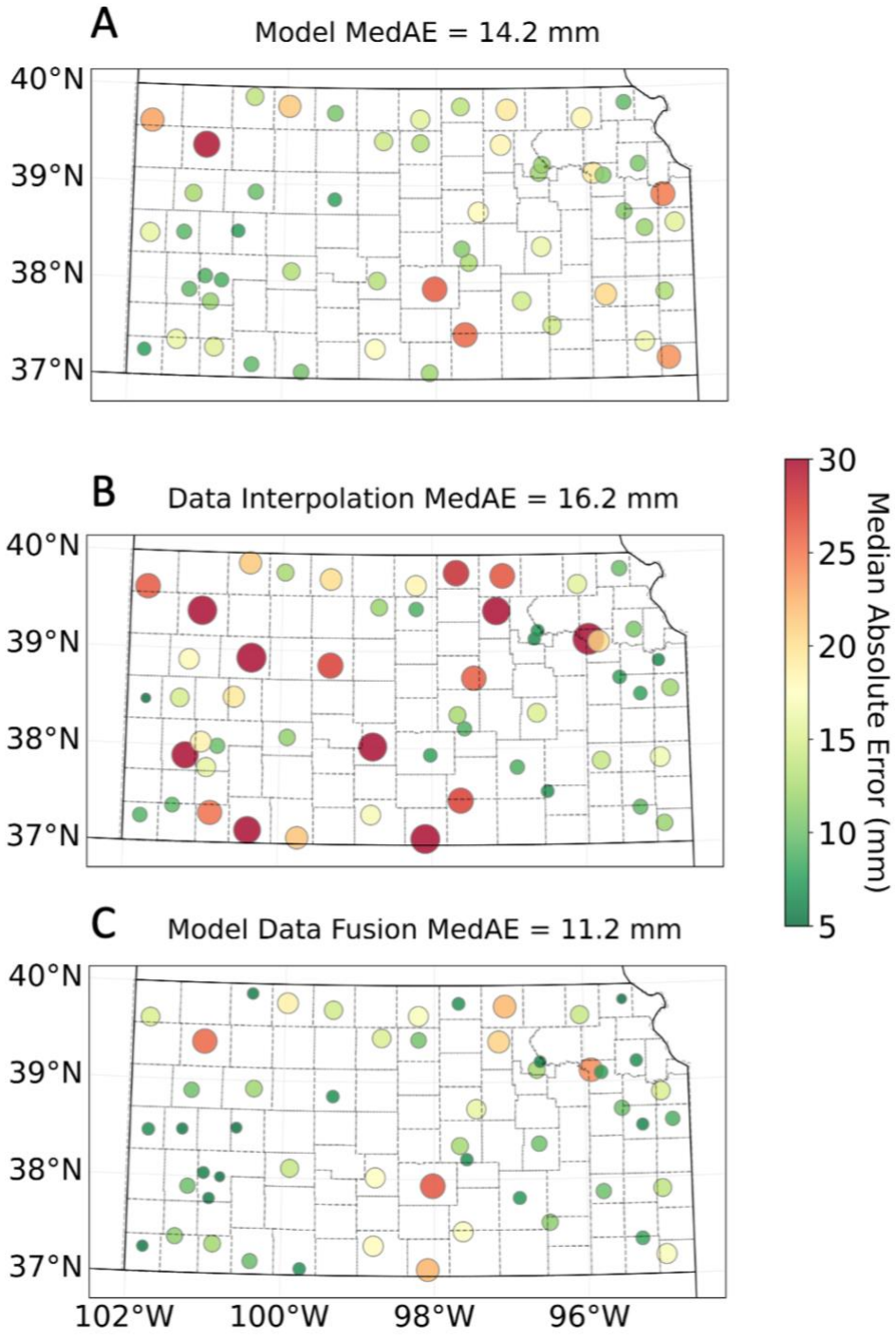


Figure 2-4. Median absolute error (MAE) of the soil water storage model (A), interpolated *in situ* observations (B), and the proposed model-data fusion approach (B) based on a leave-one-out cross-validation approach.

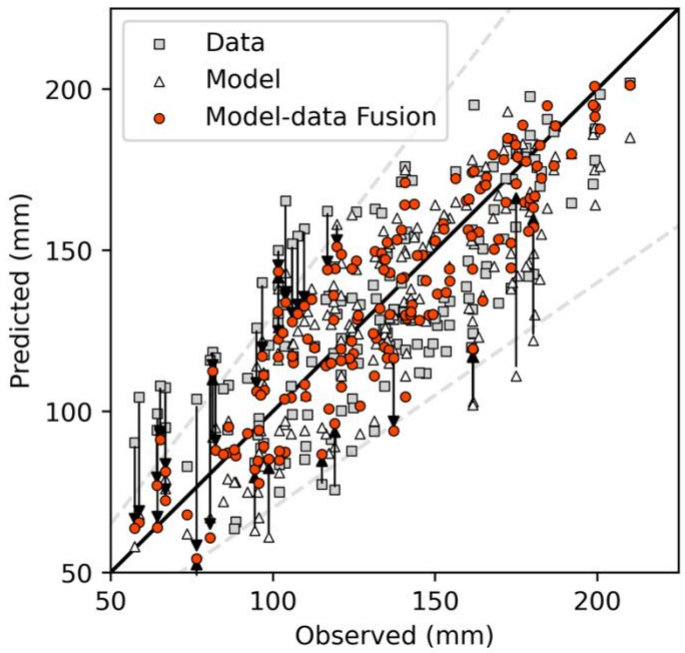


Figure 2-5. Results of the leave-one-out cross-validation exercise using 5% randomly selected points. Arrows depict that for large errors (i.e., relative error >30%), the model-data fusion reduced the uncertainty of model estimations and performed better than interpolating neighboring *in situ* observations.

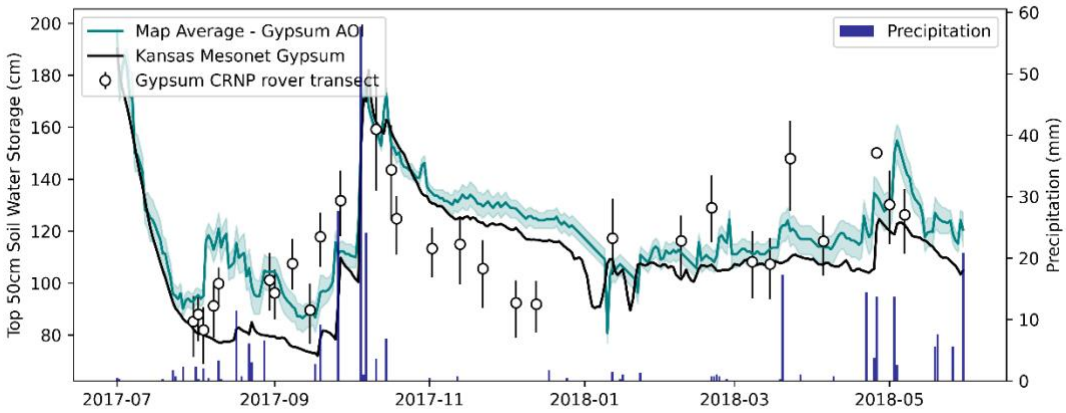


Figure 2-6. Time series of daily soil water storage in the top 50 cm obtained from the Gypsum station of the Kansas Mesonet, estimated from rover surveys, and with the data-assimilation method. Shaded area of the soil moisture time series from the data-assimilation method represents one standard deviation. Daily precipitation was obtained from the Gypsum station of the Kansas Mesonet.

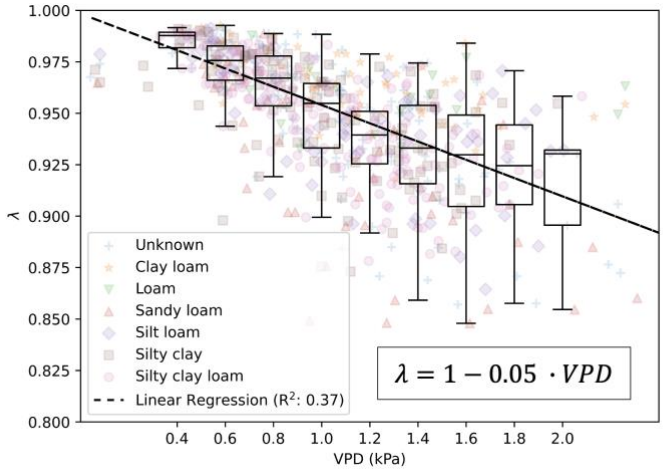


Figure 2-7. Vapor pressure deficit (VPD) negative linear relationship with the soil water storage loss coefficient (λ) for the 485 drydown events analyzed. Different soil textural classes represented with different markers were not related with the λ parameter value.

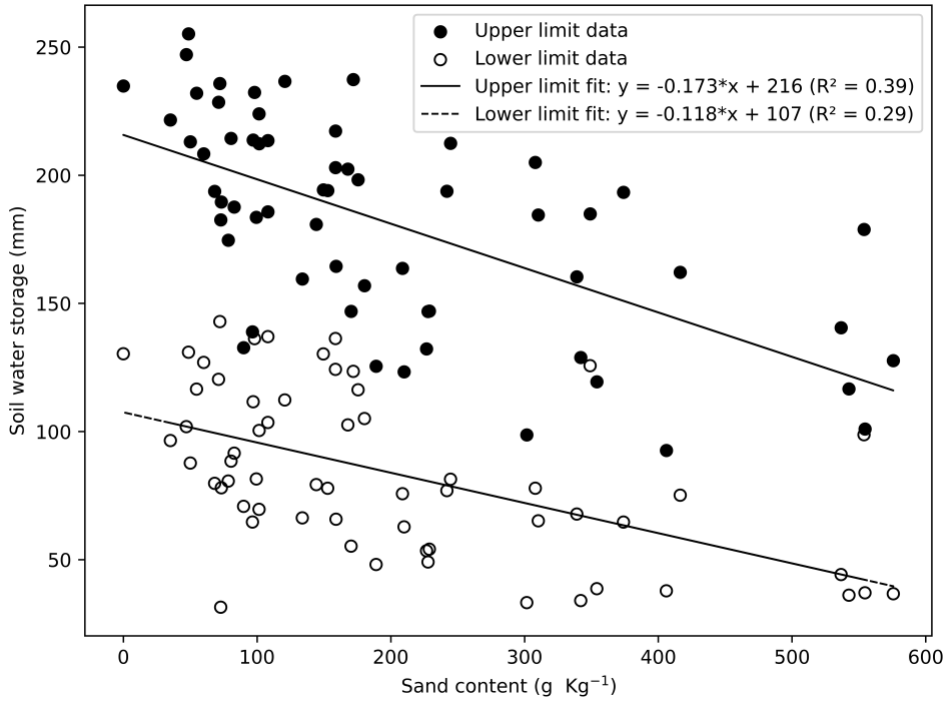


Figure 2-8. Fitted linear models between sand content extracted from SoilGrids and the observed upper and lower limits of soil water storage at each location of the Kansas Mesonet.

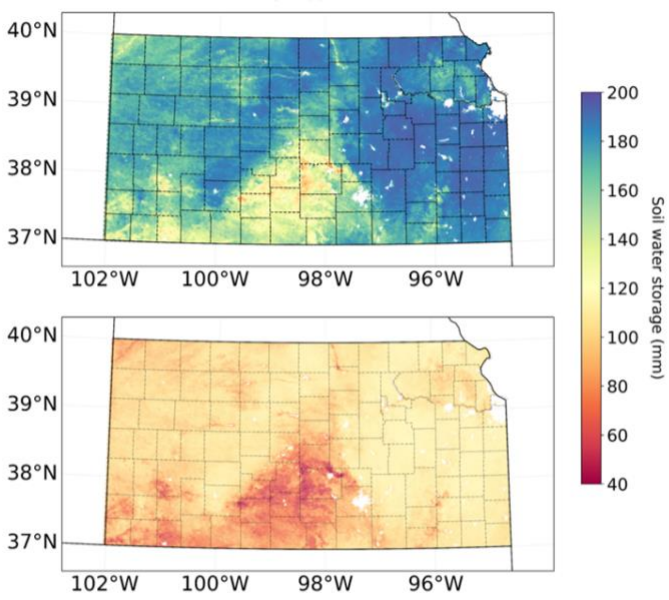


Figure 2-9. Maps of soil water storage upper (top) and lower limit (bottom) for the state of Kansas created using the adjusted linear regression between sand content and soil water storage limits.

Chapter 3 - CRNPy: An Open-Source Python Library for Cosmic-Ray Neutron Probe Data Processing

Joaquin Peraza[1], Tyson E. Ochsner[2], Andres Patrignani[1]

[1]Department of Agronomy, Kansas State University, Manhattan, KS 66506

[2]Department of Plant and Soil Sciences, Oklahoma State University, Stillwater, OK 74078

This manuscript has been submitted to the Journal of Open Source Software

Abstract

Accurate soil moisture estimates are vital for efficient water resources management and critical for agriculture, hydrology, and climate studies. Traditional methods for measuring soil moisture, such as gravimetric sampling and time-domain reflectometry, provide accurate estimations but over a small sensing volume. Cosmic-ray neutron probes (CRNPs) measure the number of neutrons produced by cosmic rays interacting with hydrogen atoms in the environment, including soil water content. Correcting cosmic-ray neutron count measurements with atmospheric conditions and other undesired interactions allows estimating soil moisture at a field scale (10 - 30 ha). This methodology has been increasingly used due to its large-scale footprint and non-invasive nature, providing continuous field-scale soil moisture monitoring. However, the workflow of converting CRNP data into soil moisture, including atmospheric corrections and other noise-removing filters, can be complex. This research introduces CRNPy, a Python library specifically designed to streamline and standardize the processing of CRNP data. Drawing upon empirical functions from multiple studies, CRNPy offers a flexible workflow to estimate various correction parameters, simplifying a traditionally complex process. With a focus on usability and accessibility, CRNPy lowers the barriers for researchers and students working in

the field of cosmic-ray soil moisture monitoring. Alternative correction routines were evaluated with CRNPy finding a trade-off between operational simplicity and accuracy. In scenarios exploring smoothing techniques, the library maintained consistent MAE of 0.029 m^3 m^{-3}, affirming the flexibility in choosing the smoothing step in stationary CRNP. However, when processing CRNP rover transects, a Root Mean Square Deviation (RMSD) of 0.019 m^3 m^{-3} was observed compared with *in situ* soil moisture stations, and a larger RMSD of 0.023 m^3 m^{-3} when applying spatial averaging directly to volumetric water content estimates. CRNPy, thus, presents a significant leap forward in the efficient and accurate interpretation of CRNP data, contributing valuable insights into developing and validating models and remote sensing tools in the soil physics and hydrology domain.

