

Plants as Proxies: Mapping Hydrophytic Intrinsic Potential (HIP) from Community Science
Observations as a Soil Moisture Indicator

Rachel Buchler

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2026

Committee:

L. Monika Moskal

David E. Butman

Meghan Halabisky

David D'Amore

Program Authorized to Offer Degree:
College of the Environment
Environmental and Forest Sciences

©Copyright 2026
Rachel Buchler

University of Washington

Abstract

Plants as Proxies: Mapping Hydrophytic Intrinsic Potential (HIP) from Community Science
Observations as a Soil Moisture Indicator

Rachel Buchler

Chair of Supervisory Committee:

L. Monika Moskal

Department of Environmental and Forest Sciences

Soil moisture is a critical environmental variable, with ties to forest resilience, plant community composition, and carbon storage. Increasing drought severity and hydrological fluctuations will impact forest health, altering the spatial distribution and temporal dynamics of soil moisture, underscoring the urgency of soil moisture mapping. However, modeling and mapping soil moisture is challenging in heavily forested environments where traditional remote sensing methods cannot penetrate the canopy. This is confounded by the challenges of in situ soil moisture monitoring, stemming from the high temporal and spatial heterogeneity of soil moisture. This gap leaves land managers without the fine-scale soil moisture observations necessary for management-level decision-making. Therefore, we aimed to model surface soil moisture at a resolution of 10m or less in the Skykomish watershed. First, we tested the feasibility of using a lidar-based wetland intrinsic potential (WIP) model to map surface soil moisture. The WIP explained 60% of the variance in surface soil moisture and was calibrated to estimate Volumetric Water Content (VWC). Second, building on these results, we introduce a novel method for modeling soil moisture by coupling community science plant observations with wetland indicator status. This new approach explained 62% of the variance in summer soil

moisture. Third, we investigated where these models systematically over- or underpredict moisture using soil physical properties data. We found no significant link between model performance and soil physical properties; however, the plant-based model systematically underpredicts moisture at the wettest end of the moisture gradient. We demonstrate the promise of wetland mapping tools and community science for soil moisture applications, thereby furthering our ability to model the full moisture gradient at fine spatial scales in densely forested landscapes.

Table of Contents

Table of Contents	i
List of Figures	iii
List of Tables.....	iii
List of Equations	iii
1. Introduction, Significance, and Background	1
1.1 Soil Moisture	2
1.2 Importance of Soil Moisture	3
1.3 Remote Sensing of Soil Moisture.....	4
1.4 Plants as Bioindicators for Moisture	8
1.5 Goals and Objectives.....	10
2. Assumptions.....	12
3. Study Area	14
4. Methods.....	16
4.1 Field Data Methods	16
4.1.1 Field Methods Introduction	16
4.1.2 Site and Plot Selection Methods.....	17
4.1.3 Soil Moisture Data Collection	18
4.1.4 Soil Characteristics Data Collection.....	20
4.2 Modeling Methods	22
4.2.1 Covariates and Data Processing	23
HIP (5m) and Baseline WIP (5m) Covariates.....	23
HIP (10m) Covariates	26
TW WIP (3m) Covariates and Data.....	29
4.2.2 WIP Model	31
Baseline WIP (5).....	31
TealWaters WIP (3m)	31
4.2.3 HIP Model	32

HIP 5m	33
HIP 10m	33
4.3 Model Validation, Residual Analysis, and VWC Estimations.....	36
5. Results.....	39
5.1 Field Data Collection Results.....	39
5.2 WIP Results.....	42
TW WIP 3m.....	42
WIP 5m.....	42
5.3 HIP Results.....	43
HIP 5m.....	43
HIP 10m.....	44
5.4 Residual Analysis (RA).....	46
WIP RA results.....	46
HIP RA results.....	47
HIP 5m	47
HIP 10m.....	47
5.5 VWC Estimates from WIP and HIP.....	48
6. Discussion.....	50
6.1 Performance of Models in Predicting Soil Moisture.....	50
6.2 Where Do Models Fail to Act as a Soil Moisture Proxy?.....	58
7. Conclusion	63
7.1 Management Implications and Future Work.....	63
7.2 Concluding Remarks	66
References.....	68
Appendix A. Loss on Ignition for Mineral and Organic Soils.....	82
A 1.1. Mineral Soil LOI.....	82
A 1.1. Organic Soil LOI	85
Appendix B. Rapid Soil Particle-Size Analysis.....	86
Appendix C. Bulk Density Methods.....	88
Appendix D. SoilRx Field Protocol	90
Appendix E. Linear Mixed Effects Model.....	96

Appendix F. Topographic Wetness Index (TWI) in the Skykomish	99
Appendix G. SAR (L-band) for Soil Moisture	103

List of Figures

FIGURE 1. STUDY AREA MAP OF THE SKYKOMISH WATERSHED WITH FIELD SAMPLING LOCATIONS	16
FIGURE 2. MAP OF HIP TRAINING DATA BY WETLAND INDICATOR AND WETLAND REGION FOR WASHINGTON STATE.	36
FIGURE 3. HIP AND WIP MODEL FLOWCHART FOR SOIL MOISTURE MONITORING	39
FIGURE 4: MEAN SUMMER 2025 SOIL MOISTURE ACROSS 36 SITES IN THE SKYKOMISH WATERSHED	40
FIGURE 5. SOIL TEXTURE BY SITE IN THE LOWER SKYKOMISH WATERSHED	41
FIGURE 6. SITE BULK DENSITY (BD) AND LOI RELATIONSHIP WITH FINE BD AND TOTAL BD	42
FIGURE 7. HIP (10M) IN THE LOWER 500M OF THE WATERSHED, ZOOMED IN ON THE SITE MARSH CREEK ORGANIC	46
FIGURE 8. NORMALIZED RESIDUALS FROM THE HIP	48
FIGURE 9. TEALWATERS WIP VWC ESTIMATES ACROSS THE WATERSHED WITH A ZOOMED IN MAP OF MARSH CREEK ORGANIC	50
FIGURE 10. CONCEPTUAL DIAGRAM OF PLANTS AS MOISTURE PROXIES ALONG A MOISTURE GRADIENT	52
FIGURE 11. HIP 10M VWC ESTIMATES	57
FIGURE 12. HIP (5M) THREE CLASS MODEL RESIDUALS AND LOI	59

List of Tables

TABLE 1. HIP (5M) AND BASELINE WIP (5) COVARIATE TABLE	25
TABLE 2. HIP 10M COVARIATE TABLE.	27
TABLE 3. TW WIP 3M COVARIATE TABLE.....	29
TABLE 4. HIP MODELS: TRAINING DATASETS.	34
TABLE 5. LM RESULT TABLE FOR SOIL MOISTURE FROM HIP AND WIP (RESCALED TO 0-100)....	45

List of Equations

EQUATION 1. PARTICLE SIZE CALCULATIONS.	21
EQUATION 2. BULK DENSITY.	21
EQUATION 3. TOPOGRAPHIC WETNESS INDEX (TWI).	25

Acknowledgments

This research was supported by the SEFS McIntire-Stennis project, award no. NI24MSCFRXXXG065 Amend 2, from the U.S. Department of Agriculture's National Institute of Food and Agriculture, and was carried out in the University of Washington Remote Sensing and Geospatial Analysis Laboratory (RSGAL). I would like to thank my supervisory committee, L. Monika Moskal (chair), David Butman, Meghan Halabisky, and David D'Amore, for their guidance and support throughout my Master's. Additionally, thank you to Brittany Johnson for teaching me so much about the fundamentals of soil science. I would also like to thank the science team at Aspen Global Change Institute, particularly Tanya Petach, and Elise Osenga, for their support and guidance in planning my field season and for including me in fieldwork in Colorado, where I had the opportunity to engage with the broader soil moisture community. A huge thank you to all my lab mates and lab alums for their advice and encouragement; in particular, I would like to thank Helen Miller, Keenan Ganz, Anthony Stewart, and Lindsey Skidmore for their help with editing, brainstorming, statistics, lugging soils up and down the stairs of Bloedel, and much more. I would additionally like to thank the TealWaters team, especially Dan Miller, Steven Pestana, and Jesse Anderson, for their technical support and ideas throughout my thesis. Thank you to Arianna Goodman and Jesús M. Rueda-Becerril for helping with fieldwork and geospatial analysis code/concepts. Arianna authored all code used for the TWI analysis. I would like to thank the undergraduate students who worked with me, Skye Gearhart, Kylie Gates, and June Olivier, whose support in the lab and field was integral to the success of this project. I would also like to thank the Pacific Northwest Research Station's Pacific Wildland Fire Sciences Lab for allowing me to use their laboratory space, and the Washington Department of Natural Resources for allowing me access to my field sites. Lastly,

any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author and should not be construed to represent any official USDA or U.S. Government determination or policy. Figures, created in R, QGIS, and BioRender.com.

1. Introduction, Significance, and Background

Water plays a fundamental role in shaping terrestrial landscapes, influencing plant community composition, carbon storage, and ecological function (Porporato et al., 2004). The spatial distribution of moisture across the landscape is therefore a powerful environmental variable, informing processes such as plant assemblage, productivity, and land-atmosphere energy feedbacks (Seneviratne et al., 2010a). However, with projected shifts in precipitation and increases in drought severity, the mosaic of moisture across the landscape will shift, resulting in cascading ecological and human impacts (Crausbay et al., 2020). Washington state has a wide array of moisture regimes, from densely forested upland slopes to forested wetlands and peatlands, and a great deal of this distribution is governed by topographic position (Asada et al., 2003; Beven & Kirkby, 1979; Murphy et al., 2009). Soil serves as the foundation of many terrestrial ecosystem functions and is a key component of infrastructure, agriculture, and, thus, human health. Soil moisture is central to these services, shaping the soil biogeochemical environment and driving plant community composition (Li et al., 2022; Vereecken et al., 2022; Zhang et al., 2019).

Topography alone does not control every environmental variable or condition; however, topography can be thought of as a catalyst for a chain of processes that impact moisture distribution, soil properties, and ultimately soil moisture. The relationship between soil moisture and plant community distribution is not independent, and the control of topography over moisture is modulated by the interplay between soil, plant communities, microbes, and animals

that alter the landscape and moisture distribution over time (Beatty, 1984). However, topography remains a powerful variable for understanding the drivers behind the redistribution of precipitation, groundwater, and runoff into soils. Topography is also a principles variable for mapping soil properties beyond moisture and has been leveraged in the field of digital soil mapping for many years via soil formation mechanistic models and soil spatial predictor models (Jenny, 1941; McBratney et al., 2003).

Soil characteristics are heterogeneous, and a small area may contain multiple soil types spanning multiple scales from differing particle sizes to aggregate shapes and bulk densities. A soil catena refers to the variety of soils and stratigraphy that is commonly found along slopes. Catena in Latin means chain, highlighting the interplay and heterogeneity of soils along a slope. Soil along a slope may have differing characteristics, such as depth and texture, but still be pedagogically related (Borden et al., 2020; Milne, 1947). The concept of a catena can be utilized to isolate the impacts of topography and microtopography on soil moisture by studying soils whose differences can be explained as a function of topography. This concept is also referred to as a toposequence (Jenny, 1941). A soil's position on a toposequence isolates topographic position, holding all other soil formation variables constant, such as parent material and age. Topography is fundamental to soil moisture, as it governs moisture distribution across the landscape, making it a powerful tool for mapping soil moisture.

1.1 Soil Moisture

The main source of freshwater worldwide is driven by surface runoff (Gleick, 1993). Much of this water will either be held in or pass through soil at some point during the hydrological cycle. Water enters the soil primarily via infiltration, where is then held in the soil's

pore space, and is referred to as soil moisture or soil water (Dingman, 2015). The texture, aggregation, structure, and biotic and abiotic environment of the soil will all affect the amount of water stored in the soil and the location within the profile where it is present. There is a clear relationship between soil texture and water-holding capacity (field capacity); texture controls moisture content, and water movements through the profile (Ratliff et al., 1983). For example, soils with a coarser texture will have a lower field capacity, and water will flow through them more freely than in soils with higher clay and silt contents (Abrol et al., 1968; Ratliff et al., 1983).

Soil moisture is highly temporally and spatially variable (Peterson et al., 2019). At Fine spatial scales much of this variability is tied to topography and vegetation. On a landscape scale, soil moisture patterns are more closely tied to climate. Additionally, the distribution of moisture across the landscape varies not only laterally but also vertically (Dingman, 2015). Depth plays a role in the accumulation of water in the soil profile; not all horizons will be equally saturated. Studies have found that surface soil moisture is less variable than deeper-layer moisture during dry weather, suggesting decoupling of surface soil moisture from temperature under dry conditions, with greater variability in the root zone (Hirschi et al., 2014).

1.2 Importance of Soil Moisture

In a rapidly changing climate with subsequent hydrological regime changes, the relationship between water and soil is crucial for many ecosystem functions, given that soil moisture governs many processes in terrestrial ecosystems (Crausbay et al., 2020; Seneviratne et al., 2010b; Vereecken et al., 2022). Soil moisture's inherent heterogeneity can be challenging for land managers to account for accurately. Additionally, many soil moisture maps are at coarse scales on the order of kilometers. With projected increases in drought severity, it is increasingly

important to understand soil moisture distributions at fine scales relevant to management-level decision-making (Crausbay et al., 2020).

Across industries ranging from agriculture to forestry, fine-scale soil moisture data support operational planning and ecological protection (Ford et al., 2015; Murphy et al., 2008; Radmehr et al., 2026; Rowland & Lovelock, 2024). For example, forestry operations rely on non-saturated conditions to avoid soil degradation and infrastructure integrity (Murphy et al., 2009). Forestry implications are especially urgent in the Pacific Northwest, where forestry is a prominent industry. Fine-scale soil moisture modeling would allow land managers to factor wetness into operations planning. Additionally, advancements in the capacity to map soil moisture at fine scales will enable researchers to understand ecosystem connectivity, ecological transition zones, and possible areas of concern in the face of increasing aridity.