Introduction

The cosmic-ray neutron sensing technology is an innovative and non-invasive method for field-scale monitoring of soil moisture (Zreda et al., 2008). The sensing principle is based on the detection of fast neutrons that are generated when cosmic rays interact with atoms in the Earth's atmosphere. These secondary fast neutrons interact with hydrogen pools in the land surface, and thus, the intensity of neutrons detected near the ground surface is inversely related to the amount of water in land biomass and the top layers of the soil (Desilets et al., 2010). Typically, cosmic-ray neutron probes have a large sensing footprint, covering an area of approximately 12 hectares (about 30 acres) and can measure soil moisture up to a depth of 70 cm, although under most field conditions the effective sensing depth spans a depth ranging between 5 and 40 cm (Franz et al., 2012; Köhli et al., 2015). This makes cosmic-ray neutron probes (CRNP) a technology particularly useful for capturing spatially-averaged soil moisture over large areas, filling the gap between point-level sensors and remote sensors onboard orbiting satellites.

In hydrology, CRNP are often used to quantify surface soil water storage to improve hydrological models (Fatima et al., 2023) and better understand the soil water balance at the watershed and ecosystem scales. For example, stationary CRNP have been used to quantify soil water storage in a forest dominated by deciduous and coniferous trees in north-eastern Germany (Heidbüchel et al., 2016). Another study in two semiarid ecosystems in the southwestern United States validated the use of CRNP against a comprehensive soil moisture monitoring network, capturing soil moisture variations and enabling a deeper understanding of hydrological processes at the watershed level, including the relation between evapotranspiration and soil moisture across the region (Schreiner-McGraw et al., 2016). In agriculture, CRNP are often used to guide irrigation scheduling (Avery et al., 2018; Brogi et al., 2023; X. Han et al., 2016; Li et al., 2019),

determine water use efficiency (Chen et al., 2022), and quantify aboveground crop biomass (Franz et al., 2013; Jakobi et al., 2022; Tian et al., 2016). Because of the large sensing volume, stationary CRNP have recently been added as an integral component of mesoscale environmental monitoring networks (Bogena et al., 2022; Zreda et al., 2012). Roving versions of CRNP devices that can be mounted on moving vehicles for on-the-go soil moisture monitoring have also play an important role for characterizing mesoscale soil moisture spatial variability (Chrisman & Zreda, 2013) and to calibrate and validate remote sensing soil moisture products by conducting transects over large spatial domains with different soil types (Dong et al., 2014; Hornbuckle et al., 2012; Montzka et al., 2017). However, the process of cleaning, processing, and analyzing CRNP data involves multiple corrections and filtering steps spread across multiple peer-reviewed articles, which hinders reproducibility and represents a barrier to new investigators wanting to use this technology for soil moisture monitoring.

The objective of this work was to compile the progress made in the processing workflow for converting raw neutron counts into soil moisture since the inception of this technique into a user-friendly, versatile, and instrument agnostic library coded in the Python programming language. This new library, named “CRNPy”, was designed to be accessible to both researchers and instrument manufacturers seeking to integrate the processing workflow into their data analysis and hardware. Unlike other similar libraries (Power et al., 2021), CRNPy does not require any specific naming convention of column headers in the input data or the download of additional external data sources or reanalysis data.

Library computation routines

Using fast neutron counts collected with stationary and roving CRNP for soil moisture monitoring requires multiple correction steps (Figure 3-1 and Figure 3-2). The first step in the correction of observed raw neutron counts (N_{raw}) requires a correction including incoming neutron flux (f_i), atmospheric pressure (f_p) as recorded by collocated barometer, and air absolute humidity (f_w). The application of these three corrections factors lead to a new set of corrected neutron counts (N_{corr}):

$$N_{corr} = \frac{N_{raw} \cdot f_w}{f_p \cdot f_i} \quad (1)$$

Incoming neutron flux correction

The correction factor for the natural variation of incoming neutron intensity is calculated as the ratio of the measured neutron counts at a reference neutron monitor and a reference intensity during the study period (Andreasen et al., 2017):

$$f_i = I_m / I_0 \quad (2)$$

where f_i, is the correction factor for the incoming neutron flux, I_m is the timeseries of neutron counts obtained from a reference neutron monitor selected for a site with similar earth geomagnetic cutoff rigidity, and I_0 is a reference neutron count typically defined as the counts on the first day of the study. The geomagnetic cutoff rigidity, which is the shielding provided by the earth's magnetic field against charged cosmic ray particles, depends on the geographic latitude and altitude (Shea & Smart, 2019). This shielding is greatest at low elevations near the equator and is lowest at the poles. The functions *find_neutron_monitors* and *get_incoming_neutron_flux*

can be used to find and retrieve data from active reference neutron monitors listed in the Real-time Neutron Monitor Database (NMDB).

Atmospheric pressure correction

Raw neutron counts are also affected by variations in atmospheric pressure and air absolute humidity. For instance, the greater the atmospheric pressure and the greater absolute humidity, the greater is the attenuation in raw neutron counts. This fact also suggests that devices in locations at high altitudes tend to have greater raw neutron counts compared with CRNP devices located at low altitudes. Neutron counts variations caused by pressure changes in the atmosphere were corrected using the following equation (Zreda et al., 2012):

$$f_p = \exp[(P_0 - P)/L] \quad (3)$$

where f_p, represents the correction factor for atmospheric pressure, P_0 is the reference atmospheric pressure typically assumed to be the long-term value for the location or the atmospheric pressure at the time of the calibration of the CRNP, P is the observed atmospheric pressure in hectopascals using a collocated barometer, and L is the atmospheric attenuation coefficient (Andreasen et al., 2017; Zreda et al., 2012). The value of L is typically between 128 g cm^{-2} at high latitudes and 142 g cm^{-2} near the equator (Desilets & Zreda, 2003). For the state of Kansas a value of 130 g cm^{-2} is a good choice being also consistent with other studies in Oklahoma (Brown et al., 2023; Dong & Ochsner, 2018).

Atmospheric humidity correction

The impact of atmospheric humidity on N_{raw} was accounted for by using the following correction formula (Andreasen et al., 2017; Rosolem et al., 2013):

$$f_w = 1 + 0.0054 * \left(A - A_{ref}\right) \quad (4)$$

where f_w represents the correction factor for water vapor content, A is the observed absolute humidity measured at a similar height of the neutron detector, and A_{ref} is the reference absolute humidity. (Rosolem et al., 2013).

Biomass correction

In addition to atmospheric water vapor, water stored in plant biomass can also exert a moderating effect on raw neutron counts. Thus, to accurately estimate soil moisture conditions it is also necessary to account for this additional pool of water, particularly when devices are surrounded by high-biomass vegetation like forests or corn (*Zea mays* L.). The *bwe_correction* function was included for correcting neutron counts for the effects of aboveground biomass (Baatz et al., 2015).

$$f_b = 1 - BWE * r_2/N_{(0,BWE=0)} \quad (5)$$

$$BWE = SWB - SDB + SDB \cdot f_{WE} \quad (6)$$

where fb, represents the correction factor for water in surrounding biomass, BWE is the biomass water equivalent (kg m^{-2}) based on the stoichiometric amount of hydrogen and oxygen contained in cellulose, and $r_2/N_{(0,BWE=0)}$ is the ratio of neutron counts with biomass to neutron counts without biomass. *BWE* can be estimated with *in situ* destructive sampling of biomass, by measuring the standing wet biomass (SWB), the standing dry biomass (SDB) and including the stoichiometric ratio of water to organic carbon molecules in the plant (f_{WE}) that is approximately

0.494 (Wahbi et al., 2018). A weakness of this correction is that it only accounts for aboveground biomass ignoring the effect of water pools in root tissue. Recent studies have shown that CRNP are sensitive to root biomass, which would require improvements in the calibration. Similarly, if all other pools of hydrogen are accounted for, this methodology could be used to track field-scale root biomass (Jakobi et al., 2018). The CRNPy library implements this approach through the *biomass_to_bwe* and *correction_bwe* functions.

Road correction

Because sensing footprint of CRNP is greater in the immediate surroundings of the device, rover surveys also require a correction to account for differences between the field and underlying road soil water content. The CRNPy library includes a *road_correction* routine (Eq. 7) that accounts for: the geometry of the road (F1, Eq. 8), the difference in soil moisture between the road and the field (F2, Eq. 9), and the horizontal distance from the road center to the sensor (F3, Eq. 10) (Schrön et al., 2018):

$$f_r = 1 + F1 \cdot F2 \cdot F3 \tag{7}$$

$$F1 = p_0(1 - e^{-p_1 \cdot RW}) \tag{8}$$

$$F2 = -p_2 - p_3 \cdot \theta_{road} - \left(\frac{p_4 + \theta_{road}}{\theta_N}\right) \tag{9}$$

$$F3 = p_6 \cdot e^{-p_7 \cdot RW^{-p_8} \cdot RD^4} + (1 - p_6)e^{-p_9 \cdot RD} \tag{10}$$

where RW is the road width in meters, θ_{road} is the volumetric water content of the road, θ_N is the mean volumetric water content of the surrounding landscape, *RD* is the distance in meters from the road to the sensor, and parameters p_0 to p_8 have constant values defined in Table 1 in Schrön et al. (2018). The function *road_correction* calculates the correction factor in the CRNPy library.

Converting corrected counts into volumetric water content

Converting the corrected neutron counts (N_{corr}) into volumetric soil water content (θ) typically represents the last step of the workflow and is defined in the function named *counts_to_vwc* (Desilets et al., 2010):

$$\theta = \left(\frac{a_0}{\frac{N_{corr}}{N_0} - a_1} - a_2 - w_L \right) \rho_b \tag{11}$$

where N_0 is an instrument specific parameter that represents the corrected neutron counts on dry soil and is often obtained by conducting a field calibration, w_L is the soil lattice water, ρ_b is the soil bulk density, and parameters a_0, a_1 and a_2 are fitting parameters that usually have the following values: a_0 = 0.0808 , a_1 = 0.372 , a_2 = 0.115 (Desilets et al., 2010).

1) Uncertainty estimation

As with any instrument, it is also important to characterize the uncertainty of the measurements. The uncertainty in the corrected neutron counts is defined by the standard deviation (σ_N) of the Poissonian probability distribution (Schrön, 2016; Zreda et al., 2012):

$$\sigma_N = s\sqrt{N} \tag{12}$$

where *s* is the correction factor applied for all the involved corrections and *N* is the raw neutron count. In order to propagate the uncertainty error from the corrected neutron counts into the soil moisture estimates, the standard deviation is expressed in the form of a third order Taylor expansion (Jakobi et al., 2020):

$$\sigma_{\theta_g}(N) = \sigma_N \frac{a_0 N_0}{(N - a_1 N_0)^4}$$

$$\sqrt{[(N - a_1 N_0)^4 + 8\,\sigma_N^2 (N - a_1 N_0)^2 + 15\,\sigma_N^4]} \quad (13)$$

where $\sigma_{\theta_g}(N)$, is the standard deviation of the gravimetric soil moisture, *N* are the corrected neutron counts, N_0 is an instrument-specific parameter, and parameters a_0 and a_1 were defined in equation 10. In the CRNPy library this routine is implemented through the methods *uncertainty_counts* and *uncertainty_vwc.*

Sensing depth

The sensing depth can be obtained by computing two *e*-folding sample volume, which can be defined as the volume within which 86% (D_{86}) of the detected neutrons above the surface originate using the following equation (Franz et al., 2012; Köhli et al., 2015; Schrön et al., 2017):

$$D_{86} = \frac{1}{\rho_b}\left(8.32 + 0.14\left(0.97 + e^{-\frac{r}{100}}\right)\frac{26.42 + \theta}{0.057 + \theta}\right) \quad (14)$$

where ρ_b, is the field bulk density, θ is the estimated volumetric soil moisture estimated from the neutron counts, and r is the distance in meters between the site and the CRNP scaled by the pressure correction. The *sensing_depth* method allows the sensing depth estimation using the CRNPy library.

Outlier detection

The CRNPy library offers several methods for outliers detection using the *is_outlier* method, which includes a simple range detection based on user-provided lower and upper

boundaries, interquartile range, z-scores, and a scaled mean absolute difference (Iglewicz & Hoaglin, 1993). Outliers can be caused by hardware malfunction, vibrations, or external neutron sources (e.g., use of nearby neutron probe soil moisture meters like the 503 Hydroprobe by Instrotek, Inc.).

Exponential filter

Because observations with CRNP devices typically represent the soil moisture conditions in the top 10-20 cm, an exponential filter operator (Albergel et al., 2008) was added to the CRNPy library to extrapolate surface soil moisture conditions to the rootzone using the function *exp_filter* (Franz et al., 2020). For instance, the exponential filter has been successfully used to estimate rootzone soil water storage from *in situ* surface soil moisture measurements in on-farm cropland experiments in Kansas (Rossini & Patrignani, 2021).

Library Features

- The CRNPy library was implemented using the Python programming language, with the libraries Numpy (Harris et al., 2020), Pandas (The pandas development team, 2020), SciPy (Virtanen et al., 2020), and Matplotlib (Hunter, 2007), which are included in common Python data science bundles like the Anaconda open source ecosystem.
- Utility functions for determining site-specific information required before processing raw neutron counts include the determination of 1) lattice water (i.e., bounded water to clay particles), 2) geomagnetic cutoff rigidity (Smart & Shea, 2001), and 3) a reference neutron monitor (Klein et al., 2009).
- The library requires input data in the form of delimited text files, but does not require specific naming conventions column headers, thus allowing which increases reproducibility and minimizes human error. Each function of the CRNPy library accepts either a Numpy array or a Pandas series, enabling a more versatile, modular, and customizable workflow that adapts to instrument outputs from different manufacturers.
- For roving devices, CRNPy includes a few utility functions for spatial filtering and basic interpolation routines with cubic, linear, nearest neighbor, and inverse distance weighting interpolation methods.

Description of Use-case Scenarios

After developing the Cosmic-Ray Neutron Python library, we used it to test a series of alternatives and unexplored corrections, and post processing methods in two stationary CRNP use case scenarios and one roving CRNP use case scenario.

Use-case scenario 1 – Comparison of stationary CRNP corrections

Usually, CRNP also need to be equipped with additional research-grade atmospheric pressure and humidity sensors to implement the required correction routines (Rosolem et al., 2013; Zreda et al., 2012), which can further increase the complexity and cost of the required hardware (Patrignani et al., 2021). To find a way to circumvent some of these requirements, the first case scenario was aimed at evaluating alternative correction routines that require less inputs from auxiliary sensors onboard of stationary CRNP. The two alternative approaches that were tested are: i) applying only the pressure correction using collocated barometer observations, and ii) applying the pressure correction in addition to a fixed absolute humidity correction factor estimated using the long-term average value of absolute humidity for the study site. The evaluation was conducted using a dataset collected with a stationary CRNP (Radiation Detection Technologies, Inc., Manhattan, KS) deployed at the KONA site of the National Ecological Observatory Network (NEON) within the Konza Prairie Biological Station near Manhattan, KS (39.110 N, 96.613 W altitude 325 m a.s.l.) from 10 April 2020 to 18 June 2020. To benchmark the volumetric water content resulting from the different combination of correction routines we used the average volumetric water content from an existing array of five *in situ* soil moisture sensors at 5 and 15 cm depth from the NEON network. Mean atmospheric pressure 978 mbar and mean absolute humidity 7.7 g m^{-3} for this location were obtained from Ashland Bottoms

station of the Kansas Mesonet (Patrignani et al., 2020), which is located at ~2.5 km from the study site.

The accuracy of the resulting volumetric water content from stationary CRNP case-study scenarios was tested using the mean absolute error (MAE):

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\theta_N - \theta_{in-situ}| \quad (15)$$

where θ_N is the volumetric water content estimated with the CRNP, and $\theta_{in-situ}$ is the *in situ* volumetric water content from the KONA site of the NEON network.

Use-case scenario 2 – Temporal filter of neutron counts vs. soil moisture

A common practice to remove some of the noise in the signal of corrected neutron counts is to smooth the signal by applying a third-degree Savitzky-Golay temporal filter with an 11-hour window (Franz et al., 2020). However, because there is a non-linear relationship between raw neutron counts and volumetric water content, it remains unclear whether there is an advantage of conducting the temporal smoothing in the resulting volumetric water content compared with the typical smoothing of the raw neutron counts. The objective of this scenario was to test whether smoothing the signal of volumetric water content results in more accurate estimates of soil moisture compared with smoothing the signal of raw neutron counts. Using the same dataset and similar methodology from use case scenario 1, the accuracy of applying the smoothing filter to either corrected counts or final soil moisture estimations, was compared to the *in situ* observations from the KONA site of the NEON network.

Use-case scenario 3 – Spatial filtering of neutron counts vs. soil moisture

With roving CRNP platforms, spatial smoothing methods such as the nearest neighbor averaging, often yield more accurate results compared to temporal techniques like the Savitzky-Golay filter. This particularly occurs in areas with abundant local data points (i.e. transects with intersecting roads), where the spatial techniques can capture localized variations more effectively (Schrön et al., 2018). The aim of this use case scenario was to test the effect of smoothing the neutron counts before converting to volumetric water content *versus* directly smoothing the estimated volumetric water content.

Using 13 CRNP transects collected with a roving device (Hydroinnova LLC, Albuquerque, NM, USA) from 17 August 2021 to 16 November 2021 in the in the Kings Creek watershed of approximately 16 square kilometers located within the Konza Prairie Biological station, Kansas, USA (39.0964 N, 96.5848 W, altitude 340 m a.s.l.). The accuracy of applying a spatial average with a buffer of 800 m was compared to either corrected counts or final volumetric water content estimations. *In situ* soil moisture observations from a hydrological network consisting of 16 monitoring stations across the Kings Creek watershed, KS, was used for validation. Each station is equipped with soil moisture sensors (model TEROS 12, Meter, Inc., Pullman, WA, USA) at 5, 20, and 40 cm depth, the average of 5 and 20 cm depth sensors were used for validation. The CRNP soil moisture estimations were also compared to observations made by the Soil Moisture Active-Passive (Entekhabi et al., 2010) remote sensing product. For better consistency with previous studies, roving CRNP accuracy was evaluated using the Root Mean Square Deviation (RMSD) between the CRNP estimated volumetric water content (θ_N) and the *in situ* observations ($\theta_{in-situ}$) of the the Kings Creek hydrological network.

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\theta_N - \theta_{in-situ})^2} \quad (16)$$

Results and discussion

Use-case scenario 1 – Comparison of stationary CRNP corrections

The CRNPy library implemented correction and conversion workflow successfully allowed the estimation of volumetric water content from stationary CRNP observations using different combinations of correction factors (Figure 3-3). The stationary CRNP yielded a MAE of 0.029 m^3 m^{-3} compared to in situ observations when applying atmospheric pressure and absolute humidity corrections using collocated sensors, and incoming neutron flux corrections. The obtained MAE is consistent with results from previous studies in the same location (Patrignani et al., 2021). During the experiment, the mean hourly raw neutron count was 2170, after applying the correction factors, the mean hourly corrected count was 2280 (Figure 3-4). The MAE for the volumetric water content using only the pressure correction from the collocated barometer was 0.063 m^3 m^{-3} increasing the error by approximately 117 %. However, using the atmospheric pressure from the collocated barometer, and estimating a fixed humidity correction factor from long-term (10 years) observations of absolute humidity from the Ashland Bottoms, Kansas Mesonet station, led to a MAE of 0.037 m^3 m^{-3} being 28% higher than the MAE of using the correction factors computed from collocated sensors (Table 3-2). In the absence of a collocated pressure sensor, nearby pressure observations from environmental monitoring networks could be used for computing the pressure correction factor, as presented in Supplementary Material 1. This opens a new opportunity to reduce hardware complexity and costs by assessing the relevance of each correction when implementing CRNP systems.

During the study period, atmospheric pressure ranged from 95.7 to 99.2 kPa, which represents 66% of the mean range for this site. As a result, the pressure correction factor ranged from 0.89 to 1.16. Absolute humidity ranged from 2 to 24 g m^{-3}, leading to a correction factor

ranging from 1.01 to 1.13. The incoming neutron flux showed correction factors from 0.99 to 1.02 (Figure 3-5). The values for atmospheric and pressure corrections aligned with those reported in prior studies (Hawdon et al., 2014; Patrignani et al., 2021). This study presented smaller variations in incoming neutron flux compared to the range of approximately 0.86 to 1.22 reported by Hawdon et al. (2014).

Use case scenario 2 – Temporal filter of neutron counts vs. soil moisture

Applying the Savitzky-Golay filter to the volumetric water content resulted in a MAE of 0.029 m^3 m^{-3} when smoothing the corrected neutron counts, as proposed in previous studies (e.g., Davies et al., 2022; Franz et al., 2020; Patrignani et al., 2021). The alternative approach of smoothing the resulting volumetric water content observations led to the same MAE of 0.029 m^3 m^{-3} with negligible differences during the period of study (Figure 3-6).

Use-case scenario 3 – Spatial filtering of neutron counts vs. soil moisture

The CRNP rover transects were successfully processed using the CRNPy library. The mean volumetric water between the 13 rover surveys and the 16 *in situ* soil moisture stations within the Kings Creek watershed resulted in a RMSD of 0.019 m^3 m^{-3}. This value is similar to the accuracy of 0.03 m^3 m^{-3} reported by a previous study in cropland and rangeland areas that used CRNP rover surveys for calibrating and validating CRNP soil moisture estimations (Dong et al., 2014). A similar study evaluated the impact of roads on the accuracy of CRNP rover soil moisture estimations and reported errors between 0.020 and 0.061 m^3 m^{-3} (Schrön et al., 2018). In 10 out of the 13 rover transects, CRNP observations had a better agreement with *in situ* observations than the SMAP remote sensing product (Figure 3-7 and

Table 3-3), this highlights the potential of the cosmic-ray sensing technology and watershed *in situ* soil moisture networks to validate remote sensing products (Ochsner et al., 2013).

Applying the spatial average directly to the volumetric water content estimates decreased the accuracy, which was reflected in a larger RMSD = 0.023 m^3 m^{-3}. The non-linearity between corrected neutron counts and the resulting volumetric water content produced changes in the obtained estimations smoothing the corrected counts compared to smoothing the CRNP volumetric water content observations on the same date. These differences were consistently observed across all transects. Using a spatial filter that relies on the median value instead of the arithmetic average, prevented the occurrence of these changes, which resulted in the same values as when applying the filtering to corrected neutron counts or volumetric water content estimations.

Conclusions

- Atmospheric pressure and absolute humidity obtained from collocated sensors onboard of stationary CRNP agreed well with observations from nearby stations of the Kansas Mesonet. Our findings suggest that hourly data from nearby weather stations equipped with research-grade sensors could be used to fill in missing data or completely replace the need for installing additional instrumentation with the stationary CRNP.
- The alternative approach of applying only the pressure correction duplicates the Mean Absolute Error compared to using the complete set of corrections, suggesting a trade-off between operational simplicity and accuracy. While applying the pressure correction and fixed correction factor for absolute humidity using long-term average of absolute humidity for the site increased the error of MAE of the soil moisture estimations by 28%.
- In the case of stationary CRNP, smoothing either the corrected neutron counts or the volumetric water content had negligible impact in the final volumetric water content, offering CRNP users the flexibility to choose when to apply the temporal smoothing step.
- The spatial average filter applied to either corrected neutron counts or volumetric water content across multiple surveys conducted with a roving CRNP resulted in a 21% loss in accuracy when comparing to the watershed network estimates. Selecting an appropriate smoothing technique is necessary given the non-linearity of the data.

References

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

Andreasen, M., Jensen, K. H., Desilets, D., Franz, T. E., Zreda, M., Bogena, H. R., & Looms, M. C. (2017). Status and perspectives on the cosmic-ray neutron method for soil moisture estimation and other environmental science applications. Vadose Zone Journal, 16(8), 1–11.

Avery, W., Franz, T., Wahbi, A., Dercon, G., Heng, L., & Strauss, P. (2018). Mobile Soil Moisture Sensing in High Elevations: Applications of the Cosmic Ray Neutron Sensor Technique in Heterogeneous Terrain. EGU General Assembly Conference Abstracts, 9781.

Baatz, R., Bogena, H. R., Hendricks Franssen, H. -J., Huisman, J. A., Montzka, C., & Vereecken, H. (2015). An empirical vegetation correction for soil water content quantification using cosmic ray probes. Water Resources Research, 51(4), 2030–2046. https://doi.org/10.1002/2014WR016443

Bogena, H. R., Schrön, M., Jakobi, J., Ney, P., Zacharias, S., Andreasen, M., Baatz, R., Boorman, D., Duygu, M. B., Eguibar-Galán, M. A., Fersch, B., Franke, T., Geris, J., Sanchis, M. G., Kerr, Y., Korf, T., Mengistu, Z., Mialon, A., Nasta, P., … Blume, T. (2022). COSMOS-Europe: A European network of cosmic-ray neutron soil moisture sensors.

Brogi, C., Pisinaras, V., Köhli, M., Dombrowski, O., Franssen, H.-J. H., Babakos, K., Chatzi, A., Panagopoulos, A., & Bogena, H. R. (2023). Monitoring Irrigation in Small Orchards with Cosmic-Ray Neutron Sensors.

Brown, W. G., Cosh, M. H., Dong, J., & Ochsner, T. E. (2023). Upscaling soil moisture from point scale to field scale: Toward a general model. Vadose Zone Journal, 22(2), e20244. https://doi.org/10.1002/vzj2.20244

Chen, X., Song, W., Shi, Y., Liu, W., Lu, Y., Pang, Z., & Chen, X. (2022). Application of Cosmic-Ray Neutron Sensor Method to Calculate Field Water Use Efficiency.

Chrisman, B., & Zreda, M. (2013). Quantifying mesoscale soil moisture with the cosmic-ray rover. Hydrol. Earth Syst. Sci.

Davies, P., Baatz, R., Bogena, H. R., Quansah, E., & Amekudzi, L. K. (2022). Optimal Temporal Filtering of the Cosmic-Ray Neutron Signal to Reduce Soil Moisture Uncertainty. Sensors, 22(23), 9143. https://doi.org/10.3390/s22239143

Desilets, D., & Zreda, M. (2003). Spatial and temporal distribution of secondary cosmic-ray nucleon intensities and applications to in situ cosmogenic dating. Earth and Planetary Science Letters, 206(1–2), 21–42. https://doi.org/10.1016/S0012-821X(02)01088-9

Desilets, D., Zreda, M., & Ferré, T. P. A. (2010). Nature's neutron probe: Land surface hydrology at an elusive scale with cosmic rays: Nature's neutron probe. Water Resources Research, 46(11). https://doi.org/10.1029/2009WR008726

Dong, J., & Ochsner, T. E. (2018). Soil texture often exerts a stronger influence than precipitation on mesoscale soil moisture patterns. Water Resources Research, 54(3), 2199–2211.

Dong, J., Ochsner, T. E., Zreda, M., Cosh, M. H., & Zou, C. B. (2014). Calibration and Validation of the COSMOS Rover for Surface Soil Moisture Measurement. Vadose Zone Journal, 13(4), 1–8. https://doi.org/10.2136/vzj2013.08.0148

Entekhabi, D., Njoku, E. G., O'neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., & others. (2010). The soil moisture active passive (SMAP) mission. Proceedings of the IEEE, 98(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Fatima, E., Kumar, R., Attinger, S., Kaluza, M., Rakovec, O., Rebmann, C., Rosolem, R., Oswald, S., Samaniego, L., Zacharias, S., & Schrön, M. (2023). Improved representation of soil moisture simulations through incorporation of cosmic-ray neutron count measurements in a large-scale hydrologic model [Preprint]. Catchment hydrology/Modelling approaches. https://doi.org/10.5194/egusphere-2023-1548

Franz, T. E., Wahbi, A., Zhang, J., Vreugdenhil, M., Heng, L., Dercon, G., Strauss, P., Brocca, L., & Wagner, W. (2020). Practical data products from cosmic-ray neutron sensing for hydrological applications. Frontiers in Water, 2, 9.

Franz, T. E., Zreda, M., Ferre, T., Rosolem, R., Zweck, C., Stillman, S., Zeng, X., & Shuttleworth, W. (2012). Measurement depth of the cosmic ray soil moisture probe affected by hydrogen from various sources. Water Resources Research, 48(8).

Franz, T. E., Zreda, M., Rosolem, R., Hornbuckle, B. K., Irvin, S. L., Adams, H., Kolb, T. E., Zweck, C., & Shuttleworth, W. J. (2013). Ecosystem-scale measurements of biomass water using cosmic ray neutrons. Geophysical Research Letters, 40(15), 3929–3933. https://doi.org/10.1002/grl.50791

Han, X., Franssen, H.-J. H., Rosolem, R., Bogena, H., Alzamora, F. M., Chanzy, A., & Vereecken, H. (2016). Simultaneous soil moisture and properties estimation for a drip irrigated field by assimilating cosmic-ray neutron intensity. https://doi.org/10.1016/j.jhydrol.2016.05.050

Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hawdon, A., McJannet, D., & Wallace, J. (2014). Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia. Water Resources Research, 50(6), 5029–5043. https://doi.org/10.1002/2013WR015138

Heidbüchel, I., Güntner, A., & Blume, T. (2016). Use of cosmic-ray neutron sensors for soil moisture monitoring in forests. Hydrology and Earth System Sciences, 20(3), 1269–1288. https://doi.org/10.5194/hess-20-1269-2016

Hornbuckle, B., Irvin, S., Franz, T., Rosolem, R., & Zweck, C. (2012). The potential of the COSMOS network to be a source of new soil moisture information for SMOS and SMAP. 2012 IEEE International Geoscience and Remote Sensing Symposium, 1243–1246. https://doi.org/10.1109/IGARSS.2012.6351317

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Iglewicz, B., & Hoaglin, D. C. (1993). Volume 16: How to detect and handle outliers. Quality Press.