1.3 Remote Sensing of Soil Moisture

Soil moisture is monitored via in situ measurements and satellite remote sensing. In situ monitoring networks can be costly and challenging to interpolate over a broader landscape, given the spatiotemporal heterogeneity of soil moisture. Ground-based soil moisture monitoring was predominantly developed for field-scale agricultural tracking, and thus, remote sensing is more widely used for watershed-scale monitoring. However, remote sensing is more recently being adapted for field-scale soil moisture monitoring for precision agriculture (Robinson et al., 2008; Sreeram et al., 2026). Various methods and metrics are used to quantify and classify soil moisture, including volumetric water content, gravimetric water content, soil matric potential, and soil water potential.

The widely used volumetric water content is the ratio of the volume of water in a soil to the volume of soil, while the gravimetric method is calculated as the difference in mass between

a wet and a completely dried sample. The gravimetric method is widely used for its accuracy and relative accessibility, as it does not require expensive equipment and can be performed with just an oven and a scale (Rasheed et al., 2022). Volumetric water content is also widely utilized and can be measured with various methods, including neutron probes (Gardner & Kirkham, 1952), Time Domain Reflectometry (TDR) (Topp et al., 1980), Capacitance and Frequency Domain Reflectometry (FDR) (Gaskin & Miller, 1996), Gamma-Ray Attenuation (Moreira et al., 2007), Ground Penetrating Radar (Huisman et al., 2003), and Cosmic Ray methods (Zreda et al., 2012). For the purposes of this work, we focused on Volumetric Water Content (VWC), as it is widely used in the soil sciences due to its relative cost-effectiveness and quick application (see the appendix for more context on TDR).

Current long-term ground-based soil moisture networks in Washington state are very coarse, and the lack of in situ observations hinders understanding of the spatial distribution of soil (Rodriguez, 2023; Seneviratne et al., 2010a). Established in 2011, the National Coordinated Soil Moisture Monitoring Network ([NCSMMD](#)) is a centralized platform for soil moisture data and information, bringing together networks across the US (Tavakol et al., 2021). The network includes stations from SNOwpack TELemetry (SNOWTEL), the National Ecosystem Observation Network (NEON), and sites operated by the United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) including the NRCS/United States Forest Service (USFS) National Forest Soil Moisture Monitoring Network (FSMMN). Although these data provide valuable insights into the spatiotemporal distribution of soil moisture, they are less representative of management- or watershed-level soil moisture patterns due to their sparse spatial coverage. The United States has the most sites worldwide; in 2021, there were 1,353 sites between 13 networks spanning the contiguous United States

(Tavakol et al., 2021). However, soil moisture data coverage at the watershed scale can be very sparse or completely absent.

Given that spatially continuous data is challenging to acquire with in-situ methods, remote sensing has been widely utilized for soil moisture monitoring. The utilization of remote sensing in soil moisture research has steadily increased since the 1990s, with a more recent spike in high-resolution soil moisture remote sensing research. Numerous products have been developed to remotely sense soil moisture using methodologies such as machine learning, multispectral remote sensing, image classification, and the development of mechanistic models. Still, many lack sufficient spatial resolution for management-level decision-making, or are less effective in forested landscapes due to the sensor's inability to penetrate the forest canopy.

Soil moisture passive remote sensing products are available via satellite and airborne multispectral data. In the US, the National Agricultural Imagery Program (NAIP) provides high-resolution multispectral imagery nationwide; however, these data are only collected every 3 years, which limits their effectiveness for soil moisture monitoring given the high temporal variability of soil moisture. Spaceborne multispectral imagery is readily available via the National Aeronautics and Space Administration (NASA), the European Space Agency (ESA), and other international and private vendors. These data can be utilized to monitor landscape wetness, particularly for mapping areas on the wetter side of the spectrum, such as wetlands, using indices like the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI) (Ashok et al., 2021; Zeng et al., 2022).

However, Optical and thermal remote sensing have weak explicit ties to soil moisture and are impacted by challenges such as cloud cover and overstory vegetation. These limitations make them less widely adopted for soil moisture monitoring, especially in forested landscapes with

dense canopies and steep complex topography (Peng et al., 2021). Although, this does not discount the ability of vegetation indices to improve machine learning and modeled soil moisture predictions. Recent studies using machine learning have demonstrated the effectiveness of including multiple types of remote sensing data (Amazirh et al., 2018; Guo et al., 2024; Tyystjärvi et al., 2022; Whiting et al., 2004).

Radar remote sensing has demonstrated its ability to map soil moisture, as radar can detect soil dielectric properties, making it well-suited for mapping soil moisture. Radar-based products such as NASA's SMAP offer soil moisture estimates at global extents but at spatial resolutions of 3–36 km, making them unsuitable for watershed-scale management that requires high-resolution observations, such as timber harvest and water management (Chan et al., 2016; Peng et al., 2021). Synthetic Aperture Radar (SAR) offers finer resolution but performs poorly beneath dense forest canopies and in complex terrain (Bonnell et al., 2024) see Appendix G. SAR (L-band) for Soil Moisture for more information. These limitations underscore the need for alternative approaches in heavily forested landscapes.

Lidar data also offers insights into soil moisture and hydrological modeling; topographically derived metrics such as the Depth to Water (DTW) and Topographic Wetness Indexes (TWI) can be used as soil moisture proxies (Buchanan et al., 2013; Murphy et al., 2008, 2009; Riihimäki et al., 2021; Southee et al., 2012). TWI in particular is a widely utilized proxy for soil moisture, although the literature indicates wide variance in TWI's ability to act as a soil moisture proxy depending on the TWI calculation methods (Buchanan et al., 2013; Kopecký et al., 2021; Riihimäki et al., 2021; Southee et al., 2012; Winzeler et al., 2022). Aside from serving as standalone proxies, topographic variables are commonplace in soil moisture modeling. An effective example of this is illustrated in a study in Sweden where researchers used high-

resolution lidar-derived Digital Terrain Model (DTM) and machine learning methods to map 2-meter soil moisture at the national level (Ågren et al., 2021).

Methods such as these are not only applied to soil moisture mapping but also to wetland mapping. Wetlands are identified by three main factors: vegetation, hydrology, and hydric soils. Topography plays a significant role in governing these indicators, and thus is an effective tool in identifying wetland area (Cowardin et al., 1979). The Wetland Intrinsic Potential (WIP) tool is a peer-reviewed wetland probability model, that aims to model wetland intrinsic potential as a function of variables associated with wetland occurrence, such as topography (Halabisky et al., 2023). WIP was developed with wetlands in mind, this research tested its ability to illustrate overall wetness trends, such as soil moisture. The limitations in remote sensing of soil moisture are largely driven by a lack of in situ moisture observations for model development and validation, canopy obstruction of soil, and the inability of remote sensing methods, such as multispectral analysis, to detect soil moisture signals. These limitations underscore the need for alternative approaches to soil moisture mapping in densely forested landscapes. One such alternative leverages biological indicators of moisture, such as plant communities, to help model the spatial distribution of soil moisture.

1.4 Plants as Bioindicators for Moisture

The coupling of abiotic and biotic factors drives plant community distribution; both climatic variables and soil characteristics, such as soil moisture, play a key role in plant occurrence (Chan et al., 2016). Local topography and landscape position strongly control site hydrology, shaping the environmental attributes of a given site. Thus, plant occurrence can be attributed to their environmental preferences, and may act as proxies for these environmental

factors (Bisbing et al., 2016; Keddy, 2017). The use of plants as proxies is a common approach in ecology and has been applied to landscape classification and delineation across fields (Hajek & Knapp, 2022; Wang et al., 2019).

The relationship between soil moisture and plants is inherently linked via coupled feedback loops. For example, throughfall and stemflow are governed by plant community traits and will impact soil moisture conditions, as these are fundamental mechanisms of water input to the soil profile. In turn, soil moisture influences vegetation communities, altering these mechanisms of water input (Wang et al., 2019). Plant community composition is impacted by soil properties (Beatty, 1984), and micro topography and landscape position further control species composition and diversity (Hutchinson et al., 1999; Vivian-Smith, 1997).

An effective example of plants as bioindicators are the Ellenberg indicator variables (EIV). The EIVs use a numerical scale, and assess 7 abiotic environmental factors (light, temperature, continentality, moisture, soil reaction, nutrient content, and salinity (Ellenberg, 1974). These indicators are widely adopted throughout Europe to infer environmental factors without direct measurements, enable new plant ecological research, and even to quantify disturbance (Tichý et al., 2023).

The EIVs were developed for European flora and are not directly transferable to the United States, given the fundamental botanical differences. Although EIVs have been widely used across multiple disciplines in Europe, there are limitations and criticisms regarding their effectiveness as bioindicators of soil moisture. Research suggests that TWI shows a stronger correlation than EIV in central European forests, indicating that EIVs should be approached with caution as a proxy for soil moisture alone (Raduła et al., 2018). Topography exerts control over landscape attributes such as moisture gradients, and in turn governs community composition,

making it a principal variable in the spatial distribution of soil moisture. The absence of numerical plant indicator systems in US equivalent to the EIVs, combined with evidence that plant indicators alone are insufficient proxies for soil moisture (Raduła et al., 2018), suggests that **integrating plant-based bioindicators with topographic wetness indices may offer a more robust approach to soil moisture mapping in the forested landscapes** of the Pacific Northwest.

Although the numerical EIV methods are not as widely adopted in the United States, other plant classifications have been developed. There are several hierarchical classifications of plant communities, such as the United States National Vegetation Classification (USNVC), which classifies plant communities rather than individual species. The National Wetland Plant List (NWPL), first developed by the US Fish and Wildlife Service (Cowardin et al., 1979), offers plant classifications for wetland delineation. Species are classified based on the percentage they are found within wetlands, into 5 discrete classifications: Obligate (OBL), Facultative Wet (FACW), Facultative (FAC), Facultative Upland (FACU), and Upland (UPL) (Tiner, 1983). Although these classifications were not specifically developed as proxies for soil moisture, they do serve as suitable proxies for moisture. Plant occurrence data is widely available via community science platforms such as iNaturalist, providing spatial coverage that would be cost-prohibitive to achieve under traditional field sampling and monitoring. Community science offers a unique opportunity to scale plant-based moisture indicators across large landscapes where fine-scale soil moisture observations remain sparse.

1.5 Goals and Objectives

In the Pacific Northwest, over 85% of forests are dominated by conifers, resulting in a challenging landscape for remote sensing, as many sensors can't penetrate the dense canopy

(Northwest Climate Hub, n.d.) Although topography alone has a strong influence on the spatial distribution of soil moisture, Light Detection and Ranging (LiDAR) has enabled high-resolution topographic mapping in densely forested landscapes such as the Pacific Northwest. The compounding challenge of dense canopy cover and high spatial variability in soil moisture results in limited high-resolution soil moisture products in heavily forested landscapes.

Given the gap in high-resolution soil moisture distribution maps in heavily forested landscapes, we aimed to explore how topography can be leveraged to model proxies for soil moisture. We posed the question **1. how well do topographically based wetland probability models function as a proxy for surface soil moisture across the full moisture gradient in a heavily forested Pacific Northwest watershed?** Building on our investigation of topographic-based wetland models as soil moisture proxies, we explored novel modeling methods using community science plant occurrence data. Much work has been conducted leveraging machine learning and topography to model soil moisture, but little has been done to test the feasibility of using plant bioindicators to model soil moisture proxies in Western Washington. Given this gap in the literature, we aimed to build on previous work by using plant occurrence data to model potential proxies for soil moisture and illustrate the potential of community science data as a scalable, accessible method for mapping moisture.

Given the inherent link between soil moisture and hydrophytic plant physiology (Ahanger et al., 2021) we investigated the question **2. Do hydrophytic plant indicator models have predictive power for spatial patterns in soil moisture?** Expanding on our investigation of suitable proxies for soil moisture, we aimed to explore where models fail to serve as suitable proxies. We pose question **3. Where do these soil moisture proxies show systematic over- or underprediction, and what soil properties drive disagreement?**

Soil moisture is a powerful climatic variable, and multiple disciplines within ecology, plant physiology, carbon, and hydrology would benefit from better soil moisture models at finer scales. Moreover, land managers across federal, state, and tribal jurisdictions would benefit from advancements in soil moisture mapping, given the correlation between soil moisture and ecological functions and land uses such as agriculture or the harvest of culturally significant plants. **This research advances high-resolution soil moisture mapping in densely forested landscapes by proposing a reproducible, low-cost framework for fine-scale soil moisture proxy mapping that couples community-science plant observations with high-resolution LiDAR-derived terrain metrics.**

To meet this goal, we aimed to:

1. Aid in management- level decision making by providing methods to map soil moisture proxies at fine spatial scales,
2. offer a method within which to monitor soil moisture in the field, and
3. offer insight into the relationship between surface soil moisture proxies and soil physical properties.

2. Assumptions

There are several key assumptions underlying the methods outlined below. First, we assumed that high-resolution topographic data is representative of the ground topography and remains relatively unchanged since data acquisition. Plant occurrence data are sourced from community science observations, with underlying uncertainties in both species identification and GPS location. Here, we assume that species identifications are accurate; this assumption is

supported by research on the quality of iNaturalist observations in the US, which indicates that research-grade observations have an accuracy greater than 80% at the species level (Mesaglio et al., 2025). Geographic location poses uncertainty for community science data, especially in the heavily forested Pacific Northwest, where smartphone GPS accuracy tends to degrade under the canopy. Aside from the GPS itself, we assume that observation locations are correctly placed by users. Apart from the quality of the plant observations, this work also assumes that hydrophytic plant status data is spatially representative enough to capture the full moisture gradient. We were focused on surface soil moisture and use the top 20cm of soil for all analyses since surface soil moisture in the root zone disproportionately affects plant growth. Although a more robust analysis of soil moisture would incorporate moisture throughout the entire profile. However, for the purposes of this work, we assume that the 20cm measure is representative of surface soil moisture dynamics, as seen in similar soil moisture research.

Beyond data quality assumptions, this work is also subject to practical constraints, which limit the sample size and spatial scale of the validation data. We were limited by both time and resources in terms of the location and amount of field data collection, leading to a smaller validation dataset than initially desired. Moreover, conducting fieldwork in the Pacific Northwest (PNW) comes with unique obstacles, given the density of the forest; moving through the densely vegetated PNW forest is challenging and slow, making it difficult to collect field data across large spatial scales. These factors resulted in the soil moisture dataset being spatially and numerically constrained, with 36 soil moisture observations confined to the lower western extent of the Skykomish watershed. Moreover, none of our field plots were sampled in areas of Alluvium so all results should be met with caution in these areas. The temporal nature of soil moisture is also not addressed in the proposed proxy models, given that we don't have adequate

soil moisture data to validate temporal trends, which fall outside the scope of this work. We are most interested in modeling high-resolution spatial variability in soil moisture. For further discussion of limitations of methodological choices, refer to the discussion section.

3. Study Area

We explored the spatial distribution of surface soil moisture in the Skykomish watershed, which was selected for its representation of conifer forests in western Washington. A central goal of this study is to test soil moisture mapping methods in areas with dense forest cover, making the Skykomish an appropriate choice due to its dense conifer landscape. Moreover, the Skykomish watershed is within driving distance of the University of Washington, enabling a larger set of field data points to be collected. In addition to enabling a larger spatially distributed set of field validation points, the location also enabled repeat synoptic sampling, a method vital to understanding full summer soil moisture patterns. Large parts of the Skykomish watershed are actively managed for forestry and drinking water, making it a high priority for soil moisture monitoring.

Our collaboration with the Washington Department of Natural Resources (DNR) influenced our site selection process. The DNR actively manages land within our area of interest for recreation, forestry, and hydrology, and a robust soil moisture map is of interest for managing these lands. Additionally, the City of Everett relies on the Skykomish watershed for the city's water supply; drinking water is sourced from Spada Lake Reservoir, located within the Upper Sultan watershed (City of Everett, 2015). Lastly, the Skykomish watershed is an area of interest in ecological research, in partnership with local tribes, allowing this work to inform future research.

The Skykomish watershed is located on the western side of the Cascade Mountains, extending into the Puget Sound lowlands with elevations ranging from 5 meters to just over 2,400 meters (Figure 1). The watershed is located within Snohomish County, and is characterized by coniferous forests, featuring Western Hemlock (*Tsuga heterophylla*), Western Red Cedar (*Thuja plicata*), and Douglas-fir (*Pseudotsuga menziesii*). The Skykomish River flows westward, flanked by an Alluvial valley bottom with Holocene terraces on either side.

The Cascade Mountains were initially established during the Mesozoic and late Paleozoic ages. During the Eocene era, heavy deposition and accretion occurred as large rivers flowed through the lowlands. The modern-day Cascade Range is relatively young; uplifted about six million years ago or less (Booth et al., 2003). These tectonic factors are crucial in understanding the geomorphology of the Skykomish watershed and the soils that fall within. Later glaciation played a critical role in shaping and modifying the landscape. The most notable glacial interval is the most recent, "Vashon stade of the Fraser glaciation," during this period; ice covered much of the Puget Sound lowlands, including the Seattle area (Booth et al., 2003). This glacial interval helped form the peaks of the Cascade Mountains seen today. The geomorphic history of the Skykomish has a substantial impact on soil properties; the watershed is 40% Volcanic and 27 % Glacial till in terms of soil parent material (Soil Survey Staff, n.d.). The dominant soil order in the Skykomish watershed is Andisols (46%), which form in volcanic ash and other pyroclastic materials. Andisols are characterized by high moisture and nutrient-holding capacity and are often seen in coniferous forests (Soil Science Division Staff, 2017). The Skykomish watershed is representative of Western Washington forested areas and of interest to project partners, making it a suitable study area.

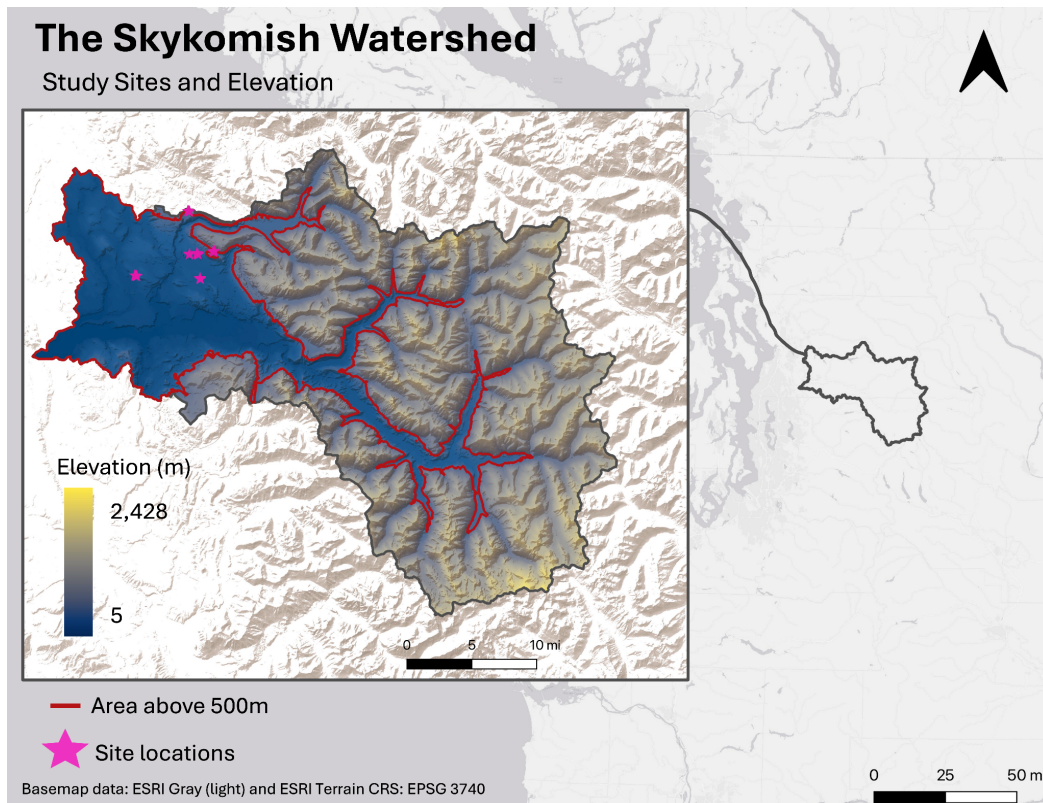


Figure 1. Study Area Map of the Skykomish Watershed with Field Sampling Locations

4. Methods

4.1 Field Data Methods

4.1.1 Field Methods Introduction

Volumetric water content (VWC) varies across both spatial and temporal scales, making it challenging to capture the full range of soil moisture variability with limited time and resources. To address this, field methods were designed to: (1) capture surface soil moisture across the full wet-to-dry gradient, (2) account for seasonal variation within the summer growing season, and (3) remain feasible for reproduction by small field crews, with the goal of informing future soil moisture monitoring efforts. Sampling was conducted across a matrix of site types spanning the wet-to-dry gradient during three 6-day campaigns. These campaigns aimed to

gather information on soil moisture, soil physical properties, to evaluate model performance, and characterize the relationship between soil moisture and physical soil properties.

4.1.2 Site and Plot Selection Methods

In-situ data were collected in the western extent of the Skykomish watershed, within the Lower Sultan River, Woods Creek, Olney Creek, and Wallace River sub-watersheds (Huc 12). All sites were located on Washington Department of Natural Resources (DNR) land. The data were collected along soil catenas at 6 study wetland sites with six plots moving upslope from the wetlands into mesic and upland areas. Resulting in a nested dataset with 6 sites and a total of 36 plots.

All sites were selected before sampling, and targeted areas with high moisture variability. Given the limited amount of high-resolution soil moisture observations and models in the area, the WIP model was used to help identify areas that might exhibit multiple moisture classes. High variability was defined by classifying the WIP probability raster into 6 classes mimicking the 6 Soil Survey Geographic Database (SSURGO) drainage classes present in the study area. From this classified WIP, the wettest class was used to identify wetland sites. The target wetland sites all fell within DNR land, 50 meters from a roadway, and no further than 1 mile from a roadway. After these wetland areas of interest were delineated, each polygon underwent visual quality control using aerial imagery, stand age data, topographic information, and wetland probability data, broken into 6 classes as defined above, to ensure accessibility and overall fit for the aim of this project. Polygons that passed this quality control filter were characterized by the potential for high moisture variability. Stand age was assessed using the DNR RS-FRIS 5.2 Combined Origin Age Dataset (Washington State Department of Natural Resources, 2025) to confirm that the surrounding areas consisted of stands greater than 30 years of age.

From these polygons, a set of 15 points were generated using a spatially constrained random sampling in ArcGIS to reduce spatial sampling bias in site selection. These 15 sites were then assigned priorities: 10 tier-one sites and 5 backup sites, based on their accessibility. During the first week of sampling, 6 sites were visited within the 6-day sampling window; all were tier 1 sites. Plots were established in the field using a digital transect of approximately 300 meters along the azimuth, with the highest moisture variability delineated using the WIP probability raster. Then plots were established based on shifts in hydrology, vegetation, and landscape position (refer to Appendix D. SoilRx Field Protocol). Sampling was structured into three field campaigns, each lasting six days. These campaigns were executed during the summer months of 2025, specifically June, July, and August. These months were selected because they represent the growing season and have been consistently used for similar sampling methods (Kemppinen et al., 2018; Riihimäki et al., 2021; Southee et al., 2012).

4.1.3 Soil Moisture Data Collection

Volumetric Water Content (VWC) was collected at each site throughout the summer months (June 23rd to 28th), July (21st to 26th), and August (25th to 30th). Data was collected with the Field Scout 150 Time Domain Reflectometer (TDR) meter. The Field Scout has an accuracy of $\pm 3.00\%$ VWC when the soil electrical conductivity is at or below 2 mS/cm (Spectrum Technologies, Inc., n.d.). The Fieldscout was selected for this project for its size, speed, and overall accuracy. Although there are some inherent drawbacks to using TDR, particularly in highly organic soils, where TDR is known to underestimate VWC. The Field Scout 150 captures reflectance observations using an electrical pulse (waveform) from the soil and then converts this waveform to a VWC measurement using an equation developed for mineral soils (Topp et al., 1980). The device offers three settings: standard, high-clay, and sand, with no organic mode

offered. Given the absence of known textural and percent organic data prior to sampling, all VWC measurements were taken in standard mode to ensure consistency. This choice introduced the potential for underestimating VWC in highly organic soils, as discussed further in Section 6.2. TDR was selected to maximize measurements across all 36 plots in 6 days, given that characterizing spatial soil moisture trends is the primary objective of this work. All measurements were taken from the top 20cm of the soil. The top 20cm was sampled as this depth captures the majority of root biomass, and disproportionately influences vegetation dynamics (Dong et al., 2022). To account for high heterogeneity within small spatial scales, all VWC measurements were taken in triplicate. Measurements were taken 20cm apart, starting with the plot center, then moving right, then left, horizontal to the slope. These measurements were then averaged to form the plot VWC metric (see Appendix D. SoilRx Field Protocol).

To illustrate moisture variability within each monthly campaign (day one to day six), a calibration line was established at a central location in the study area to enable repeatability. The calibration line was 15 meters long, and soil moisture was sampled at 1-meter intervals on day 1 and day 6 of each campaign. A paired t-test was then used to determine whether soil moisture changed significantly between Day 1 and Day 6, providing context on inter-campaign moisture variability. The average mean temperature in the study area across June, July, and August was 21.3°C, ranging from 5.0°C to 33.9°C, with a total summer precipitation of 54.1mm. During the June sampling window (June 23-28), a total of 4.83mm of rain fell, and the average temperature was 17.4 °C. There was 0mm during the July window, and an average temperature of 18.8°C (July 21-26). Lastly, during the August window (August 25-30), 1.27mm of rain fell, and the average sampling-week temperature was 21.3°C. All weather data was pulled from the National

Oceanic and Atmospheric Administration (NOAA) Monroe weather station (USC00455525), located approximately 20km from the study area.

Given the target resolution of this project (<10m) and the density of the forest canopy, highly accurate GPS is both essential and challenging. To combat this, the geographic location was recorded using a research-grade Global Navigation Satellite System (GNSS), the JAVAD Triumph 2. All GNSS data was post-processed using JAVAD Justin 3 software, resulting in precise point positioning (PPP) solutions with a minimum RMS of less than 1 meter. We developed all soil moisture proxy models at a 5-meter resolution, then scaled the Hydrophytic Intrinsic Potential (HIP) modeling framework to 10 meters. Thus, 1-meter GPS resolution is adequate for the purposes of this project.

4.1.4 Soil Characteristics Data Collection

Soil characteristic data were collected from a 20cm soil pit. The soil texture was estimated using a field hand texture method according to soil taxonomic USDA standards, along with Soil matrix color using a Munsell Soil Color book (Munsell Color, 2012). Soil texture was estimated in the lab using a rapid wet-sieve soil texture analysis method (Kettler et al., 2001). These methods are performed on soil particles less than 2mm, and all soils were chemically dispersed with Sodium Hexametaphosphate (HMP) prior to wet sieving. The sand fraction was separated using a 0.053-mm sieve (# 270). The silt and clay fractions are then separated using sedimentation methods; clay and silt particles are suspended in water for at least 90 minutes, and the material that settles at the bottom is the silt. Then, the clay is carefully poured out, leaving only the silt fraction. Both the sand and silt fractions are then dried back to a consistent weight at 105 °C. The clay percentage is then back-calculated using the total sample mass (Equation 1). Bulk Density samples were collected in the center of each soil pit, 2 samples were taken per pit.

Samples were taken using a PVC cylinder with a known volume, then dried to a consistent weight, and bulk density was calculated for the entire soil sample and the fine earth fraction with all material larger than 2mm removed (Equation 2), following the methods outlined in (Clancy et al., 2023).