Jakobi, J., Huisman, J. A., Fuchs, H., Vereecken, H., & Bogena, H. R. (2022). Potential of Thermal Neutrons to Correct Cosmic-Ray Neutron Soil Moisture Content Measurements for Dynamic Biomass Effects. Water Resources Research, 58(8). https://doi.org/10.1029/2022WR031972

Jakobi, J., Huisman, J. A., Schrön, M., Fiedler, J., Brogi, C., Vereecken, H., & Bogena, H. R. (2020). Error Estimation for Soil Moisture Measurements With Cosmic Ray Neutron Sensing and Implications for Rover Surveys. Frontiers in Water, 2, 10. https://doi.org/10.3389/frwa.2020.00010

Jakobi, J., Huisman, J. A., Vereecken, H., Diekkrüger, B., & Bogena, H. R. (2018). Cosmic Ray Neutron Sensing for Simultaneous Soil Water Content and Biomass Quantification in Drought Conditions. Water Resources Research, 54(10), 7383–7402. https://doi.org/10.1029/2018WR022692

Klein, K.-L., Steigies, C., & Nmdb Team. (2009). WWW.NMDB.EU: The real-time Neutron Monitor database. EGU General Assembly Conference Abstracts, 5633.

Köhli, M., Schrön, M., Zreda, M., Schmidt, U., Dietrich, P., & Zacharias, S. (2015). Footprint characteristics revised for field-scale soil moisture monitoring with cosmic-ray neutrons. Water Resources Research, 51(7), 5772–5790. https://doi.org/10.1002/2015WR017169

Li, D., Schrön, M., Köhli, M., Bogena, H., Weimar, J., Bello, M. A. J., Han, X., Gimeno, M. A. M., Zacharias, S., Vereecken, H., & Franssen, H.-J. H. (2019). Can Drip Irrigation be Scheduled with Cosmic-Ray Neutron Sensing? https://doi.org/10.2136/vzj2019.05.0053

Montzka, C., Bogena, H., Zreda, M., Monerris, A., Morrison, R., Muddu, S., & Vereecken, H. (2017). Validation of Spaceborne and Modelled Surface Soil Moisture Products with

Cosmic-Ray Neutron Probes. Remote Sensing, 9(2), 103. https://doi.org/10.3390/rs9020103

Ochsner, T. E., Cosh, M. H., Cuenca, R. H., Dorigo, W. A., Draper, C. S., Hagimoto, Y., Kerr, Y. H., Larson, K. M., Njoku, E. G., Small, E. E., & Zreda, M. (2013). State of the Art in Large-Scale Soil Moisture Monitoring. Soil Science Society of America Journal, 77(6), 1888–1919. https://doi.org/10.2136/sssaj2013.03.0093

Patrignani, A., Knapp, M., Redmond, C., & Santos, E. (2020). Technical Overview of the Kansas Mesonet. Journal of Atmospheric and Oceanic Technology, 37(12), 2167–2183. https://doi.org/10.1175/JTECH-D-19-0214.1

Patrignani, A., Ochsner, T. E., Montag, B., & Bellinger, S. (2021). A Novel Lithium Foil Cosmic-Ray Neutron Detector for Measuring Field-Scale Soil Moisture. Frontiers in Water, 3, 673185. https://doi.org/10.3389/frwa.2021.673185

Power, D., Rico-Ramirez, M. A., Desilets, S., Desilets, D., & Rosolem, R. (2021). Cosmic-Ray neutron Sensor PYthon tool (crspy 1.2.1): An open-source tool for the processing of cosmic-ray neutron and soil moisture data. Geosci. Model Dev.

Rosolem, R., Shuttleworth, W. J., Zreda, M., Franz, T. E., Zeng, X., & Kurc, S. A. (2013). The Effect of Atmospheric Water Vapor on Neutron Count in the Cosmic-Ray Soil Moisture Observing System. Journal of Hydrometeorology, 14(5), 1659–1671. https://doi.org/10.1175/JHM-D-12-0120.1

Rossini, P., & Patrignani, A. (2021). Predicting rootzone soil moisture from surface observations in cropland using an exponential filter. Soil Science Society of America Journal, 85(6), 1894–1902.

Schreiner-McGraw, A. P., Vivoni, E. R., Mascaro, G., & Franz, T. E. (2016). Closing the water balance with cosmic-ray soil moisture measurements and assessing their relation to evapotranspiration in two semiarid watersheds. Hydrology and Earth System Sciences, 20(1), 329–345. https://doi.org/10.5194/hess-20-329-2016

Schrön, M. (2016). Cosmic-ray neutron sensing and its applications to soil and land surface hydrology [Doctoralthesis]. Universität Potsdam.

Schrön, M., Köhli, M., Scheiffele, L., Iwema, J., Bogena, H. R., Lv, L., Martini, E., Baroni, G., Rosolem, R., Weimar, J., Mai, J., Cuntz, M., Rebmann, C., Oswald, S. E., Dietrich, P., Schmidt, U., & Zacharias, S. (2017). Improving calibration and validation of cosmic-ray neutron sensors in the light of spatial sensitivity. Hydrology and Earth System Sciences, 21(10), 5009–5030. https://doi.org/10.5194/hess-21-5009-2017

Schrön, M., Rosolem, R., Köhli, M., Piussi, L., Schröter, I., Iwema, J., Kögler, S., Oswald, S. E., Wollschläger, U., Samaniego, L., Dietrich, P., & Zacharias, S. (2018). Cosmic-ray Neutron Rover Surveys of Field Soil Moisture and the Influence of Roads. Water Resources Research, 54(9), 6441–6459. https://doi.org/10.1029/2017WR021719

Shea, M., & Smart, D. (2019). Re-examination of the First Five Ground-Level Events. International Cosmic Ray Conference (ICRC2019), 36, 1149.

Smart, D., & Shea, M. (2001). Geomagnetic cutoff rigidity computer program: Theory, software description and example.

The pandas development team. (2020). pandas-dev/pandas: Pandas (latest) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.3509134

Tian, Z., Li, Z., Liu, G., Li, B., & Ren, T. (2016). Soil water content determination with cosmic-ray neutron sensor: Correcting aboveground hydrogen effects with thermal/fast neutron ratio. Journal of Hydrology, 540, 923–933. https://doi.org/10.1016/j.jhydrol.2016.07.004

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wahbi, A., Heng, L., Dercon, G., Wahbi, A., & Avery, W. (2018). In Situ Destructive Sampling. Cosmic Ray Neutron Sensing: Estimation of Agricultural Crop Biomass Water Equivalent, 5–9.

Zreda, M., Desilets, D., Ferré, T. P. A., & Scott, R. L. (2008). Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. Geophysical Research Letters, 35(21), L21402. https://doi.org/10.1029/2008GL035655

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., & Rosolem, R. (2012). COSMOS: The COsmic-ray Soil Moisture Observing System. Hydrology and Earth System Sciences, 16(11), 4079–4099. https://doi.org/10.5194/hess-16-4079-2012

Table 3-1. Average raw neutron counts per minute, atmospheric pressure, absolute humidity and reference incoming neutron flux obtained from the Dourbes (3.18 GV, 225 m a.s.l) neutron monitor in Belgium part of the NMDB (Klein et al., 2009) network, for each CRNP rover transect collected in the Kings Creek watershed.

Date	Mean raw neutron counts (cpm)	Mean Atmospheric Pressure (kPa)	Mean Atmospheric Absolute Humidity ($g\ m^{-3}$)	Incoming neutron flux (cps)
18 Aug 2021	318	971.8	16.8	114
23 Aug 2021	329	967.1	15.9	113
02 Sep 2021	321	966.0	19.1	113
05 Sep 2021	286	974.6	14.3	115
08 Sep 2021	300	974.5	10.1	113
13 Sep 2021	320	968.5	14.2	114
20 Sep 2021	335	964.7	15.4	112
28 Sep 2021	338	965.9	12.2	112
04 Oct 2021	332	974.0	7.7	115
11 Oct 2021	330	960.7	8.1	113
23 Oct 2021	331	963.9	9.6	112
01 Nov 2021	274	983.6	5.0	111
17 Nov 2021	302	974.4	2.4	112

Table 3-2. Correction factors range and Mean Absolute Error (MAE) of three different approaches, using only the atmospheric pressure correction sourced from the collocated barometer, using the atmospheric pressure correction factor from the collocated barometer and a correction factor for absolute humidity computed from long term site average observed in the Ashland Bottoms Kansas Mesonet station, and using the atmospheric pressure, absolute humidity, correction factor from collocated sensors and the incoming neutron flux corrections.

Applied corrections	**Factor Range**	**MAE**
Obs. Pressure	0.89 – 1.16	0.063 m^3 m^{-3}
Obs .Pressure + LT - Abs Humidity	0.90 – 1.17	0.037 m^3 m^{-3}
Obs. Pressure + Obs. Abs. Humidity + Incoming flux	0.90 – 1.15	0.029 m^3 m^{-3}

Table 3-3. Average corrected neutron count and estimated volumetric water content for each rover transect of the Kings Creek watershed. *In situ* observations of sensors at 5 and 15 cm depth that are part of the NEON KONA site, and value of surface volumetric water content (0 – 5 cm) of a SMAP pixel that covers the watershed.

Date	Corrected neutron counts (cpm)	CRNP Avg. VWC (m^3 m^{-3})	In situ Avg. VWC (m^3 m^{-3})	SMAP Avg. VWC (m^3 m^{-3})
18 Aug 2021	330	0.202	0.192	0.181
23 Aug 2021	330	0.225	0.233	0.235
02 Sep 2021	323	0.205	0.182	0.240
05 Sep 2021	298	0.342	0.345	0.282
08 Sep 2021	309	0.286	0.306	0.247
13 Sep 2021	320	0.237	0.257	0.195
20 Sep 2021	330	0.204	0.217	0.196
28 Sep 2021	331	0.197	0.192	0.162
04 Oct 2021	334	0.200	0.188	0.217
11 Oct 2021	302	0.286	0.240	0.276
23 Oct 2021	314	0.242	0.242	0.245
01 Nov 2021	297	0.322	0.314	0.293
17 Nov 2021	301	0.298	0.331	0.309

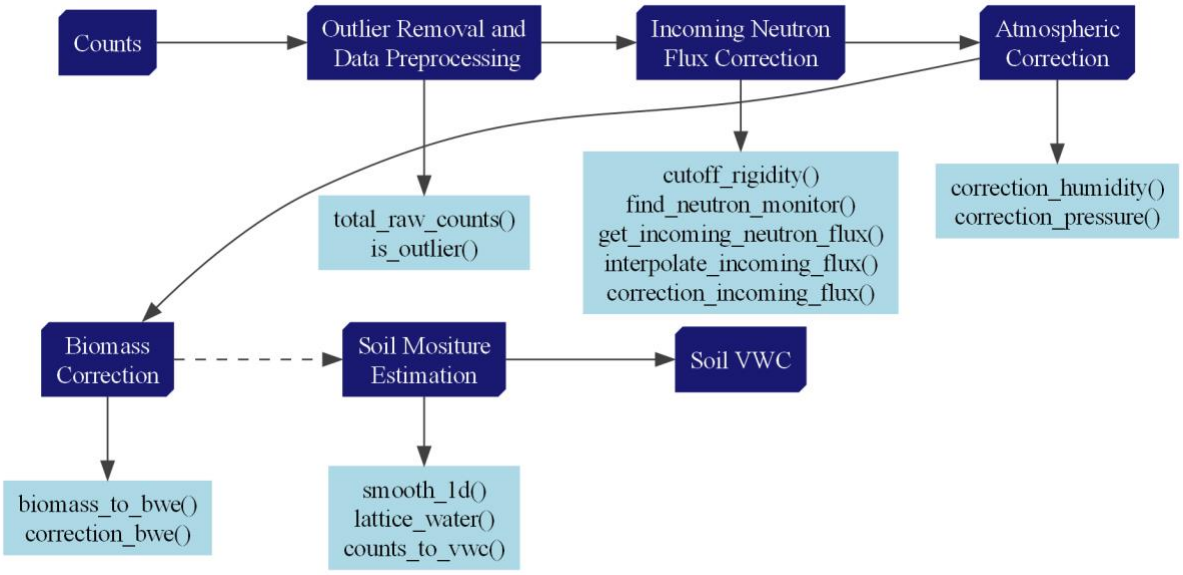


Figure 3-1. Example workflow for stationary CRNP, dashed lines represent optional steps.

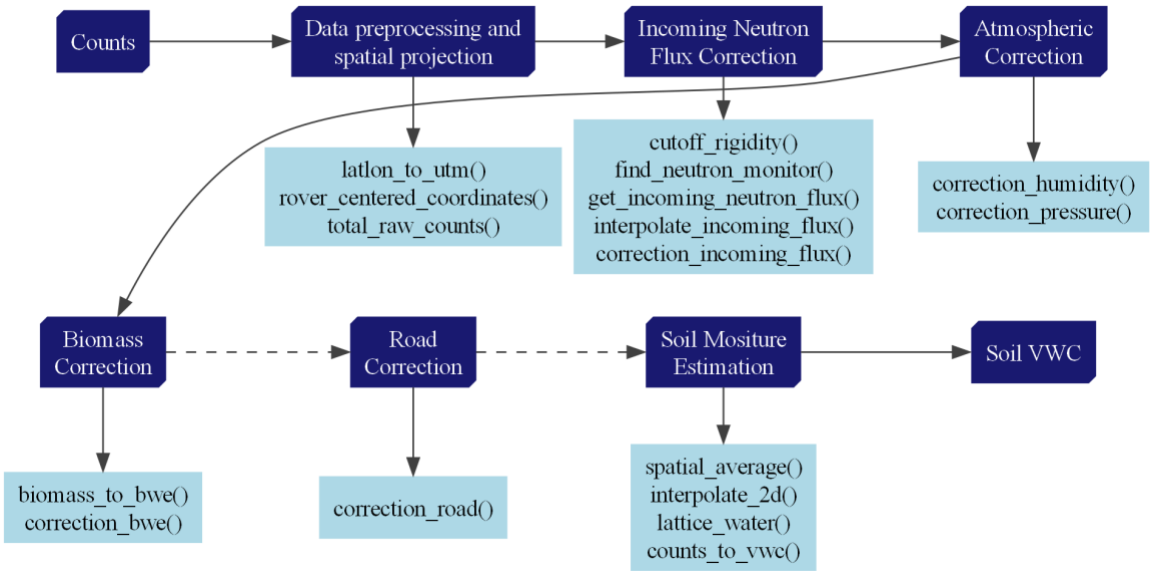


Figure 3-2. Example workflow for roving CRNP, dashed lines represent optional steps.

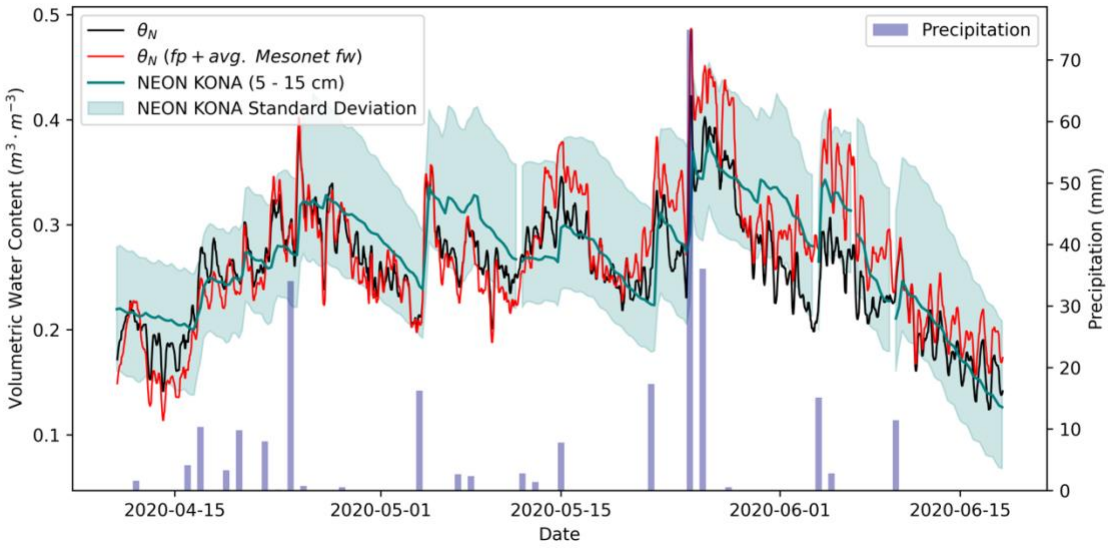


Figure 3-3. Volumetric water content estimated from 1) neutron counts corrected by hourly pressure, humidity, and incoming neutron flux (black line) and 2) neutron counts corrected by hourly pressure and mean humidity for the site (red line). The benchmark for this test was the mean volumetric water content obtained from five in situ sensors at 5 and 15 cm depth that are part of the NEON KONA site (teal line) with the standard deviation band. Daily precipitation observations were obtained from the NEON KONA site.

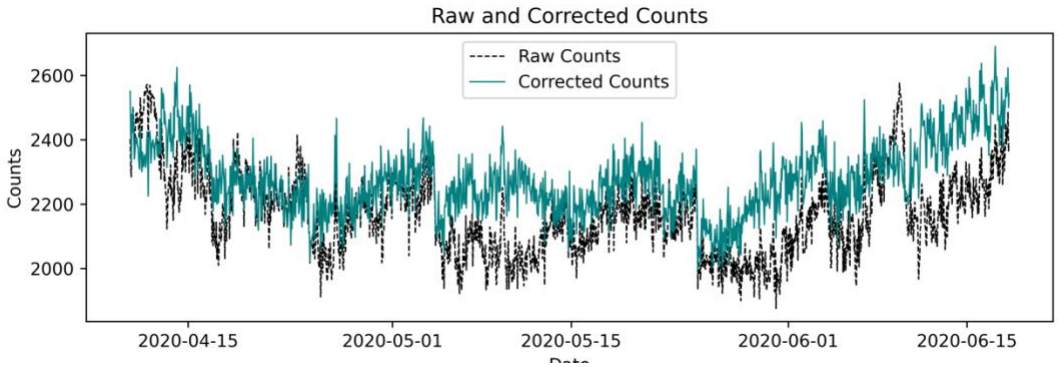


Figure 3-4. Raw and corrected neutron counts recorded with a stationary CRNP at the Konza Prairie Biological Station near Manhattan, Kansas.

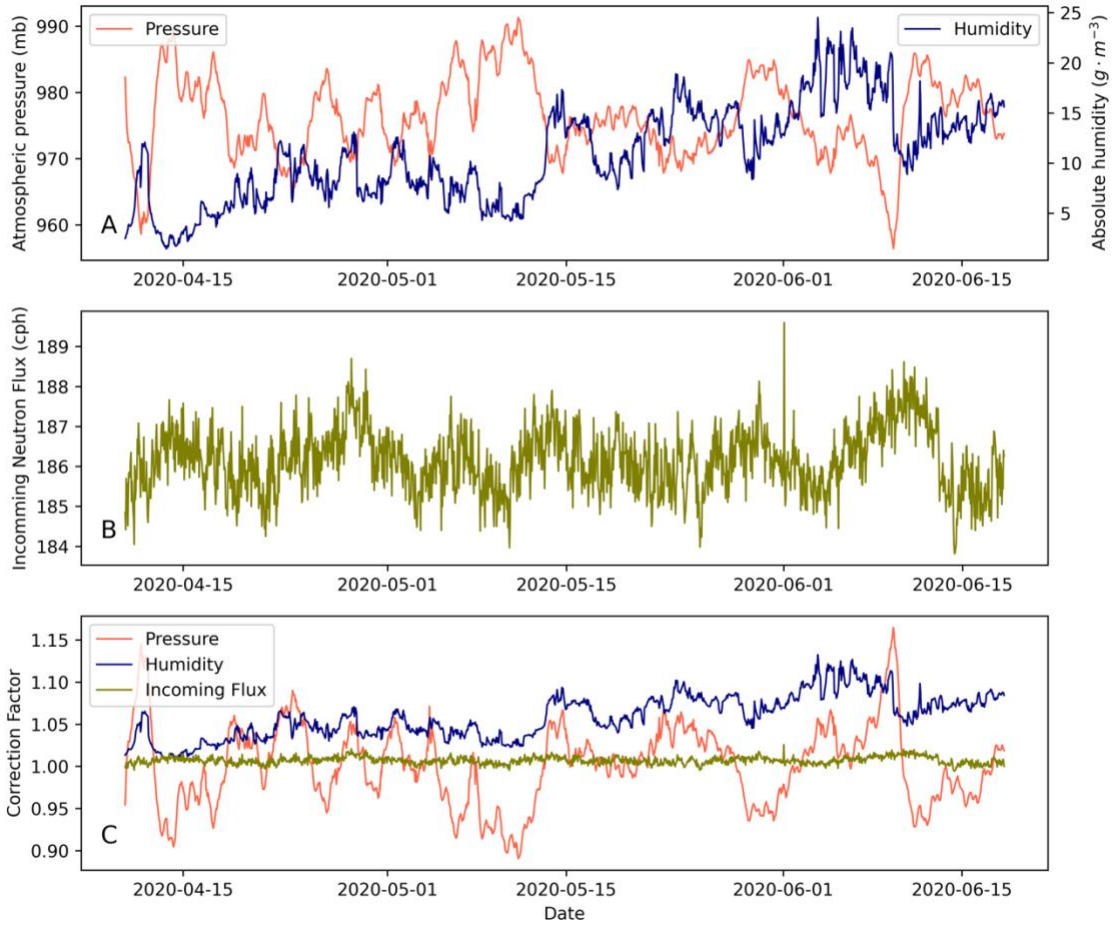


Figure 3-5. A) Barometric pressure and air absolute humidity, B) incoming neutron flux recorded by the Irkutst Neutron Monitor part of the Neutron Monitor Database (NMDB), which has similar elevation and cutoff rigidity as the experimental site, and C) correction factors for the stationary CRNP within the Konza Prairie Biological Station near Manhattan, Kansas.

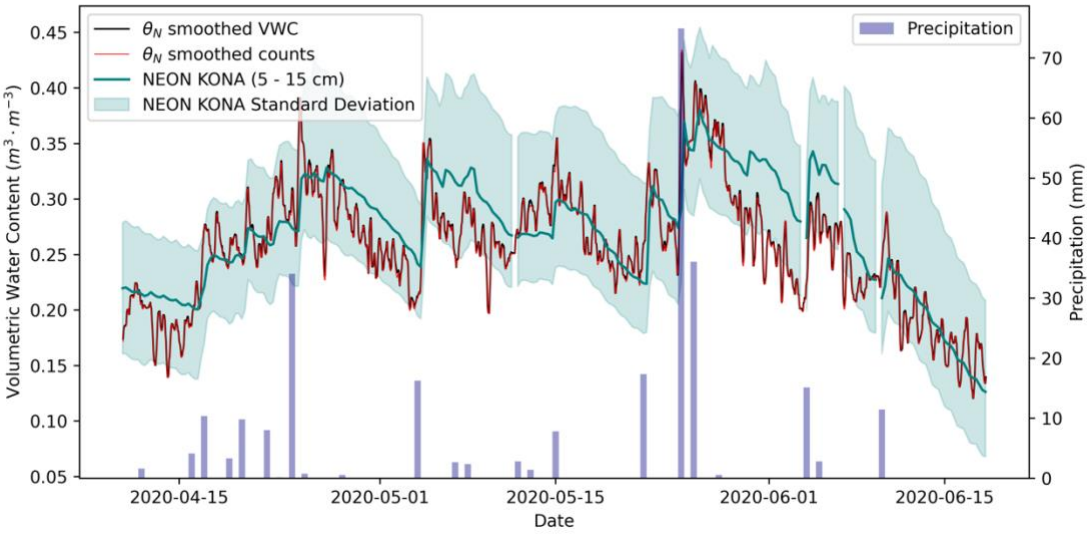


Figure 3-6. Volumetric water content obtained by smoothing using an 11-hour third-order Savitzky-Golay filter in the corrected neutron counts signal (red line) and smoothing the resulting volumetric water content (black line). The benchmark for this test was the mean volumetric water content obtained from five *in situ* sensors at 5 and 15 cm depth that are part of the NEON KONA site (teal line) with the standard deviation band. Daily precipitation observations were obtained from the NEON KONA site and plotted in bars.

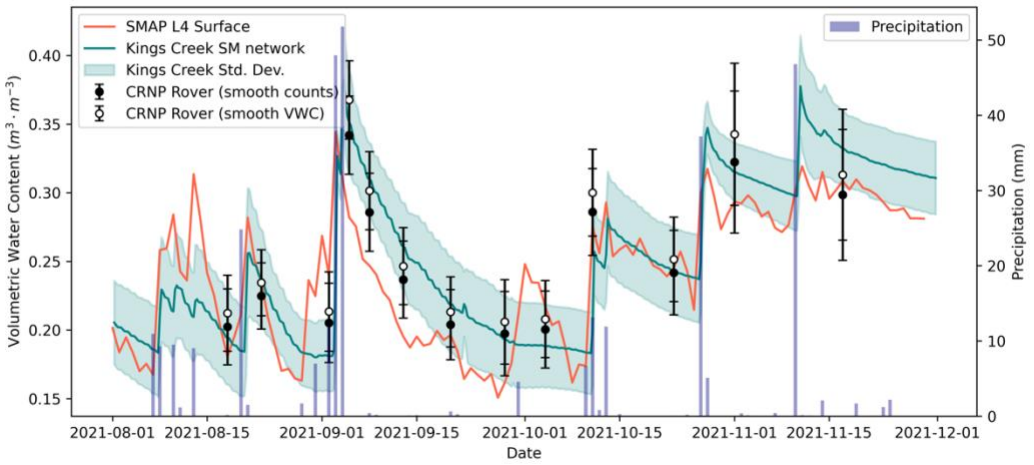


Figure 3-7. Volumetric water content from the Kings Creek *in situ* soil moisture monitoring network, averaging 5 and 20 cm depth sensors (teal line) with the standard error band, daily SMAP L4 surface soil moisture values for the pixel covering the watershed (red line), and the average of each CRNP roving transects with the standard deviation of the corrected counts (black markers) or volumetric water content estimations (white markers) after applying a 2D moving average in each variable. Daily precipitation observations were obtained from the Kings Creek network, the average value is represented in bars.

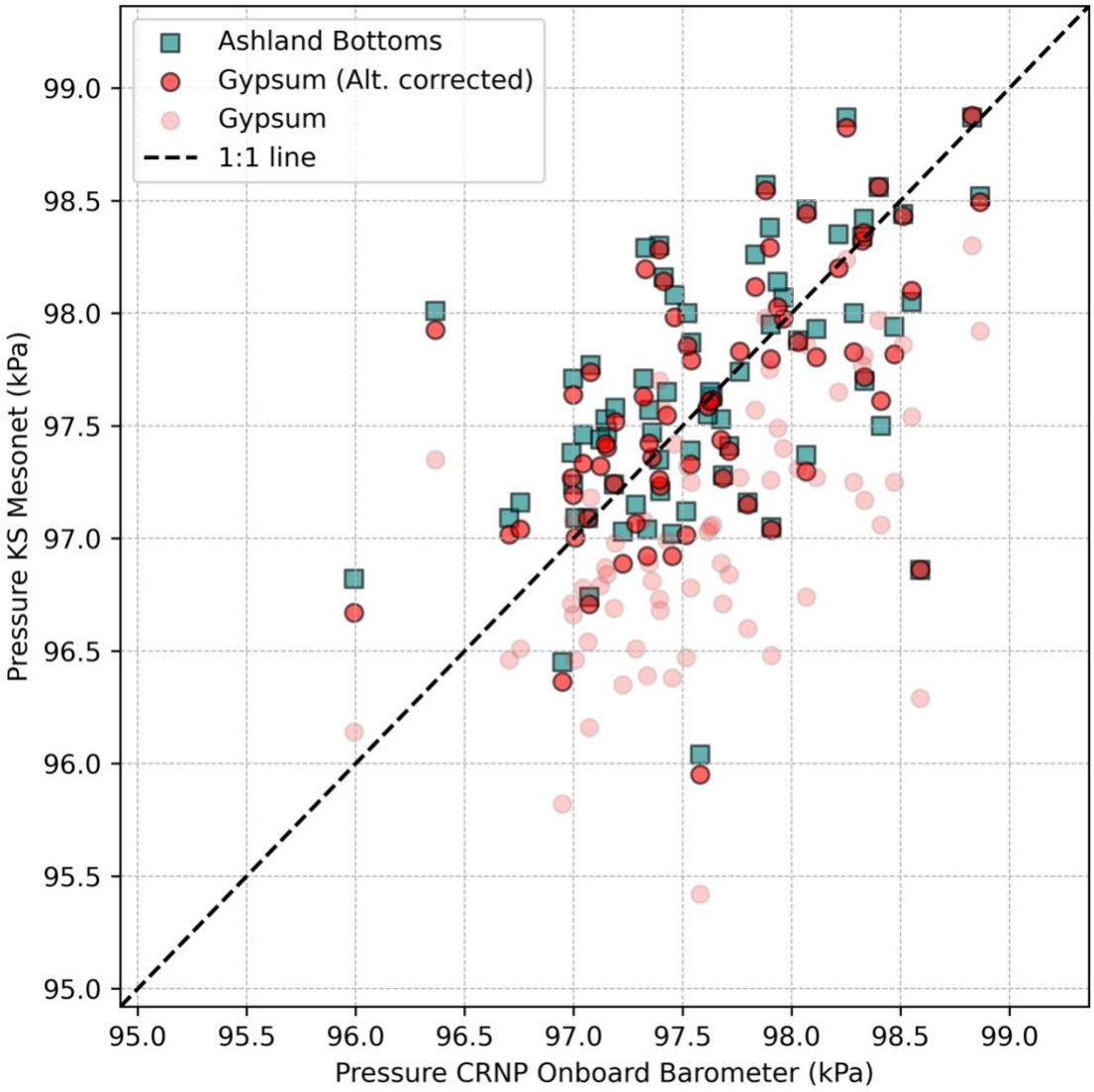


Figure 3-8. Comparison of pressure observations from the CRNP onboard barometer, and different stations of the Kansas Mesonet. The Ashland Bottoms station (teal square markers) located 2.5 km from the CRNP device was located at the same altitude (39.126 N, −96.637 W, altitude 325 m a.s.l.), while the Gypsum, KS station (38.725 N, −97.444 W, altitude 373 m a.s.l.) located 90 km away presented a negative bias (shadow markers) removed after correcting for the altitude difference using Eq. 41 from Bellamy (1945).