To estimate organic matter content, Loss on Ignition (LOI) was performed. Samples were processed following United State Geologic Survey (USGS), and Natural Resources Conservation Service (NRCS) LOI methods, samples were first processed with a 2 hour run in the muffler furnace set to 500c, a subset (12 out of the 36) of highly organic samples were then processed at 16 hours at a slightly lower temperature of 400c see LOI Standard Operating Procedure (SOP) for more information (Mercury Research Laboratory., 2016; Soil Survey Staff., 2014.). A longer time is preferred for organic soil, as it allows all organic matter to burn off, providing a more accurate metric in highly organic soils.

Equation 1. Particle Size Calculations.

$$\text{Sand \%} = (\text{oven dry sand mass} \div \text{original sample mass}) \times 100$$

$$\text{Silt \%} = (\text{oven dry silt mass} \div \text{original sample mass}) \times 100$$

$$\text{Clay \%} = 100 - (\text{sand\%} + \text{silt\%})$$

Note: The above equations are from Kettler et al (2001). The original sample mass refers to the dry, sieved sample with particles less than 2mm.

Equation 2. Bulk Density.

$$BD\ total = Ms \div Vs$$

$$BD\ fine = Mf \div Vf$$

Note: Where M_s refers to the mass of the total bulk density sample, and M_f refers to the mass of the fine earth portion of the sample (all samples less than 2mm), V_s refers to the total volume of the sample, and V_f refers to the fine earth fraction (Clancy et al., 2023).

4.2 Modeling Methods

We developed and tested two models as possible proxies for soil moisture. Rather than modeling soil moisture directly, we leveraged machine learning and the underlying drivers of soil moisture distribution to model proxies for soil moisture, which we then calibrate to estimate VWC. First, a novel peer-reviewed Wetland Intrinsic Potential (WIP) model is tested as a proxy for soil moisture (Halabisky et al., 2023). We tested two versions of WIP: the baseline WIP at 5 meters and the TealWaters WIP (TW WIP) at 3 meters. Then, building off the WIP framework, we proposed a new model using iNaturalist plant observations to train a random forest model to predict Hydrophytic Intrinsic Potential (HIP). Both models tested here aimed to explain spatial distribution of soil moisture as a function of topographic relief and hydrological indices. The use of a random forest model allows for non-linearity in the relationships between dependent and independent variables. The proposed models were intended to serve as proxies for soil moisture rather than model VWC directly. The WIP model was fit to wetland occurrence data, and the HIP model was fit to wetland indicator occurrence data derived from plant observations. Although neither model predicts VWC directly, both outputs can be calibrated to estimate VWC using a linear model, as discussed further in Section 4.3 Model Validation, Residual Analysis, and VWC Estimations.

4.2.1 Covariates and Data Processing

The covariates selected in this study were guided by variables that control soil moisture distribution (McBratney et al., 2003; Murphy et al., 2009). These variables will vary across spatial scales; at finer scales, topography and hydrology will exert greater control over soil moisture distribution. At broader spatial scales, such as regional or state scales, climate, particularly precipitation and evapotranspiration, will have increasingly important controls on soil moisture. This was considered when selecting covariates for different spatial scales and extents. Initial models were developed at 5-meter resolution using topographic and hydrological covariates. Later iterations of the HIP model tested scaling the model to 10-meter resolution, leveraging statewide datasets from project partner TealWaters, a wetland-mapping company. The 10-meter model incorporated climate variables because they exert larger control over soil moisture at broader spatial scales, and the model was fit with data from multiple watersheds in Washington state, enabling a larger set of plant observations for training (Figure 2). Covariates for each model are listed below by model, HIP at five meters, and the baseline WIP share all covariates, with the 10-meter HIP and TealWaters WIP covariates slightly differing.

HIP (5m) and Baseline WIP (5m) Covariates

For the 5-meter model, topography and hydrological variables were selected. The topographic metrics used in this model were: Deviation from Mean Elevation (DEV), Planform Curvature (PLAN), Profile Curvature (PROF), Gradient (GRAD), and Elevation (see Table 1). These metrics were all calculated in ArcGIS Pro using a freely available custom toolbox built for the WIP model (Halabisky et al., 2023). Deviation from Mean Elevation is expressed as $DEV = (Z - Z_{mean}) / \sigma$, where Z is the elevation of the pixel, and Z_{mean} is the mean elevation of the neighborhood of pixels controlled by the length scale selected; σ is the standard deviation of the

pixels in the neighborhood. The length scales were user-defined, and can be run at multiple scales to capture landscape variability. DEV can be thought of as a means of quantifying areas with differing elevations to their neighborhood of cells; it can be utilized to identify both depression and mound or rise landforms, both of which will impact soil moisture dynamics. Profile Curvature (PROF), Planform Curvature (PLAN), and Gradient (GRAD) were calculated using the methods outlined by Zevenbergen and Thorne (1987). These methods treat the surface of a Digital Terrain Model (DTM) cell as a smooth polynomial surface matching the adjacent cells. The method used a circular neighborhood, allowing for GRAD, PLAN, and PROF to be calculated over multiple length scales. A more detailed explanation can be found in Halabisky et al (2023). These metrics were calculated at the following scales: 50, 100, 500, and 1000 meters. These scales were selected based on the scale of the landscape features found within the Skykomish watershed. Lastly, the parent DEM was also used to capture elevation. The DEM was resampled from a 1-meter Light Detection and Ranging (LIDAR) derived DTM (NV5, 2014-2022).

The hydrological variables selected were the Topographic Wetness Index (TWI) and the Depth to Water Index (DTW). The Topographic Wetness Index (TWI) was first proposed by Beven and Kirkby (1979). It describes wetness as a function of topography, using total catchment area and slope to identify areas where water is likely to accumulate (Equation 3). For the HIP and WIP models at 5 meters, TWI was calculated in R using Whitebox Tools (WBT), from a 5-meter raster, and filled using the WBT fill depression function and the WBT D-infinity flow accumulation algorithm. The Depth to Water (DTW) index introduced by Murphy et al. (2007) is an effective proxy for groundwater and maps "depth to water" using least-cost paths from open water (Murphy et al., 2007, 2008, 2009). Here, DTW was calculated in R following

the methods developed by Stewart et al (2023), based on the original DTW algorithm (Murphy et al., 2007, 2008, 2009). For the open water in this implementation, the National Hydrography Data (NHD) flowline and surface water were used. Low DTW values are effectively closer to open water because they have a low "cost"; as you move further from open water or higher upslope, the cost increases. This mimics hydrological conditions and can act as an effective proxy for moisture. A high-level overview of model covariates for both the five-meter HIP and the baseline WIP is provided in Table 1.

Equation 3. Topographic Wetness Index (TWI).

$$TWI = \ln\left(\frac{A}{\tan(\theta)}\right)$$

Note: TWI can serve as a relative proxy for moisture, and is calculated using specific contributing areas expressed as A , and the ground- surface gradient expressed as θ .

Table 1. HIP (5m) and Baseline WIP (5) Covariate Table

Covariate	Length scale(m)	Processing Methods	Data source	References
Deviation from Mean Elevation (DEV)	50,100,500,1,000	WIP ArcGIS Pro custom toolbox $DEV = (Z - Z_{mean})/\sigma$	LiDAR-derived DTM	Halabisky et al., 2023; Newman et al., 2018
Planform Curvature (PLAN)	50, 100, 500, 1,000	WIP ArcGIS Pro custom toolbox, using Fortran code from Miller et al (2003)	LiDAR-derived DTM	Halabisky et al., 2023
Profile Curvature (PROF)	50, 100, 500, 1,000	WIP ArcGIS Pro custom toolbox, using Fortran code from Miller et al (2003)	LiDAR-derived DTM	Halabisky et al., 2023

Gradient (GRAD)	50, 100, 500, 1,000	WIP ArcGIS Pro custom toolbox, using Fortran code from Miller et al (2003)	LiDAR-derived DTM	Halabisky et al., 2023
Elevation	5	Resampled from a 1m Lidar-derived DTM using Bilinear interpolation	LiDAR-derived DTM	Halabisky et al., 2023
Topographic Wetness Index (TWI)	5	Whitebox Tools (WBT), DTM was filled (fill depression function), D8	LiDAR-derived DTM	Bevin and Kirkby, 1979
Depth to Water Index (DTW)	5	Executed in R using the terra package	LiDAR-derived DTM, National Hydrography Data (NHD) flowlines, and surface water	Murphy et al., 2007, 2008, 2009; Stewart et al 2023; Hijmans et al., 2026

HIP (10m) Covariates

The topographic variables employed in the 10-meter HIP were GRAD, PLAN, PROF, and Elevation as described above. The topographic position index (TPI) was used in place of DEV. TPI is calculated as $E - E_{mean}$, where E is the elevation of a given pixel, and E_{mean} is the mean elevation within a designated neighborhood of pixels; this is the same as DEV, except that the standard deviation is not taken. For the Hydrological variables, TWI and Height Above Nearest Drainage (HAND) were selected. The TWI layer was generated using a 10-meter DEM with all NHD+ flow lines burned in with a radius of 10 meters and a depth of 25 meters. Depressions were then drained by locating pour points within the DEM, and carving the flow path from each pour point to the lowest point within the depression with an elevation equal to or less than the lowest point (Soille, 2004; Soille et al., 2003). The described raster processing methods create a hydrologically conditioned DEM maintaining as much topographic detail as

possible, by creating defined surface connections without filling depressions or smoothing the landscape more than necessary. Total catchment areas were defined using a D-infinity in non-channel areas and a D8 in all burned in channels. The gradient was calculated within a 25-meter radius, effectively smoothing the DEM's slope.

Height Above the Nearest Drainage (HAND), first introduced by Nobre et al (2011), was calculated using the same 10-meter DEM with no generalizations or smoothing applied (Nobre et al., 2011). The D8 with Local Transversal Deviation (D8-LTD) algorithm was used to define drainage wings, and all elevations or heights were then calculated based on perennial streams defined using a lidar-derived DTM (NV5., 2014-2022). The D8-LTD is preferred for this method because it accounts for bias introduced by using only 8 flow directions (Orlandini et al., 2003). All topographic metrics were created using the USGS 1/3 arcsecond 10-m DEM, and hydrological data were sourced from a high-resolution lidar-derived DTM.

Climatic variables were all sourced from PRISM data; all variables are 30-year averages, at a resolution of 30 meters (Daly et al., 2008). The 30-meter rasters were then down sampled to 10 meters to meet the resolution of all other covariates. The climate variables selected were precipitation, solar radiation, dew point, temperature, and vapor pressure deficit. More details on each covariate can be found below in Table 2.

Table 2. HIP 10m Covariate Table.

Covariate	Length scale(m)	Processing Methods	Data source	References
Topographic Position Index (TPI)	50, 150, 300	$TPI = (Z - Z_{mean})$	USGS 1/3 arcsecond 10-m DEM	Halabisky et al., 2023

Planform Curvature (PLAN)	50, 150, 300	Following the methods developed by Halabisky et al. (2023), and executed in Python	USGS 1/3 arcsecond 10-m DEM	Halabisky et al., 2023
Profile Curvature (PROF)	50, 150, 300	Following the methods developed by Halabisky et al. (2023), and executed in Python	USGS 1/3 arcsecond 10-m DEM	Halabisky et al., 2023
Gradient (GRAD)	50, 150, 300	Following the methods developed by Halabisky et al. (2023), and executed in Python	USGS 1/3 arcsecond 10-m DEM	Halabisky et al., 2023
Elevation	10	no additional processing	USGS 1/3 arcsecond 10-m DEM	Halabisky et al., 2023
Topographic Wetness Index (TWI)	10	flow lines were burned in, and TWI was calculated on a hydrologically conditioned DEM, using a D8 in changes and D-infinity elsewhere.	USGS 1/3 arcsecond 10-m DEM, flowlines derived from the lidar DEM	Beven & Kirkby, 1979; Soille, 2004; Soille et al., 2003
Height Above Nearest Drainage (HAND)	10	The D8-LTD algorithm was used to define drainage wings, and HAND was calculated using channel evaluation using code from Miller et al (2003)	USGS 1/3 arcsecond 10-m DEM, flowlines derived from the lidar DEM	Nobre et al (2011)
Precipitation	10	Prism, 30-year average	Prism, 30-year average, resampled from 30m to 10m	PRISM Climate Group, Oregon State University, 2008

Solar Radiation	10	Solar radiation under clear-sky conditions, Solar radiation adjusted for slope and aspect, and Total solar radiation	Prism, 30-year average, resampled from 30m to 10m	PRISM Climate Group, Oregon State University, 2008
Dew point	10	Mean dew point temperature	Prism, 30-year average, resampled from 30m to 10m	PRISM Climate Group, Oregon State University, 2008
Temperature	10	Temperature min and max	Prism, 30-year average, resampled from 30m to 10m	PRISM Climate Group, Oregon State University, 2008
Vapor pressure deficit (VPD)	10	VPD Min and Max	Prism, 30-year average, resampled from 30m to 10m	PRISM Climate Group, Oregon State University, 2008

TW WIP (3m) Covariates and Data

Covariates for the TW WIP were hydrological and topographic, including TWI, HAND, GRAD, PROF, PLAN, and TPI. We calculated TPI at 50m and derived PLAN, PROF, and GRAD at 50m, 300m, and 1,000m (Table 3). TWI and HAND were calculated following the methods described in the HIP (10m) Covariates Section. All TW WIP covariates were calculated at 3-meter resolution from a high-resolution Lidar-derived DTM.

Table 3. TW WIP 3m Covariate table.