Chapter 4 - Using an Observation Operator to Translate Soil Moisture Conditions Across Land Covers

Joaquin Peraza[1], Ignacio A. Ciampitti[1], Eduardo Santos[1], Andres Patrignani[1]
Department of Agronomy, Kansas State University, Manhattan, KS 66506

Abstract

Soil moisture influences the development of crops during the entire growing cycle, however, obtaining soil moisture conditions in cropland areas is challenging. Most large-scale in situ soil moisture monitoring networks predominantly target grassland vegetation, presenting differences in soil water dynamics due to growth and development differences. This research explores the root-zone soil water storage dynamics between grassland and different cropland rotations within Kansas. The objectives were (i) to draw a comparison of root-zone soil water storage dynamics between grassland cropland rotations (ii) to build an observation operator for translating grassland soil water storage in the top 50 cm of the soil profile to cropland soil water storage using a Random Forest (RF) as an observation operator. We found that using a 60-days window of grassland soil water storage observations extracted from the adjacent permanent sensors and EVI differences between grassland and cropland vegetation it is possible to estimate the cropland soil water storage in the top 50 cm of the soil profile with a RMSE of 27 mm. The study recognized a lag of 15 days between high vegetation development and soil moisture depletion in the cropland rotations, and a lag of 7 days on average between the vegetation differences increase and the soil water storage operator increase, concluding that the impact of diverse land covers on soil moisture is not instantaneous and varies considerably throughout the annual cycle.

Introduction

Soil moisture is a fundamental variable influencing agricultural and hydrological processes such as rainfall partitioning into infiltration and runoff (Crow et al., 2018) and grassland and cropland productivity. For example, soil moisture influences the grain yield and development of summer crops including corn at different phenological stages (Denmead & Shaw, 1960). Similarly, a study in Canada observed that soil moisture was the most significant weather variable when predicting grain yield of winter wheat during the entire crop cycle, highlighting a higher impact of soil moisture on yield than precipitation (Baier & Robertson, 1968). However, monitoring soil moisture dynamics in cropland can be difficult because soil moisture sensors and associated hardware can conflict with farming operations.

Environmental monitoring networks that provide scientists and end users with soil moisture information have the potential to be used as a proxy for soil moisture conditions in nearby cropland. sensors differ between neighboring fields in landscapes with diverse land covers. For instance, grasslands might retain moisture due to their dense canopy, and adjacent croplands might exhibit different moisture patterns influenced by the crop's growth stage and management practices (Rodriguez-Iturbe et al., 2001; Wells et al., 2014). Understanding these is crucial, as soil moisture directly impacts crop yield, irrigation strategies, and long-term land use planning.

Remote sensing soil moisture from satellite missions (Entekhabi et al., 2010) has provided valuable insights into soil moisture dynamics, but the spatial resolution of these remote sensors is often not suitable for landscapes with intermixed land covers (Ochsner et al., 2013). An alternative is to use soil moisture sensors installed at the edge of the field. Similarly, a study in the Upper Cedar Creek watershed in Indiana investigated soil moisture dynamics using soil

moisture sensors installed at the edge of the field and found that all the point-based estimates of permanent sensors were underestimating the field soil moisture conditions of corn and soybeans rotations, observed negative biases of -0.276 and -0.135 $m^3 m^{-3}$ between field average and permanent adjacent sensors (Heathman et al., 2012). More sophisticated methods use sensors at the edge-of-the-field or existing environmental monitoring networks equipped with soil moisture sensors in combination with an observation operator. A study Oklahoma showed that ANN can be successfully used as an observation operator to relate in situ soil moisture observations under grassland (warm-season grasses) to represent soil moisture conditions in nearby winter wheat fields (cool season crop) (Patrignani & Ochsner, 2018). Since both land covers were exposed to similar environmental conditions, differences in soil moisture dynamics were mostly driven by differences in canopy development, and thus hindering the direct use of *in situ* soil moisture information from mesoscale environmental monitoring networks for inferring cropland soil moisture conditions.

A similar study in the same watershed determined that it is possible to identify time-stable locations to install soil moisture sensors and successfully implemented observation operators to upscale the point-level observations to estimate field conditions using the cumulative distribution function (CDF) matching method (Han et al., 2012). The challenge remains in harmonizing large-scale observations with local *in situ* measurements, especially in heterogeneous landscapes.

The development of observation operators enable the removal of most of the systematical errors from soil moisture observations including instrument or methodology biases (Drusch, 2005; Han et al., 2012; Patrignani & Ochsner, 2018) or the extrapolation of surface soil moisture measurements to different points of the soil profile (Gao et al., 2019). Several studies have been

using CDF matching techniques (De Lannoy et al., 2007; Drusch, 2005; Gao et al., 2019; Han et al., 2012), but recently implementations using machine-learning techniques had arisen to obtain observation operators for complex scenarios with non-linear interactions (Patrignani & Ochsner, 2018). The application of these observation operators can help bridge the gap between field-scale soil moisture and in situ soil moisture networks (Han et al., 2012; Patrignani & Ochsner, 2018). These operators can be particularly useful in translating soil moisture measurements from permanent sensors to cropland environments, where soil moisture dynamics can be significantly different due to factors like crop type, growth stage, and management practices. The sustained deployment of soil moisture sensors in cropland fields poses operational challenges, primarily due to the frequent farming activities that necessitate sensor removal and reinstallation.

Recognizing this gap, the objective of our study was to use an observation operator to estimate cropland soil water storage in the top 50 cm using the observed differences in vegetation development between grassland and cropland.

Materials and Methods

Soil moisture data

In situ Soil moisture observations from October 2020 to July 2023 were obtained from three different sources (Figure 4-1): i) the Ashland Bottoms (ASB) station of the Kansas Mesonet (Patrignani et al., 2020) that records daily soil moisture observations at 5, 10, 20, and 50 cm depth using soil water reflectometers (model CS655, Campbell Scientific), ii) an adjacent replicated plot experiment with five different crop rotations where weekly soil moisture measurements in the top meter of the soil profile were collected using a capacitance probe (model Diviner 2000, Sentek) at 10-cm intervals, and iii) a nearby cropland field located within the Konza Prairie Biological Station at about 2.5 km from the Mesonet station that is part of the National Ecological Observatory Network (NEON KONA site) with five multi-depth sensors (EnviroSCAN, Sentek) spanning the top 2 meters of the soil profile. Across all three sites, the soil water storage in the top 50 cm of the soil profile was computed using the trapezoidal rule of integration. This depth of 50 cm was selected because is the maximum depth common across all three sites and it represents the standard rooting depth across all stations of the Kansas Mesonet (Patrignani et al., 2020).

Vegetation data

Vegetation dynamics at the ASB station of the Kansas Mesonet were represented by the enhanced vegetation index (EVI, 10-m spatial resolution) retrieved from Sentinel-2 Level-2A using a polygon delineating the support area (2,300 m^2) of the station. For each crop rotation, vegetation conditions were represented using periodic observations of green canopy cover determined using downward-facing images. To match the vegetation dynamics of the other sites,

these ground-based measurements of green canopy cover were scaled to equivalent EVI values using a linear model. The linear model to scale GCC values was created using an additional dataset of 380 collocated observations of GCC and EVI for multiple cropland fields (Figure 4-2). EVI from Sentinel-2 Level-2A was also retrieved for a supporting area of 2000 m^2 around the NEON KONA cropland soil moisture monitoring site used for validation.

The diversity and intensification level of the different crop rotations (CR) in the replicated plot experiment provided a wide range of vegetation and soil moisture dynamics that allowed us to capture the typical variability crop growth and development of crops in this region. Specifically, crop rotations consisted of the following crop sequences: one through four presented varied scenarios showcasing winter wheat cultivated under different intensities and management practices. Meanwhile, rotation five stands out due to its continuous cover of Alfalfa (*Medicago sativa* L.). In years 2022 and 2022 Triticale (*x Triticosecale* Wittmack) was intermixed within rotation five, adding an extra layer of variation. This arrangement ensures that the study sites reflect a broad representation of vegetation and crop rotations, providing comprehensive insights into their growth patterns and interactions (Figure 4-3). NEON KONA site provides a different crop rotation that includes winter wheat, corn (*Zea mays* L.), and longer fallow periods in a nearby location, allowing a more accurate validation of the methodology.

Observation operator

In this study we used a Random Forest (RF) machine-learning model as the observation operator to estimate cropland soil moisture conditions based on grassland soil moisture at the ASB station of the Kansas Mesonet. The RF model was selected primarily for its robustness in handling complex datasets with potential non-linear relationships and interactions. The RF

operator consisted of estimating the residual soil water storage between land covers (S_{res}) based on the residual of the EVI between land covers ($\mathrm{EVI_{res}}$), and the day of the year (DOY):

$$RF = \hat{\mathrm{S}}_{res} = f(EVI_{res}, DOY) \quad (1)$$

Both residual terms can be expressed as:

$$\mathrm{S_{res}} = \mathrm{S_{grass}} - \mathrm{S_{crop}} \quad (2)$$

$$\mathrm{EVI_{res}} = \mathrm{EVI_{grass}} - \mathrm{EVI_{crop}} \quad (3)$$

where $\mathrm{S_{grass}}$ is the soil water storage in the top 50 cm under grassland vegetation for a particular day of the year, $\mathrm{S_{crop}}$ is the soil water storage in the top 50 cm under cropland, $\mathrm{EVI_{grass}}$ is the EVI value of grassland, and $\mathrm{EVI_{crop}}$ is the EVI value for cropland.

The trained RF regression model consisted of 200 decision trees and a minimum of three samples required to split an internal node. It was trained on the calibration data using the previous 60 days of vegetation residuals and the day of the year (DOY) for the purpose of estimating soil water storage residuals. RF ensemble approach, which aggregates the predictions of multiple decision trees, offers a natural resistance to over-fitting, especially when dealing with intricate soil moisture dynamics.

The RF model was validated using the CR2-VAL rotation that consisted of winter wheat followed by double-crop soybean (*Glycine max* L. Merr.), and the observations from the NEON KONA site. Crop soil moisture $\hat{\theta}_{crop}$ was estimated by using the observation operator $\hat{\mathrm{S}}_{res}$ predicted using the trained RF model shown in Equation 3 and the observed Kansas Mesonet station soil water storage observation θ_{grass}:

$$\hat{\mathrm{S}}_{crop} = \mathrm{S}_{grass} - \hat{\mathrm{S}}_{res} \quad (4)$$

The accuracy of the predicted soil water storage using the RF observation operator was evaluated using the Root mean squared error (RMSE, mm):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_{crop,i} - \theta_{crop,i}\right)^2} \quad (5)$$

where N is the number of observations, $\theta_{crop,i}$ is the estimated cropland soil water storage for the ith date and $\theta_{crop,i}$ is the observed cropland soil water storage for the ith date.

Additionally, the Mean Bias Error (MBE) was also calculated as:

$$\text{MBE} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_{crop,i} - \theta_{crop,i}\right) \quad (6)$$

Results and discussion

The proposed Random Forest (RF) observation operator successfully assimilated the soil water storage observation from the adjacent ASB station and vegetation residuals to estimate cropland soil water storage. For the CR2 validation site the root mean squared error RMSE (%) was 27.1 mm and the MBE error was of -7.1 mm (Figure 4-4). The validation at the cropland field NEON KONA resulted in a RMSE (%) of 26.7 mm and MBE of 6.2 mm (Figure 4-5). These results align well with a previous study that obtained RMSE (%) between 24.3 mm and 34.4 mm with MBE ranging from -6.6 mm to 5.7 mm, although they were obtained when analyzing plant available water (Patrignani & Ochsner, 2018).

As reported similar studies, the predicted cropland soil water storage after applying the RF observation operator were more representative of the cropland soil water storage than the raw measurements from the adjacent Ashland Bottoms Kansas Mesonet station in both validation sites (Han et al., 2012; Patrignani & Ochsner, 2018). During periods of rapid crop growth in the cropland rotation, the vegetation was denser than the surrounding grass at the Kansas Mesonet station. In these situations, the operator reduced the estimated cropland soil water storage. Conversely, during fallow periods when the natural grass of the Kansas Mesonet station was more developed, the operator positively adjusted the estimated cropland soil water storage.

The NEON KONA site exposed the RF operator to new scenarios, including a distinct rotation scheme (Figure 4-3), this site exhibited more pronounced vegetation development during the 2022 summer season compared to the other training and validation sites (Figure 4-6).

On average, the vegetation in the CR showed a lag of 15 days between high vegetation development and soil moisture depletion. Rotations with winter wheat or summer crops presented higher EVI values than the observed at the ASB station, this was also associated with

posterior higher soil moisture depletions. Long fallow periods observed during winter at rotation CR1 to CR3 and winter fallows as seen in rotation NEON KONA, showed similar values of soil water storage compared to the ASB observations (Figure 4-6).

The RF ensemble model parameters provides a clear understanding of the importance of each feature in the predictions. As shown in Table 4-1, different periods of the 60 days window play a significant role in predicting the soil water storage operator. Day of Year (DOY) stands out as the second most important feature, highlighting the differences in the operator's response to vegetation residuals throughout the year. On average, there is a lag of 7 days (Figure 4-6) between the vegetation differences and the soil water storage operator, suggesting that the impact of different land covers is not immediate.

The results from this study have significant implications for agricultural practices. Accurate soil moisture estimates can guide irrigation decisions, optimize water use, and improve crop yield predictions. The ability of the observation operator to translate grassland soil moisture into cropland soil moisture can be particularly beneficial for farmers, who typically do not have soil moisture sensors in their fields (Mpanga & Idowu, 2021). By leveraging data from nearby environmental monitoring stations and vegetation differences between intermixed land covers, it was possible to estimate the soil water storage for adjacent croplands.

Conclusion

This study evaluated the potential of developing an observation operator using machine learning (Random Forest model) to bridge the gap between grassland and cropland soil moisture dynamics. While the model demonstrated promising results in our specific context, it is crucial to acknowledge the inherent challenges associated with predicting soil moisture dynamics. Factors such as unexpected rainfall events, irrigation practices, and soil heterogeneity can introduce variability that the model may not fully capture when estimating adjacent soil water storage conditions.