Covariate	Length scale(m)	Processing Methods	Data source	References

Topographic Position Index (TPI)	50, 150, 300	$TPI = (Z - Z_{mean})$	USGS 1/3 arcsecond 10-m DEM	Halabisky et al., 2023
Planform Curvature (PLAN)	50, 300, 1,000	Following the methods developed by Halabisky et al. (2023), and executed in Python	Lidar-derived DTM (3m)	Halabisky et al., 2023
Profile Curvature (PROF)	50, 300, 1,000	Following the methods developed by Halabisky et al. (2023), and executed in Python	Lidar-derived DTM (3m)	Halabisky et al., 2023
Gradient (GRAD)	50, 300, 1,000	Following the methods developed by Halabisky et al. (2023), and executed in Python	Lidar-derived DTM (3m)	Halabisky et al., 2023
Topographic Wetness Index (TWI)	3	flow lines were burned in, and TWI was calculated on a hydrologically conditioned DEM, using a D8 in changes and D-infinity elsewhere.	Lidar-derived DTM (3m), and flowlines derived from the lidar DEM	Beven & Kirkby, 1979; Soille, 2004; Soille et al., 2003
HAND	3	The D8-LTD algorithm was used to define drainage wings, and HAND was calculated using channel evaluation using code from Miller et al (2003)	Lidar-derived DTM (3m), and flowlines derived from a lidar DTM	Nobre et al (2011)

4.2.2 WIP Model

Baseline WIP (5)

Random forest modeling is becoming increasingly popular in ecological modeling, including in wetness mapping (Ågren et al., 2021; Halabisky et al., 2023; Kemppinen et al., 2018; Lidberg et al., 2020). Random Forest is used here because of its ability to learn complex nonlinear relationships between predictors. The WIP model is a topographic and hydrology-driven version of the original model presented in Halabisky et al. (2023). This model is trained on a binary set of wet and dry points; the National Wetland Inventory (NWI) polygons are used to create the wet class of training points, and the dry are created using a spatially randomized point generator in the areas outside of the NWI polygons, and areas classified as open water. The Random Forest model is then run in R using the randomForest library (Liaw & Wiener, 2022). The model is set with 200 trees, and the output was generated using the probability function (Prob), so the model predicts the probability of the pixel to be “Wet” (Halabisky et al., 2023). The covariates for this model are Elevation, DEV, PLAN, PROF, GRAD at (50, 100, 500, and 1000 meters), TWI, and DTW (Table 1). The training dataset consisted of 2,406 dry points (generated in areas outside NWI polygons) and 1,511 wet points (generated in NWI polygons). This WIP model will be referred to as the Baseline WIP.

TealWaters WIP (3m)

A second WIP model we tested was generated by TealWaters at a 3-meter resolution and will be referred to as the TW WIP. Similar to the Baseline WIP, described above, this model utilized hydrological and topographic variables. The Covariates for this model were TWI, HAND, GRAD, PROF, PLAN, and TPI. The TPI was calculated at 50m, and PLAN, PROF, and GRAD were derived at 50m, 300m, and 1,000m. The model was fit using the scikit-learn

RandomForestClassifier in Python, with 65 trees and 1 to 5 wetland-upland split (4,015 dry points and 700 wet points) (Pedregosa et al., 2011). The output was generated by predicting the probability of a wet pixel, resulting in the WIP probability raster with values ranging from 0 to 1. This model will be referred to as TealWaters (TW) WIP.

4.2.3 HIP Model

The HIP model used the same model architecture and covariates as WIP, differing only in training data. We used the iNaturalist Research-grade Observation dataset hosted by the Global Biodiversity Information Facility (GBIF) (GBIF.org, 2026). We pulled all Plantae observations in Washington state from 2000 to 2026; these dates were selected to maximize observation density while remaining within the period of consistent iNaturalist adoption. Observations were filtered to a GPS accuracy of 5 meters or less, since 5 meters was the target resolution for modeling.

The plant observations alone do not inform the model of moisture content. Given this, we coupled the plant observations with the most recent version (2022) of the U.S. Army Corps of Engineers National Wetland Plant List (version 3.6). This list includes wetland indicator status, a common botanical designation used in wetland delineation (Lichvar & Gillrich, 2011). There are 5 wetland indicator statuses: Obligate (OBL), Facultative Wet (FACW), Facultative (FAC), Facultative Upland (FACU), and Upland (UPL). These delineations are associated with a plant's occurrence in wetlands (Lichvar & Gillrich, 2011). We coupled wetland indicators with the plant occurrence data from iNaturalist, resulting in a spatially explicit dataset of wetland indicator status from 2000 to 2026 (GBIF.org, 2026).

HIP 5m

The wetland indicator status spatial dataset was used to create a binary and three-class training dataset by binning wetland indicators. For the binary model, UPL and FACU were binned together to form the dry class, and FACW and OBL were binned together to create the wet class. The three-class models pull apart the wet bin by adding FACW as a separate transitional bin (Table 4) FAC species were excluded from the training dataset due to their generalist nature; they are found across the entire moisture gradient, from wetlands into uplands. So, including FAC species would reduce the model's ability to learn clear relationships between hydrophytic indicator status (wet or dry) and covariates, because these species span multiple hydrological regimes and cannot be reliably assigned to either the wet or dry training class.

The plant observations were clipped to the extent of the Skykomish watershed to match all covariates for the 5-meter HIP model. The same random forest outlined for the WIP model was fit with wetland indicator occurrence training data: one model with all data from 2000 to 2026, one with 2025 observations only, and one three-class model using all years (Table 4). A probability raster was generated to predict the probability of the Wet class across all pixels in the Skykomish watershed.

HIP 10m

The 5-meter HIP model in the Skykomish was fit with a total of 1,274 plant observations, with only 274 in the wet bin. To increase the training dataset size, we tested this HIP model framework at a coarser spatial scale of 10 meters, using data spanning all of Washington state and available to us via a partnership with TealWaters. The availability of these 10-meter covariates allowed the model to be fit on plant observations from a much larger extent. The 10-meter HIP model covariates differed from the 5-meter model; this difference was driven by the

data available at this larger spatial scale. The 10-meter HIP consisted of topographic, hydrological, and climatic variables. Climate variables are an important addition to this model given that it spans a much larger extent, where climate exerts greater control over soil moisture distribution. All topographic metrics processing and data acquisition are described in detail in the methods (Section 4.2.1, page 23).

Plant observations for the 10-meter HIP were filtered to only include observations within the Western Mountains, Valleys, and Coast (WMVC) wetland region of Washington state. The WMVC is an Army Corps of Engineers delineated wetland region covering most of Washington state; it excludes areas in eastern Washington that fall in the Arid West region (Figure 2). This filtering step was important because NWPL indicators differ across wetland regions, and the ecology and mechanisms that impact soil moisture differ between them. After extent and GPS accuracy filtering, 158,779 observations remained; however, 50,754 observations did not have a wetland indicator status, leaving 108,025 observations of plants with wetland indicator status. This dataset was divided into binary and three-class models, similarly to the 5-meter HIP. With the larger dataset, we also tested adding *Sphagnum* species to the wet training class given their prevalence in peatlands. See Table 4 for all 10-meter model training dataset configurations. The 10-meter model was fit in R using the same random forest methods as the 5-meter model, with 500 trees. The probability raster was generated using the probability function, predicting the likelihood of the wet class pixel-wise across the watershed.

Table 4. HIP Models: Training Datasets.

Training dataset	Splits	Training points	Covariates
HIP 10m Binary 2025	Wet = OBL and FACW,	Dry = 16,291, Wet = 3,037	TPI, PROF, PLAN, GRAD, HAND, TWI, Precipitation,

	Dry = UPL and FACU		Solar radiation, Dew point, VPD
HIP 10m Binary All Years	Wet = OBL and FACW Dry = UPL and FACU	Dry = 66,322, Wet = 13,717	TPI, PROF, PLAN, GRAD, HAND, TWI, Precipitation, Solar radiation, Dew point, VPD
HIP 10m Three Class	Wet = OBL Transitional = FACW Dry = UPL and FACU	Dry = 65,688, Transitional = 8,506, Wet = 5,061	TPI, PROF, PLAN, GRAD, HAND, TWI, Precipitation, Solar radiation, Dew point, VPD
HIP 10m <i>Sphagnum</i> 2025	Wet = OBL, FACW, and <i>Sphagnum</i> species Dry = UPL and FACU	Dry = 16,291, Wet = 3,046	TPI, PROF, PLAN, GRAD, HAND, TWI, Precipitation, Solar radiation, dew point, VPD
HIP 10m <i>Sphagnum</i> All Years	Wet = OBL, FACW, and <i>Sphagnum</i> species Dry = UPL and FACU	Dry = 66322, Wet = 13748	TPI, PROF, PLAN, GRAD, HAND, TWI, Precipitation, Solar radiation, Dew point, VPD
HIP 5m Binary 2025	Wet = OBL and FACW Dry = UPL and FACU	Dry = 291, Wet = 108	DEV, PROF, PLAN, GRAD, TWI, DTW, ELV
HIP 5m Binary All Years	Wet = OBL and FACW Dry = UPL and FACU	Dry = 1023, Wet = 274	DEV, PROF, PLAN, GRAD, TWI, DTW, ELV
HIP 5m Three Class	Wet = OBL Transitional = FACW Dry = UPL and FACU	Dry = 1,023, Transitional = 172, Wet = 102	DEV, PROF, PLAN, GRAD, TWI, DTW, ELV

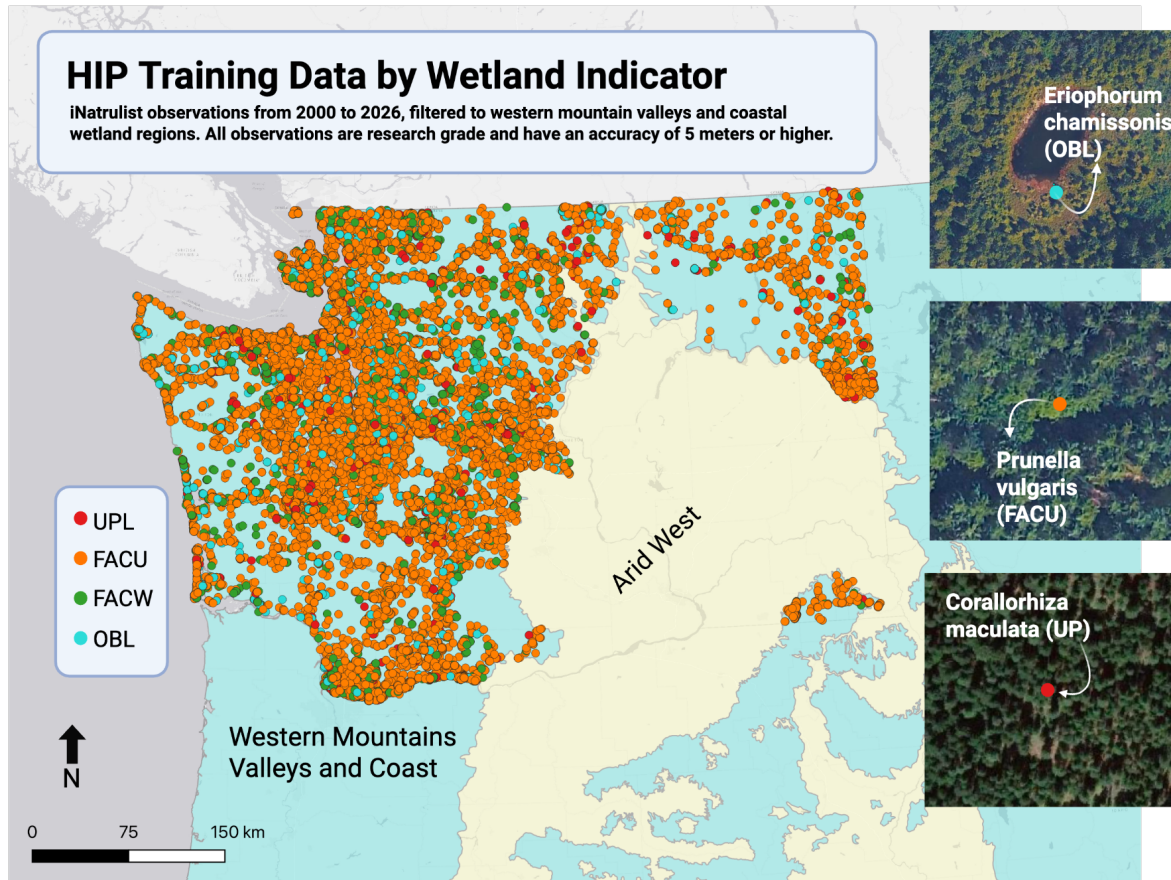


Figure 2. Map of HIP Training Data by Wetland Indicator and Wetland region for Washington State.

4.3 Model Validation, Residual Analysis, and VWC Estimations

We quantified how well the HIP and WIP models perform as a proxy for soil moisture by testing whether the HIP and WIP have any correlation with field soil moisture measurements using a linear model (LM) (Figure 3). This approach was taken because the LM accounts for scale mismatches, allowing us to quantify the relationship between the model output and the VWC without rescaling or normalizing either variable. Moreover, the LM allows us to establish an empirical relationship to calibrate models to estimate VWC. LMs are also widely adopted in the literature for their relative simplicity and effectiveness (Halter et al., 2025; Winzeler et al., 2022). The empirical relationship between HIP and WIP outputs and observed summer VWC

was established using a linear model as described above. Estimated VWC was derived using the intercepts and slopes reported from the linear models using the following formula:

$$\text{Estimated VWC} = \mathbf{B} + \mathbf{M} \times (\text{model probability})$$

where B is the intercept, and M is the slope as reported in Table 5.

To compare across linear models Akaike's Information Criterion adjusted for small sample sizes (AICc) was calculated. While the LM provides useful metrics for correlation and fit, AICc helps distinguish between models with similar correlation metrics, such as R-squared. AICc is a statistical method that can be used to compare models with the same dependent variable. Models with lower AICc scores are considered better (Akaike, 1998; Sugiyura, 1978). AICc was calculated using the AICc function from the MuMIn library in R (Bartoń K, 2025). Alongside comparing AICc scores, ΔAICc was used for interpretability; ΔAICc is calculated by setting the lowest AICc score to 0 and expressing all other models as the difference from that value. Models with a low ΔAICc are considered better, and a score of 0 is considered the best fit. Although models with an ΔAICc of less than 2 are not statistically distinguished from the best model (Burnham & Anderson, 2004). The utilization of AICc coupled with LM statistics allows for us to compare among models and judge the fit of the HIP and WIP as a soil moisture proxy.

In addition to analyzing statistics from the LM described above, the model residuals were also investigated. Residuals are calculated by subtracting observed values from predicted values, and are utilized here to help deepen our understanding of where models can be used as direct proxies for soil moisture. Residuals were chosen here, given their relative simplicity and ability to test the models' direct relationship with soil moisture beyond just correlation. Given that residuals are scale independent, they express error in the original units in which they are

calculated, making them more easily interpreted and allowing them to remain on a scale relevant to saturation. This is an important step in model interpretation, as the analysis of residuals allows us to identify specific areas or attributes where the model output diverges from on-the-ground moisture measurements. Moreover, residuals allow for the identification of the distribution of error, helping to identify systematic over- or underestimation.

We analyzed residuals to identify how soil properties (LOI, Bulk Density, and percent of sand, silt, and clay) affect soil moisture beyond the model capabilities. The residuals were analyzed based on VWC and continuous soil properties, including percent sand, percent silt, percent clay, bulk density, and LOI. Additionally, residuals were grouped into classes designed to mimic soil drainage classes, with bins delineated using VWC percent. For the drainage class approach, two binning methods were used. The first segmented the VWC data into six drainage classes within the study area. Given the small sample size, a less aggregated approach was also used, with three classes: well drained (0-10%), moderately well drained to somewhat poorly drained (10-30%), and poorly drained to very poorly drained (30% and above). Prior to calculating residuals, both model outputs and VWC data were normalized to a 0-1 scale using a min-max approach, with the maximum model output set to 1 and the minimum set to 0, and the same applied to all VWC data. The normalized VWC measurements can be thought of as a percentage of maximum saturation rather than volume of water. Residuals were then calculated using both the normalized model values and the normalized VWC values (normalized VWC minus normalized model output). Normalizing both model and VWC values to the same scale enables a more accurate assessment of residuals, as WIP operates on a 0-1 (or 0-100) probability scale while VWC rarely exceeds approximately 60%; normalization allows both metrics to be compared on the same scale.

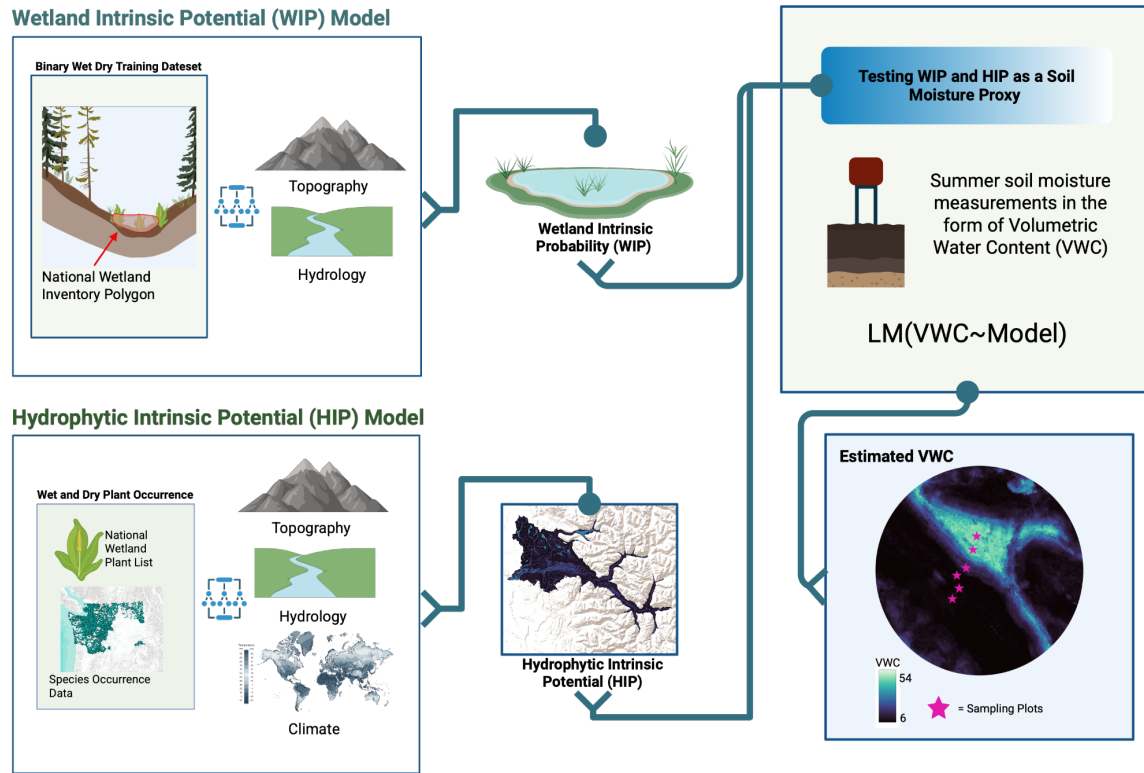


Figure 3. HIP and WIP Model Flowchart for Soil Moisture Monitoring

5. Results

5.1 Field Data Collection Results

A calibration transect was established to explore variation in VWC between day 1 and day 6 of each campaign. Calibration transect data were analyzed using a paired t-test. The results indicate that there is a statistically significant difference between the first and last days of sampling in August only ($p = .005$), with a mean decrease of 1.55% VWC. July showed a marginally significant trend ($p = .052$, mean decrease = 0.77% VWC), while June showed no significant change ($p = .335$, mean difference = 1.14% VWC). Across all months, the largest mean change in VWC was observed in August at 1.55%, although, for the purposes of this

sampling campaign, this is relatively marginal given the moisture meter's reported accuracy of $\pm 3\%$. So, the 1.55% change could be attributed to instrument error rather than meaningful shifts in VWC across the sampling window. Moreover, these methods aim to gain a better understanding of the overall spatial distribution of surface soil moisture, so a 1.55% difference will not negatively impact the quality of the VWC data for these purposes. Generally, soil moisture decreased in all months, indicating that soils consistently dried throughout June, July, and August at the calibration Line. This is consistent with observations from other field sampling locations. The difference between day 1 and day 6 during August is a limitation of this dataset, and future work may consider sampling a calibration line before and after each plot to enable post-sampling correction. Across all 36 plots Volumetric Water Content (VWC) data was aggregated to a mean summer soil moisture measurement. Most of the mean summer soil moisture across our study plots were under 25% VWC, with minimal observations in the 25% to 40% VWC range (Figure 4).

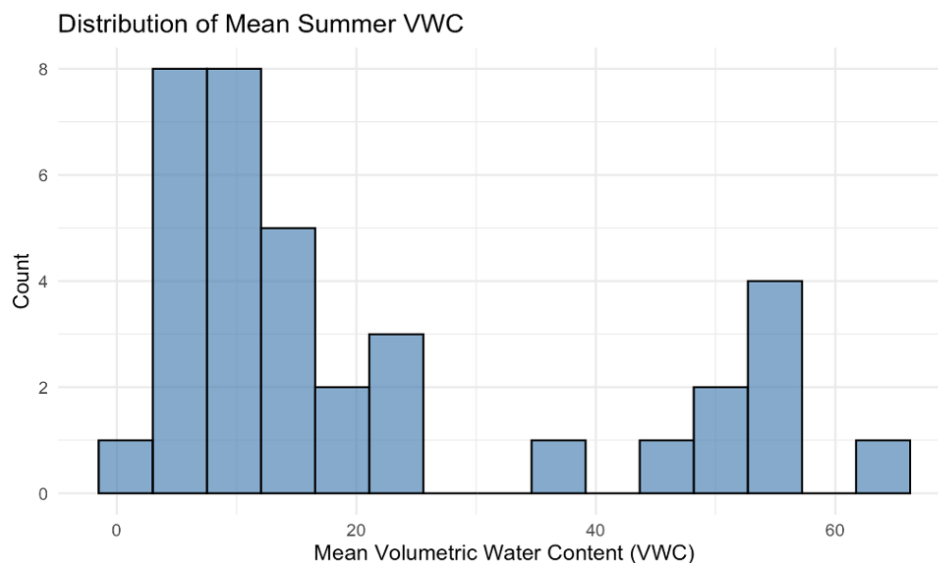


Figure 4: Mean Summer 2025 Soil Moisture Across 36 Sites in the Skykomish Watershed

Laboratory analysis of soil texture following Kettler et al (2001) revealed that most soils within the study area have a high content of sand, with clusters of sites categorized as sandy clay loams and sandy loams, with several samples in the sandy clay and two samples in the clay region (Figure 5). These results are consistent with expectations for forest soils, although the 2 clay soils seen at Marsh Creek Organic (MCO) could be due to overestimation of clay by the Kettler method. The minimum fine bulk density (4.1.4 Soil Characteristics Data Collection) was 0.0138 g/cm³, and the max fine bulk density was 0.725 g/cm³; these low bulk density measurements are consistent with the highly organic soils found at the study sites (Figure 6).

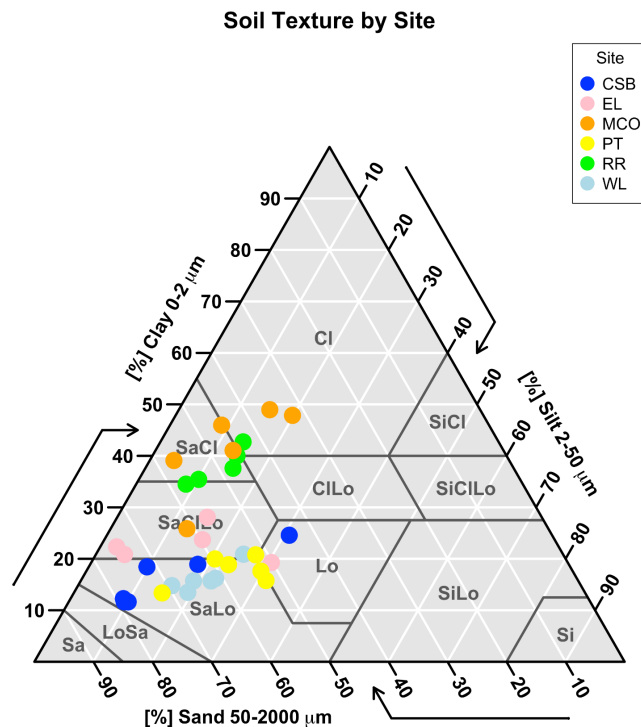


Figure 5. Soil Texture by Site in the Lower Skykomish Watershed

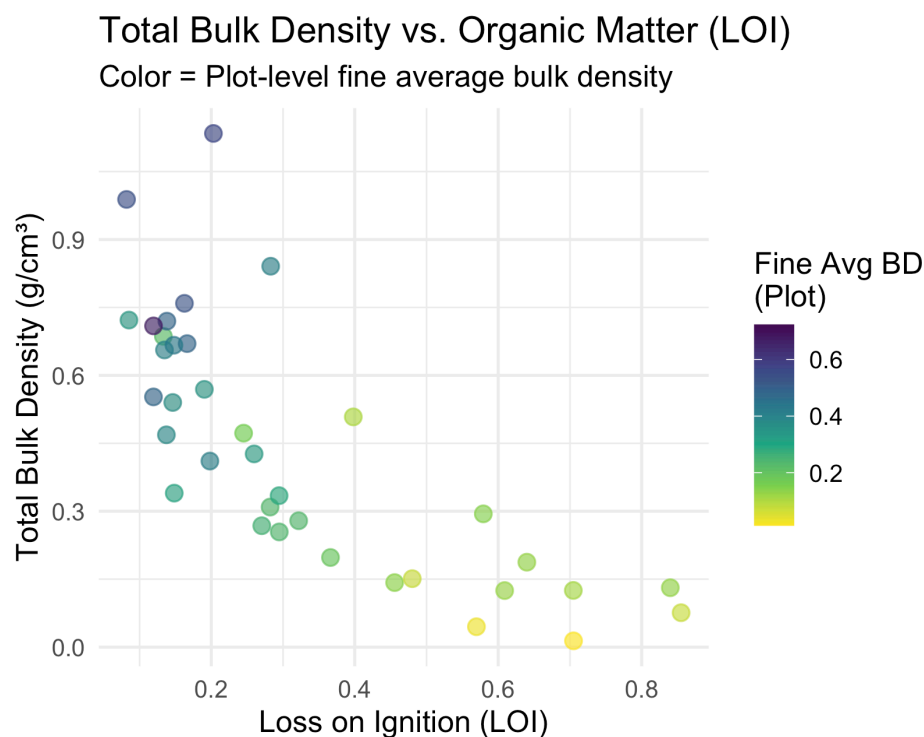


Figure 6. Site Bulk Density (BD) and LOI Relationship with Fine BD and Total BD

5.2 WIP Results

TW WIP 3m

The overall accuracy for the 3m TW WIP model was 93% in a binary cross-validation, with per-class precision of 95% in non-wetlands and 78% in wetlands. The top three highest-ranking variables for feature importance were GRAD at 50m, HAND, and GRAD at 300 meters. Comparing the TW WIP output with VWC data from 2025 showed a significant relationship between WIP and overall moisture, with an R^2 of .60 (Table 5).

WIP 5m

The 5m WIP had an Out of Bag (OOB) overall error rate of 4.4%. Per-class OOB error rates were 5.03% in the dry class and 3.31% in the wet class, reflecting higher model uncertainty in the drier areas. The top three highest-ranking variables for mean decreased Gini were GRAD

at 50 meters, DTW, and GRAD at 100 meters. TWI ranked notably low, contributing little to the model's predictive power. The linear model fit with summer VWC indicated a statistically significant correlation between soil moisture and model output, with an R^2 of .46 (Table 5).

5.3 HIP Results

Model outputs and soil moisture correlation were affected by the model training data, as models were trained with varying binning methods and temporal extents (Table 4). All models performed poorly in high-elevation areas due to a lack of training data. Given this, high-elevation areas were masked out, and correlation statistics should not be extrapolated across elevations greater than 500 meters within the watershed. Over 70% of all training data falls under 500 meters, with only 10% in the 500-1,000 range, 7% in the 1,000-1,500 range, and 1% in the 1,500-2,000 range. Given the distribution of training data, all areas above 500 meters were masked out, and no soil moisture data was taken above 500 meters.