Furthermore, the random forest model performance should be validated across different geographical regions, soil types, and land uses, emphasizing the need of in situ soil moisture observations for different crop rotations. As the availability of environmental networks continues to expand, leveraging advanced techniques and methodologies to extract useful insights from the collected data becomes essential. Future research could focus on integrating additional data sources, such as soil texture information, to enhance prediction accuracy further. There is also potential to explore other machine learning algorithms or hybrid models that combine the strengths of multiple approaches.

References

Baier, W., & Robertson, G. W. (1968). The performance of soil moisture estimates as compared with the direct use of climatological data for estimating crop yields. Agricultural Meteorology, 5(1), 17–31. https://doi.org/10.1016/0002-1571(68)90020-4

Crow, W. T., Chen, F., Reichle, R. H., Xia, Y., & Liu, Q. (2018). Exploiting Soil Moisture, Precipitation, and Streamflow Observations to Evaluate Soil Moisture/Runoff Coupling in Land Surface Models. Geophysical Research Letters, 45(10), 4869–4878. https://doi.org/10.1029/2018GL077193

De Lannoy, G. J. M., Houser, P. R., Verhoest, N. E. C., Pauwels, V. R. N., & Gish, T. J. (2007). Upscaling of point soil moisture measurements to field averages at the OPE3 test site. Journal of Hydrology, 343(1–2), 1–11. https://doi.org/10.1016/j.jhydrol.2007.06.004

Denmead, O. T., & Shaw, R. H. (1960). The Effects of Soil Moisture Stress at Different Stages of Growth on the Development and Yield of Corn 1. Agronomy Journal, 52(5), 272–274. https://doi.org/10.2134/agronj1960.00021962005200050010x

Drusch, M. (2005). Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. Geophysical Research Letters, 32(15), L15403. https://doi.org/10.1029/2005GL023623

Gao, X., Zhao, X., Brocca, L., Pan, D., & Wu, P. (2019). Testing of observation operators designed to estimate profile soil moisture from surface measurements. Hydrological Processes, 33(4), 575–584. https://doi.org/10.1002/hyp.13344

Han, E., Heathman, G. C., Merwade, V., & Cosh, M. H. (2012). Application of observation operators for field scale soil moisture averages and variances in agricultural landscapes. Journal of Hydrology, 444, 34–50. https://doi.org/10.1016/j.jhydrol.2012.03.035

Heathman, G. C., Cosh, M. H., Merwade, V., & Han, E. (2012). Multi-scale temporal stability analysis of surface and subsurface soil moisture within the Upper Cedar Creek Watershed, Indiana. Catena, 95, 91–103. https://doi.org/10.1016/j.catena.2012.03.008

Mpanga, I. K., & Idowu, O. J. (2021). A Decade of Irrigation Water use trends in Southwestern USA: The Role of Irrigation Technology, Best Management Practices, and Outreach Education Programs. Agricultural Water Management, 243, 106438. https://doi.org/10.1016/j.agwat.2020.106438

Ochsner, T. E., Cosh, M. H., Cuenca, R. H., Dorigo, W. A., Draper, C. S., Hagimoto, Y., Kerr, Y. H., Larson, K. M., Njoku, E. G., Small, E. E., & Zreda, M. (2013). State of the Art in Large-Scale Soil Moisture Monitoring. Soil Science Society of America Journal, 77(6), 1888–1919. https://doi.org/10.2136/sssaj2013.03.0093

Patrignani, A., Knapp, M., Redmond, C., & Santos, E. (2020). Technical Overview of the Kansas Mesonet. Journal of Atmospheric and Oceanic Technology, 37(12), 2167–2183. https://doi.org/10.1175/JTECH-D-19-0214.1

Patrignani, A., & Ochsner, T. E. (2018). Modeling transient soil moisture dichotomies in landscapes with intermixed land covers. Journal of Hydrology, 566, 783–794. https://doi.org/10.1016/j.jhydrol.2018.09.049

Rodriguez-Iturbe, I., Porporato, A., Laio, F., & Ridol, L. (2001). Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress I. Scope and general outline. Advances in Water Resources, 11.

Wells, M. S., Reberg-Horton, S. C., & Mirsky, S. B. (2014). Cultural strategies for managing weeds and soil moisture in cover crop based no-till soybean production. Weed Science, 62(3), 501–511. https://doi.org/10.1614/WS-D-13-00142.1

Table 4-1. Random Forest (RF) Feature importance. Daily values of vegetation residuals from the 60 days window passed to the model were grouped by week for clarity purposes. DOY feature represents the day of the year (0-365).

Feature	Importance (%)
Week 4	19.75
DOY	16.70
Week 7	13.08
Week 9	12.00
Week 6	8.96
Week 8	6.60
Week 0	6.29
Week 5	5.99
Week 1	5.59
Week 2	5.01

Figure 4-1. Orthoimage illustrating the location of the Ashland Bottoms station (ASB) of the Kansas Mesonet, the adjacent plot experiment that was used for both calibration and validation of the observation operator, and the KONA cropland site of the National Ecological Observatory Network (NEON KONA) that was also used for validation of the observation operator.

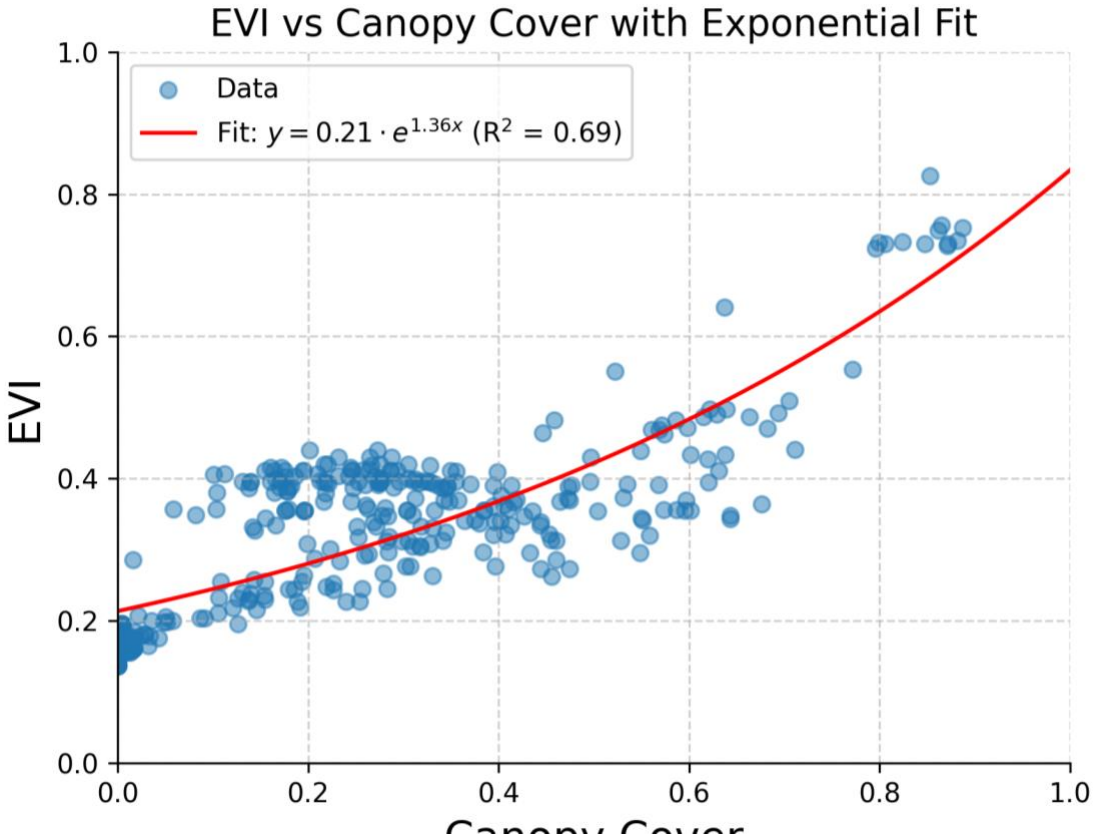


Figure 4-2. Exponential model used to estimate EVI from canopy cover observations. The dataset consisted of 380 georeferenced images in 6 different cropland field across Kansas.

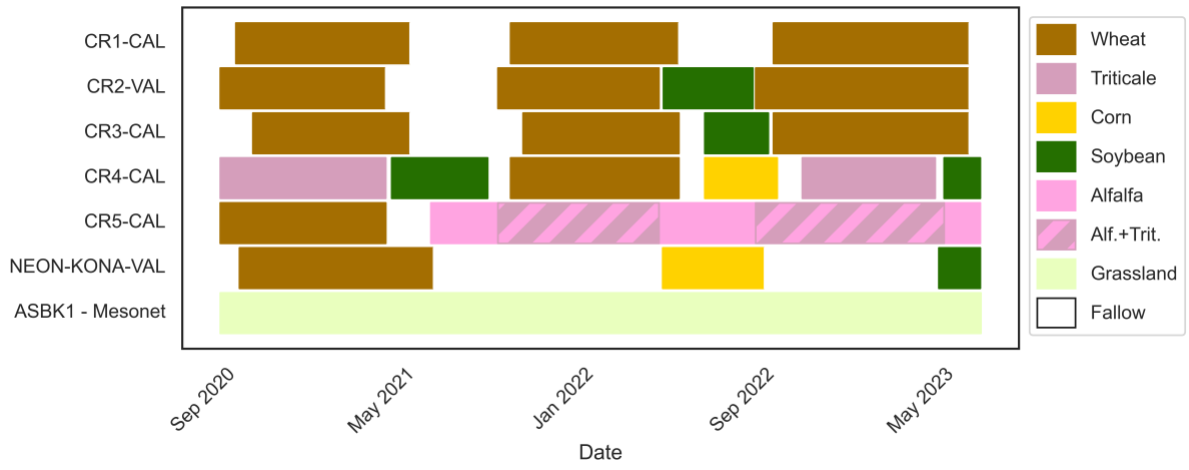


Figure 4-3. Timeline showing the different land covers for each crop rotation of the replicated plot experiment (CR1 – CR5), the KONA site of the National Ecological Observatory Network (NEON), and the Ashland Bottoms station of the Kansas Mesonet.

.

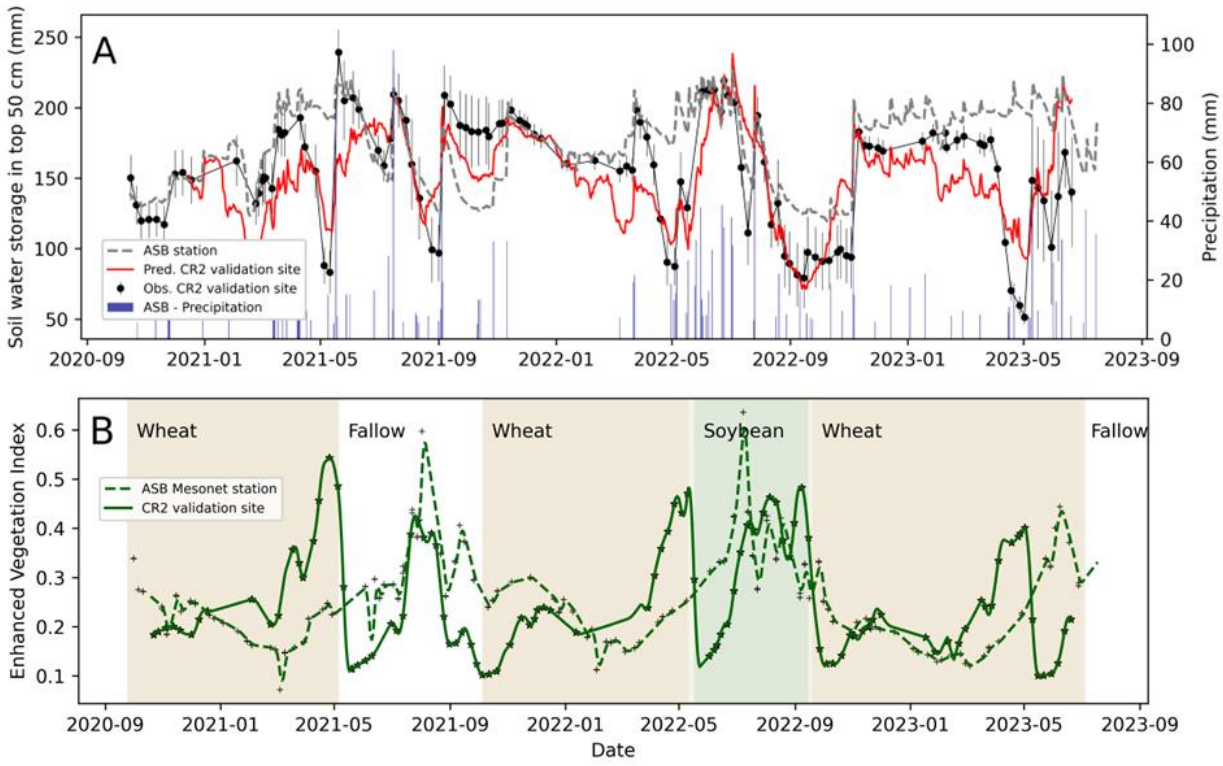


Figure 4-4. A) Daily precipitation, observed soil water storage for the ASBK1 station of the Kansas Mesonet, observed cropland soil moisture for the CR2 validation rotation of the replicated plot experiment, and predicted soil water storage for the rotation site using the random forest observation operator. Depicted soil water storage considers the top 50 cm of the soil profile from October 2020 to August 2023. B) enhanced vegetation index (EVI) time series of observed values in the ASB station and CR2 validation rotation.