HIP 5m

The binary models, trained with 2000–2026 and 2025-only occurrence data, showed notably different performance. The 2000–2026 model had an OOB error rate of 17.58%, and a Kappa of 0.38, with per-class OOB error rates of 1% in the dry and 62% in the wet. The top three mean decreased Gini features were DEV 500, DEV 100, and TWI, with DTW and elevation also ranking notably high at 5th and 6th, respectively. The binary model fit with data from all years had a notably lower correlation with summer soil moisture ($R^2 = .34$, see Table 5)

The 2025-only model had a slightly higher OOB error rate of 21.55%, an accuracy of 0.7845, and a Kappa of 0.40, with per-class OOB error rates of 9% in the dry and 55% in the wet. The top mean decreased Gini variables shifted to DEV 500, TWI, and DEV 1,000, with

planform and profile curvature ranking lowest. The 2025-only model had the strongest correlation with summer soil moisture of the two binary models and the lowest AICc among all 5-meter models ($R^2 = .56$, see Table 5).

The three-class model was fit with 1,023 dry points, 172 transitional points, and 102 wet points from the 2000–2026 dataset. The overall OOB error rate was 21.51%, and the Kappa was 0.21. Per-class OOB error rates were 4% in the dry, 90% in the transitional zone, and 81% in the wet, consistent with expectations given the difficulty of classifying transitional areas. The top three mean-decreased Gini features for the wet class were TWI, DEV 500, and DEV 100, as in the binary models, DTW ranked notably high at 5th place. The probability raster predicted for the wet class had an R^2 of .49 when tested against summer soil moisture (Table 5).

HIP 10m

Out of all five models tested at 10 meters, the binary models fit with 2000–2026 and 2025-only occurrence data performed similarly across all metrics. The 2000–2026 model had an OOB error rate of 13.26%, an accuracy of 0.87, and a Kappa of 0.43, with per-class OOB error rates of 3% in the dry and 62% in the wet. The 2025-only model had a nearly identical OOB error rate of 13.25%, and Kappa of 0.36. Across both binary models, the top-ranking mean decreased Gini variables were HAND, elevation, and GRAD at 50 meters, while climatic variables ranked lowest across all models, with vapor pressure deficit ranked last, supporting local topography as the primary driver of wetness distributions at this scale. The inclusion of *Sphagnum* observations in both binary models produced negligible changes across all metrics and is not discussed further. The 2000–2026 and 2025-only models had R^2 values of .58 and .55

and AICc scores of 288.22 and 290.72, respectively, when tested against summer soil moisture (Table 5).

The three-class model was fit with 65,688 dry, 8,506 transitional, and 5,061 wet points from the 2000–2026 occurrence dataset. The overall OOB error rate was 14.95%, slightly higher than the binary models, with a Kappa of 0.34. Per-class OOB error rates were 3% in the dry, 84% in the transitional zone, and 62% in the wet, consistent with expectations. The highest-ranking mean-decreased Gini features were HAND, elevation, and GRAD at 50 meters, with TPI and TWI ranked 4th and 5th; climatic variables again ranked lowest with vapor pressure deficit at the bottom. The three-class model had the highest correlation with soil moisture across all models, with an R^2 of .63 and the lowest AICc score of 283.60 (Table 5).

Table 5. LM Result Table for Soil Moisture from HIP and WIP (rescaled to 0-100).

Model	Interce pt (B)	Slope (M)	R^2	Adj. R^2	Residual Std. Error	F statistic	p value	AICc
TW-WIP	5.7721	0.4848	.60	.58	12.16	50.01	<.001	286.75
Baseline WIP	8.6393	0.3809	.47	.45	13.99	29.53	<.001	296.81
HIP 5m 2025	1.8765	0.7292	.56	.54	12.75	42.50	<.001	290.12
HIP 5m 26- Year	5.7686	0.516	.36	.34	15.28	19.26	<.001	303.16
HIP 5m 3- Class	6.6717	1.3678	.49	.47	13.72	32.04	<.001	295.41
HIP 10m 2000-2026	-0.2727	0.9062	.58	.57	12.42	46.65	<.001	288.22
HIP 10m <i>Sphagnum</i>	1.1895	0.841	.57	.56	12.55	44.87	<.001	289.02
HIP10m 2025	-1.0595	1.0218	.55	.54	12.85	41.23	<.001	290.72

HIP 10m <i>Sphagnum</i> 2025	-1.0595	1.0218	.55	.54	12.85	41.23	<.001	290.72
HIP 10 3- Class	6.6951	1.2847	.63	.62	11.64	57.69	<.001	283.60

Hydrophytic Intrinsic Potential (HIP)

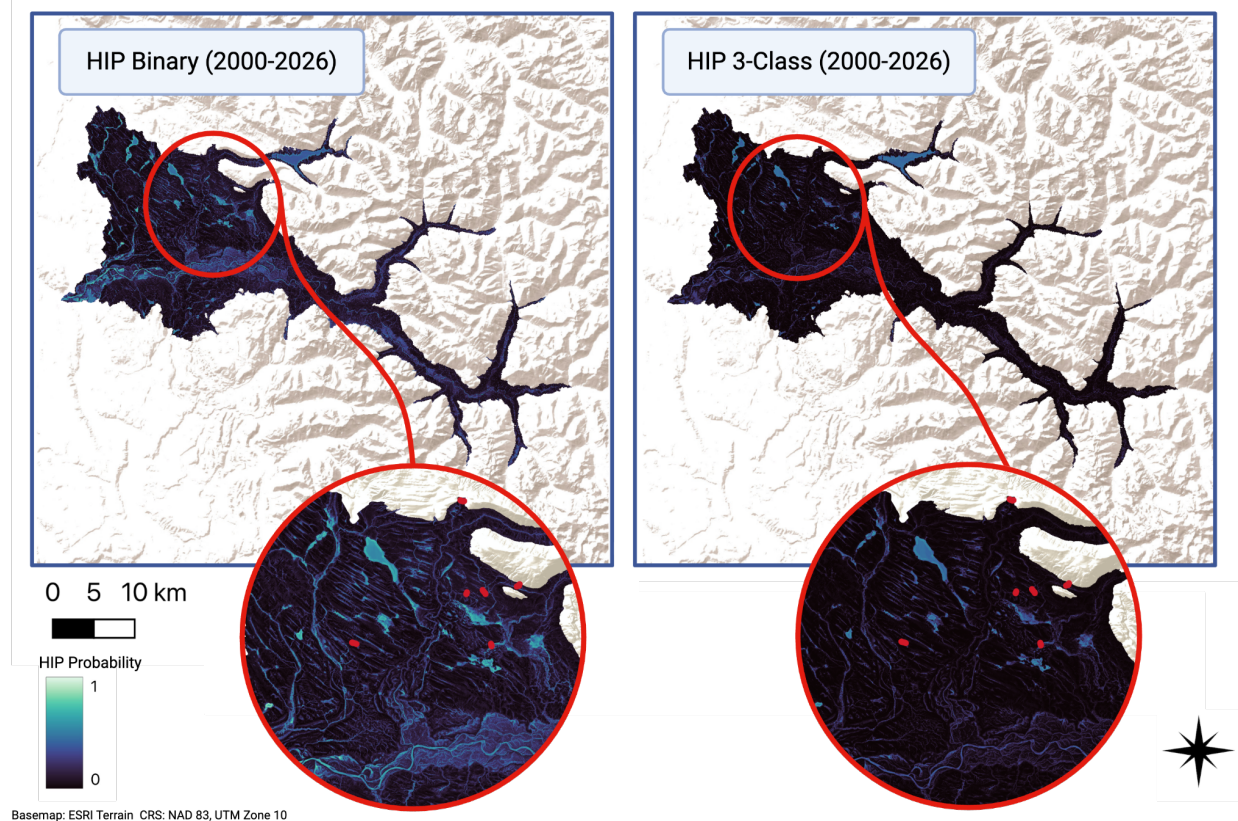


Figure 7. HIP (10m) in the Lower 500m of the Watershed, Zoomed in on the Site Marsh Creek Organic

5.4 Residual Analysis (RA)

WIP RA results

Residual analysis revealed minimal correlation (Pearson's r) between soil properties and normalized residuals across both the TW WIP and the Baseline WIP. In the TW WIP, there was no statistically significant positive correlation between normalized model residual and VWC ($r = .21, p = .231$). There was also no statistically significant correlation between TW residuals and

LOI. The baseline WIP normalized residuals were also not correlated with VWC ($r = .25$, $p = .142$), and LOI showed no statistically significant negative correlation ($r = -.14$, $p = .422$). The only continuous soil physical property with any correlation was clay, which showed a weak positive correlation with baseline WIP model residuals ($r = .42$, $p = .025$). Both models have relatively similar performance in terms of residual variance across moisture zones, although the Baseline 5m WIP has higher residual variance in the moderately well-drained/somewhat well-drained soil (10-30% VWC). Percent organic matter inferred via LOI does not have a large control over the model's ability to act as a proxy for soil moisture. The weak positive correlation with clay percentage in the baseline WIP indicates that the model's outputs were lower in high-clay soils.

HIP RA results

HIP 5m

All 5m HIP models were correlated with VWC, with no correlation to other soil physical properties. The binary HIP trained on data from 2000 to 2025 had the lowest correlation with VWC compared to the 2025 and 3-class models; the r was .38, and the p value was .022. The 2025 HIP residuals had the highest correlation with VWC ($r = .48$, $p = .003$). The three-class model also had a statistically significant correlation with VWC ($r = .46$, $p = .004$). The positive correlation between model residual and VWC indicates that all the 5-meter HIP models were underestimating moisture towards the wetter end of the moisture spectrum.

HIP 10m

Across all five 10-meter HIP models, residuals revealed no statistically significant correlation with any of the soil physical properties. All models were positively correlated with

VWC, indicating the model's predictive power as a proxy for soil moisture decayed in the wetter part of the moisture gradient. The three-class HIP fit with data from all years had the highest residual correlation with VWC ($r = .66, p = <.001$). The lowest correlation between model residuals and VWC was seen in the HIP fit with data from 2000-2026, with ($r = .57, p = <.001$). The remaining models had very similar correlations ($r = .6$ to $.61, p = <.001$). The increase in model residuals in the wetter zone is shown in Figure 8. We can also see the wider range in model residuals in the transition zones, with the best fit in the driest zone (0-10% VWC).

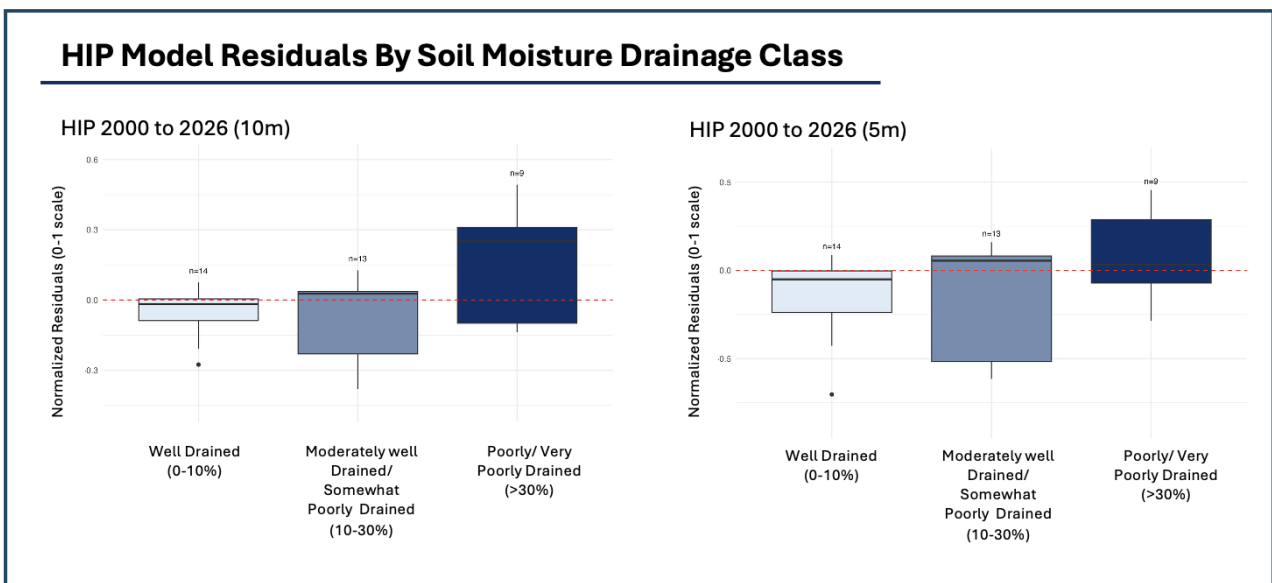


Figure 8. Normalized Residuals from the HIP

5.5 VWC Estimates from WIP and HIP

The VWC calibration framework was applied to the three best-performing models: the HIP 10m binary (2000–2026), the HIP 10m three-class, and the TW WIP. The residual standard error, which expresses the average deviation of predicted VWC values from observed VWC, was used to characterize the uncertainty of each VWC estimate. For the TW WIP, the residual standard error was 12.16% VWC, indicating that VWC estimates derived from this model should

be interpreted with an uncertainty of approximately $\pm 12\%$ VWC, with a maximum VWC estimate of 54 and a minimum of 6, as seen in Figure 9. The resulting spatially explicit VWC layers should be interpreted as estimates of relative moisture distribution across the landscape rather than precise point measurements of VWC.

When the linear calibration was applied across the HIP raster layers, predicted VWC values were extrapolated beyond the VWC range observed in the summer soil moisture dataset. The 10-meter HIP fit to data from 2000 to 2026 yielded a minimum of -0.27 and a maximum of 87.45% VWC. The three-class HIP at 10 meters estimated a minimum of 6.70 and a maximum of 122.58% VWC. Both 87% and 122% VWC are extraordinarily high and unrealistic values, indicating that the model is extrapolating outside the expected boundaries for soil VWC. However, values above 60%, these values are clustered in areas of surface water and represent less than 2% of all pixels.