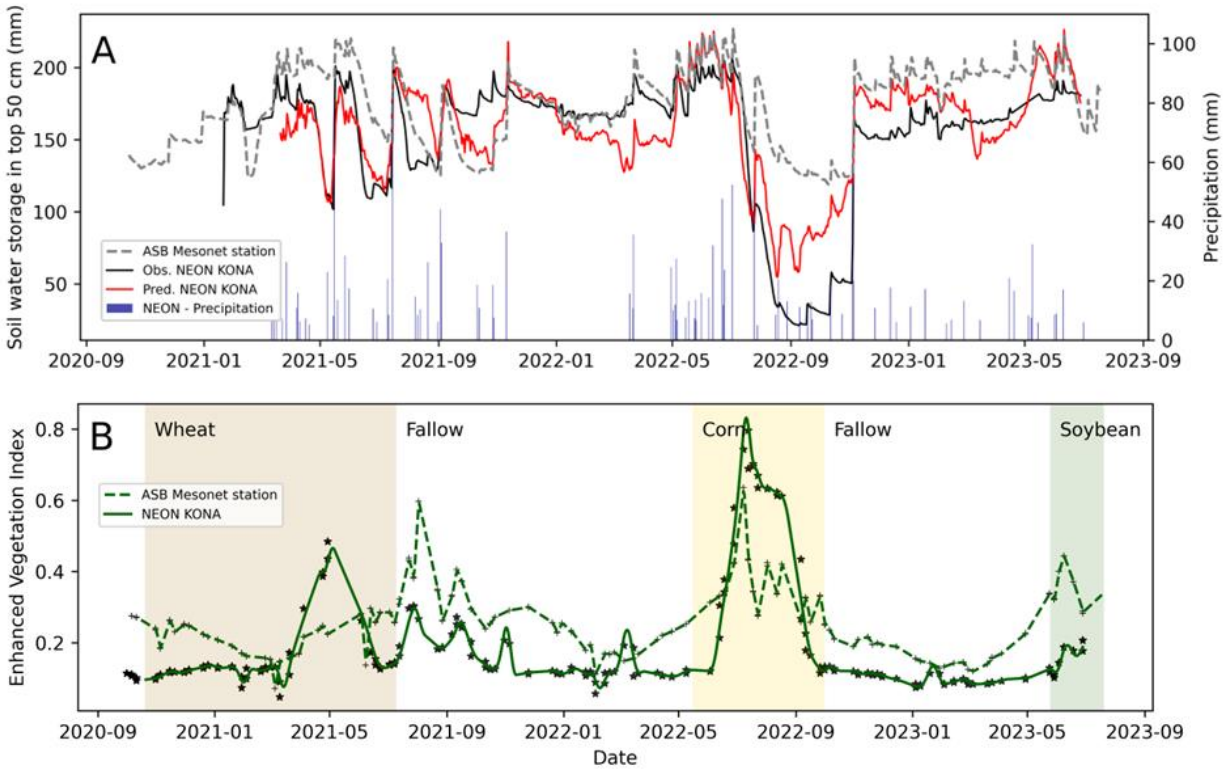


Figure 4-5 A) Daily precipitation, observed soil water storage for the ASB station of the Kansas Mesonet, observed cropland soil moisture for the KONA station of the National Ecological Observatory Network (NEON), and predicted soil water storage for the NEON KONA site using the random forest observation operator. Soil water storage is for the top 50 cm of the soil profile from September 2020 to August 2023. B) enhanced vegetation index (EVI) time series of observed values in the ASB station and NEON KONA validation rotation.

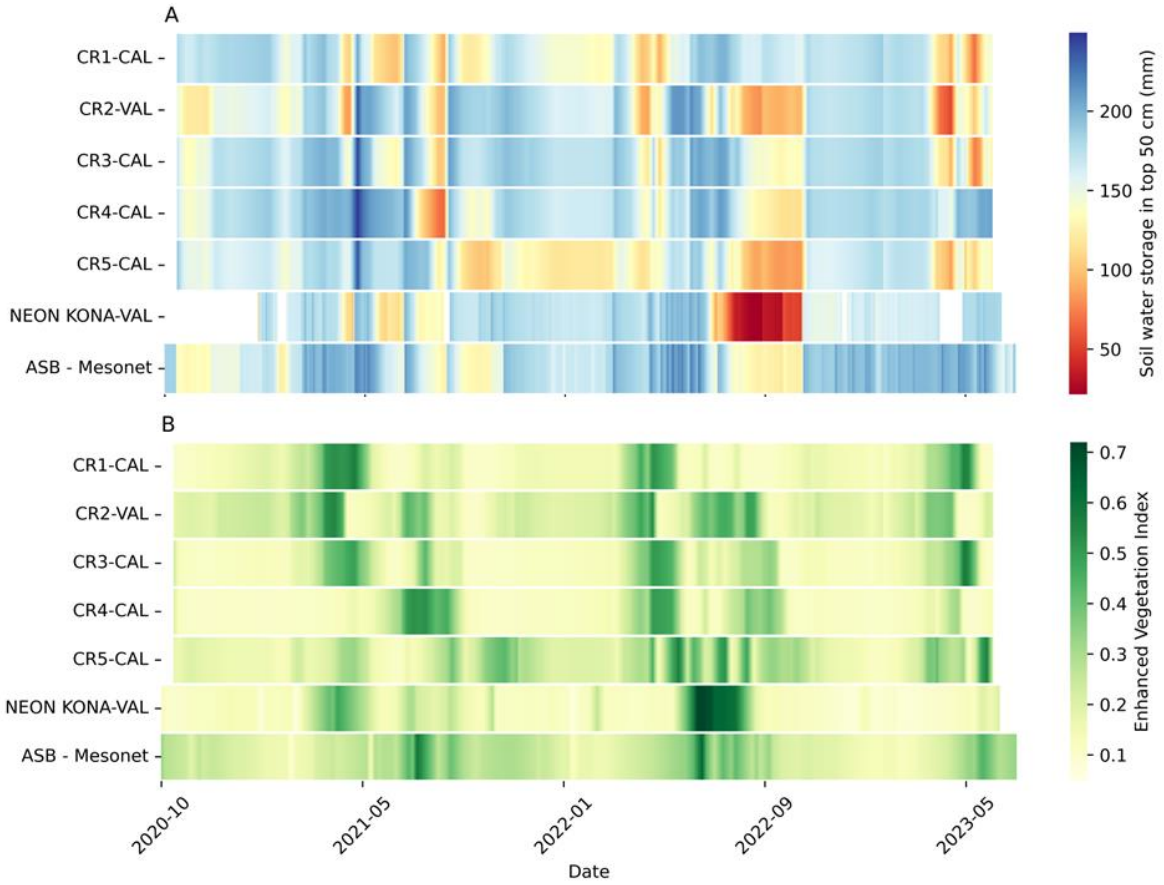


Figure 4-6. Timeline of the observed values of soil water storage (A) and Enhanced Vegetation Index (B) dynamics for each crop rotation of the replicated plot experiment (CR1 - CR5), the KONA site of the National Ecological Observatory Network (NEON), and the Ashland Bottoms station of the Kansas Mesonet from October 2020 to August 2023.

Chapter 5 - General conclusion

There is a growing number of mesoscale environmental monitoring networks (also known as mesonets) that include *in situ* soil moisture observations that can be used for a wide range of meteorological, ecological, hydrological, and agricultural applications. Mesoscale networks provide accurate and high-temporal resolution point-level soil moisture information, but mesoscale networks are sparse and leave large gaps of unmonitored area between stations. This dissertation is centered on leveraging the existing infrastructure of the Kansas Mesonet by upscaling point-level observations to represent statewide soil moisture conditions. The specific question that we tried to answer in this thesis is: *can we upscale point-level soil moisture observations from sparse monitoring networks by integrating spatial model estimates and in situ observations through data assimilation*?

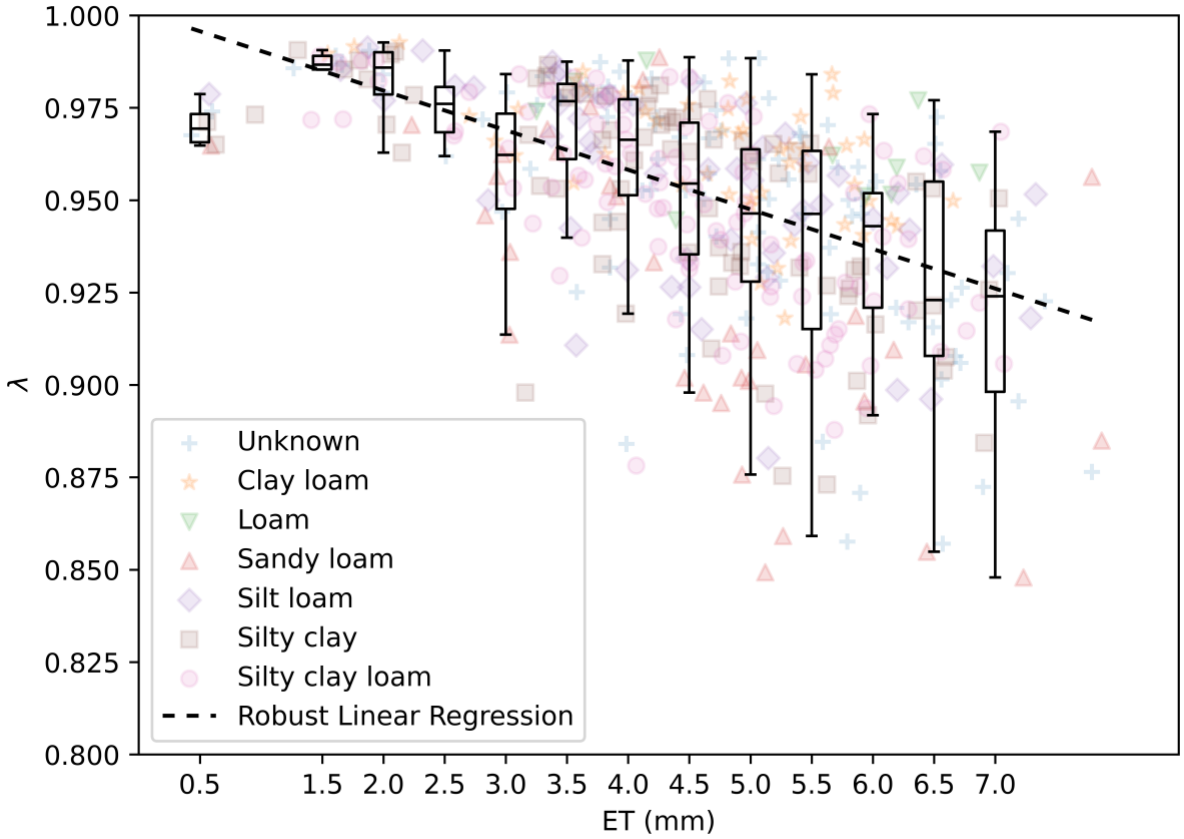
The first chapter of this thesis was aimed at developing a 250-m spatial resolution map of rootzone soil moisture across the state of Kansas by integrating sparse *in situ* observations from 54 stations of the Kansas Mesonet with spatial soil moisture estimates generated with a simple drydown model using the conditional merging assimilation method. The model-data assimilation approach reduced the error in rootzone soil moisture estimates by approximately 27% compared to the model alone and by 40% compared to interpolating the *in situ* observations alone using ordinary kriging. The resulting high-spatial resolution soil moisture maps were validated using surveys made with a roving cosmic-ray neutron probe, which resulted in a median absolute error of 16.1 mm. The validation step also revealed some discrepancies between the rover surveys and the data-assimilation method, suggesting that in areas with intermixed land covers *in situ* soil moisture observations from mesoscale networks may not accurately reflect that of nearby cropland fields.

As part of the validation process of the soil moisture maps in the previous chapter, we also developed a new Python library called Cosmic-Ray Neutron Python (CRNPy) that compiles correction and filtering routines published in >12 peer-reviewed articles since the inception of this technology in the late 2000s. To illustrate the application of the newly developed library, we tested different correction routines and filtering steps that remained unanswered in the scientific literature. The first case-study scenario consisted of computing correction factors using readily available information from the nearest station of the Kansas Mesonet instead of using collocated observations, which resulted in a mean absolute error of 0.037 m^3 m^{-3}, suggesting that acceptable soil moisture estimates could be obtained using meteorological variables from the Kansas Mesonet, but that collocated observation are required for greatest accuracy. The second case-study scenario consisted of applying a filtering step to the resulting time series of volumetric water content rather than the corrected neutron counts, which resulted in the same mean absolute error of 0.029 m^3 m^{-3}. Similarly, the third case-study scenario consisted of applying a spatial average filtering step to the volumetric water content measured with a roving device rather than the corrected neutron counts, which in this case resulted in a slightly greater root mean squared error of 0.023 m^3 m^{-3} compared to a value of 0.019 m^3 m^{-3} when applying the correction to the corrected counts.

The last study was motivated by the challenges in capturing soil moisture conditions in cropland areas using soil moisture from nearby stations of the Kansas Mesonet observed in the second chapter. To translate soil moisture conditions from grassland to nearby cropland we tested an observation operator based on a Random Forest machine learning model. Our findings concluded that by leveraging differences between grassland and cropland vegetation conditions using a remote sensing vegetation index, such as the enhanced vegetation index, soil moisture in

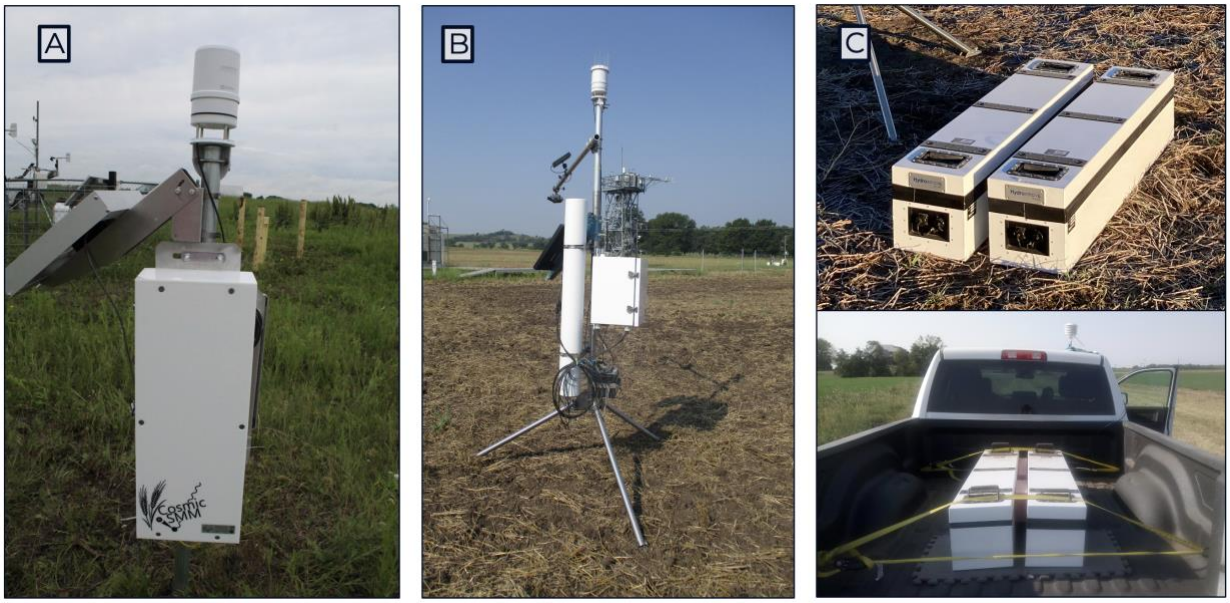
adjacent cropland fields can be estimated with an error of 27 mm, underscoring the potential of using machine learning models for upscaling soil moisture information in landscapes with intermixed land cover.

Appendix A - Reference ET vs. λ



Appendix A Figure A.1. Reference ET negative linear relationship with the soil water storage loss coefficient (λ) for the 485 drydown events analyzed. Different soil textural classes represented with different markers were not related with the λ parameter value.

Appendix B - CRNP devices



Appendix A Figure A.2. A) Stationary detector manufactured by Radiation Detection Technologies, Inc. (Manhattan, KS). B) Stationary detector manufactured by Hydroinnova, Inc. (Albuquerque, NM). C) Roving detector manufactured by Hydroinnova, Inc. (Albuquerque, NM).