This suggests that while the HIP models demonstrated the strongest statistical fit to observed soil moisture (highest *R*-squared, lowest AICc), the raster layers contained probability values that exceeded those in the calibration dataset, leading the models to extrapolate beyond the values they were fit with. As a result, VWC estimates derived from both HIP models should be interpreted with caution, and values above 60% likely represent areas of open water or full saturation rather than soil moisture. The TW WIP model produced predicted VWC values that remained within the expected range of observed soil moisture, suggesting it is a more conservative and reliable predictor for spatial estimation of VWC across the landscape, despite its lower statistical fit in comparison to HIP.

TealWaters WIP Estimates of VWC

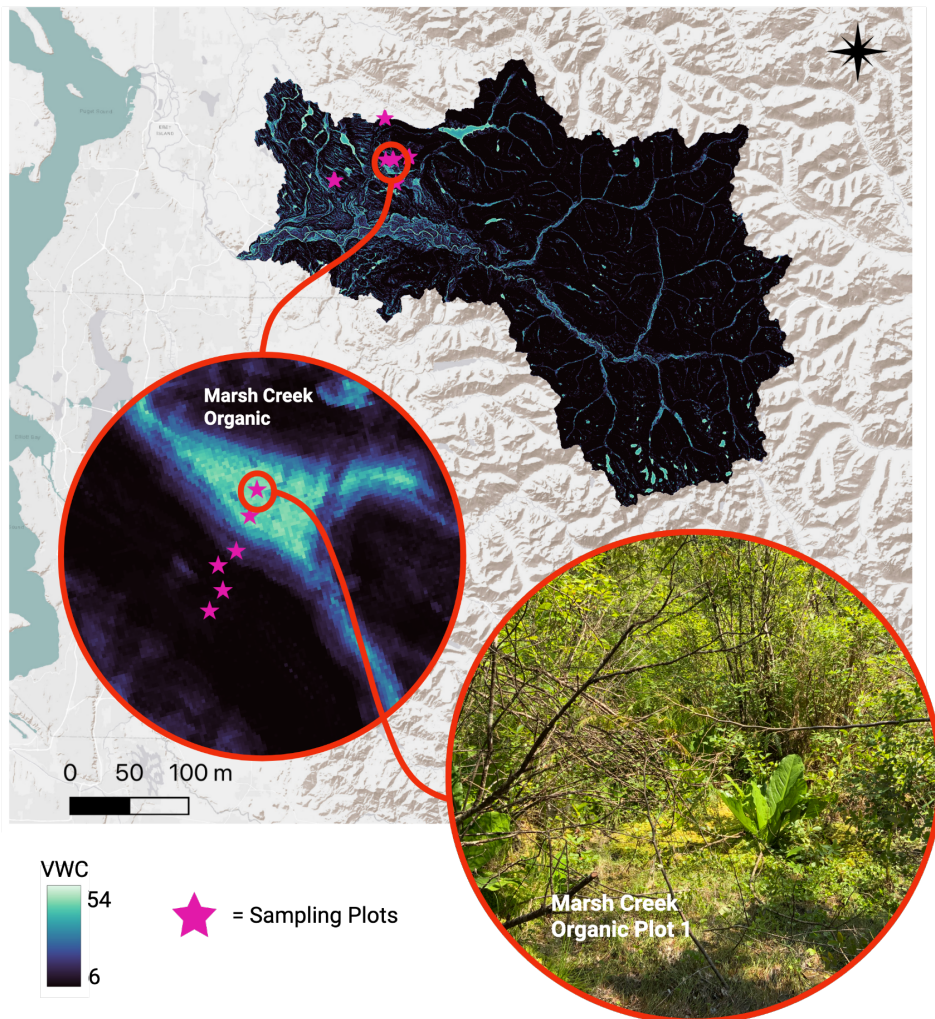


Figure 9. TealWaters WIP VWC Estimates Across the Watershed with a Zoomed in Map of Marsh Creek Organic

6. Discussion

6.1 Performance of Models in Predicting Soil Moisture

Our results support the effectiveness of using high-resolution topographic indexes to model soil moisture. This study extended the prediction of soil moisture into the complex forested terrain of Western Washington, similar to work in other regions using high-resolution topographic data for soil moisture applications (Ågren et al., 2021; Kemppinen et al., 2018;

Murphy et al., 2009; Riihimäki et al., 2021). Many soil moisture modeling methods rely on extensive soil moisture observations to directly model soil moisture. The work presented here offers a suitable alternative, highlighting the feasibility of utilizing wetland mapping and plant occurrence data in areas with sparse soil moisture observations.

The Hydrophytic Intrinsic Potential (HIP) model performed well in predicting soil moisture, due to the strong connection between moisture and plant community occurrence. Plant distribution can be explained as a function of the environmental factors, biotic, and abiotic drivers that facilitate occurrence, including soil properties such as moisture (Keddy, 2017). Conifer species occurrence along moisture gradients in southeast Alaska is driven by interlocking abiotic and biotic factors, including shifting moisture conditions, soil pH, nitrogen availability, and competition for light, operating across multiple spatial scales (Bisbing et al., 2016). Abiotic factors alone do not control community composition; rather, it is the feedback between abiotic and biotic factors that informs the distribution of species along the moisture gradient (Bisbing et al., 2016). These findings support our hypothesis of using plants as bioindicators of moisture along forested moisture gradients, given the control of topography over both abiotic and biotic processes.

The strong correlation between HIP and VWC observed here supports this fundamental link, suggesting that the probability of hydrophytic plants captures meaningful soil moisture variation at fine spatial scales. Soil moisture measurements can be challenging to implement, as soil moisture exhibits very high temporal variability. These metrics are thus just snapshots of moisture conditions, so multiple samples or longer-term monitoring sites must be established to paint the full soil moisture picture. Plants can help capture soil moisture signals, as their success and establishment are tied to moisture. This effectively allows plants to act as multiyear wetness

observations (Figure 10). The results from the HIP model validate the use of plant occurrence for moisture mapping, hinging on these fundamental drivers of plant distribution. Although soil moisture observations remain a vital tool, plant occurrence data can be used in tandem with them to depict wetness at spatially explicit scales, especially in areas with sparse soil moisture monitoring. While the concept of using wetland indicator status and wetland occurrence to assess soil moisture is novel, using plants as biological indicators is not.

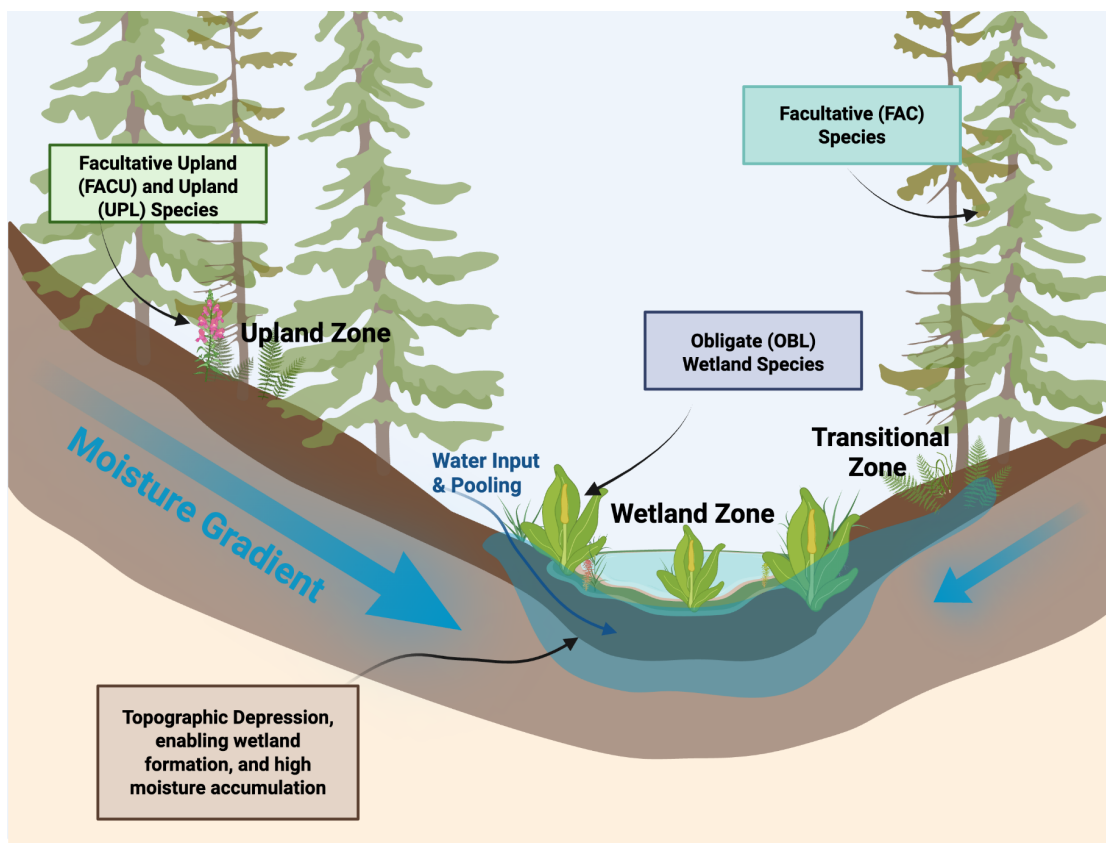


Figure 10. Conceptual Diagram of Plants as Moisture Proxies along a Moisture Gradient

In the US, the absence of explicit numeric indicators of soil moisture species occurrence, such as EIVs, led to the implementation of discrete moisture classifications. The USDA uses discrete classifications, such as drainage classes and soil moisture regimes, based on long-term moisture and temperature, and these regimes have associated plant community indicators. Recent

work by Paxton (2024) tested soil moisture regimes against field data, finding that plant-inferred soil moisture regime classifications did not align with in-situ soil moisture sensor data. Paxton also considered plant indicators and used the line-point intercept (LPI) method to assess plant community agreement with moisture regimes. This work found these indicators are not suitable proxies for soil moisture regimes, highlighting the importance of field soil moisture data (Paxton, 2024). Here, we explored similar questions in testing plants as proxies for moisture, but rather than testing plant communities' ability to act as direct proxies for a discrete moisture regime classification, we used the probability of indicator species occurrence or hydrophytic vegetation to model continuous soil moisture patterns.

Similar work presented in Ågren et al. (2021) tested a machine learning modeling framework, trained on a binary wet/dry dataset. Ågren demonstrates that high-resolution topographic indices can model soil moisture classes and probabilities at a national scale; however, this approach depends on a significant amount of field data (Ågren et al., 2021). Their model was fit with over 19,000 observations; this density of soil moisture observations is prohibitive for our study region, with more sparse soil moisture monitoring. Therefore, HIP and WIP model offer an alternative to the modeling methods presented in Ågren et al. (2021) in areas with sparse soil moisture data availability, while still leveraging key topographic-based indices such as TWI and DTW. The statistical evaluation methodologies between Ågren et al. (2021) and the work presented here differ. Ågren utilizes kappa values and overall accuracy, although we have calculated kappa values and OOB error rates; these metrics are not directly compared, given that we are not fitting models to soil moisture data directly but rather using regression to understand the relationship between model outputs and field soil moisture. Nevertheless, the

highest performing HIP explained 62% of the variance in soil moisture observations, suggesting that the model captures a meaningful proportion of the soil moisture gradient.

Out of all tested models, the statistics indicate that the HIP (10m) three-class model has the highest correlation with summer soil moisture compared to its binary model counterparts, although examination of the model output (Figure 7) revealed that the model is predicting much lower probabilities of wetness than the model fit with a binary wet and dry training split. This could be due to an inherent limitation of fitting a multiclass random forest model. When we predicted on the wet class, the output reflects the probability of a pixel being wet, not transitional or dry. Given that we have a transitional class, the model does not predict these as having a high probability of being wet. This is why the three-class model appears less bright or wet, because the transitional zones are not being mapped as high probability or even moderate probability of being wet, because the model knows they are transitional.

This helps to explain the higher numerical correlation with VWC. Although the model is "drier," the linear model, where we are getting the high R-squared, does not test if a probability is directly proportional to moisture, but rather the correlation of the relationship between the two, meaning when probabilities go up at a pixel, the VWC is also high, or vice versa. The three-class model's higher correlation with summer VWC, despite its weaker spatial display, could be attributed to its cleaner training data architecture. The three-class model is fit with UPL and FACU as the dry class, and FACW as the transitional class, with OBL as the wet class. Pulling apart the FACW and OBL allows the model to build stronger relationships at the wettest end of the spectrum, resulting in a higher R-squared with summer soil moisture.

The TWI WIP explained about 60% of the variation in summer soil moisture, suggesting that the WIP probability raster can be used as a proxy for moisture beyond pure wetland detection. Wetlands have historically been mapped as discrete polygons, with harsh boundaries. This work challenges this method and suggests that the WIP can serve as a continuous metric of moisture beyond the wet-dry binary. The "gray" in the WIP output represents areas of transitional moisture that are vital and should not be left out of wetland, moisture gradient modeling, and management decisions, especially in the face of shifting hydrological regimes due to climate change.

The Baseline WIP only explained 46% of the variation in summer soil moisture, indicating that modeling methodology, such as training data structure, resolution, and algorithm choice for TWI or DTW/HAND, plays a role in the model correlation to VWC. This is likely a product of differing training datasets, given that a model is driven by the quality of the data it's fit to; thus, the training data architecture and quality hugely impact model outputs. The TW WIP was fit with a smaller training dataset, with 700 wet points and 4,015 dry points, while the Baseline WIP was fit with 2,405 upland points and 1,511 wetland points. The TW WIP's training split reflects a more accurate proportion of wetland to upland area, as wetlands cover only a small portion of the land surface, and the TW WIP better reflects this.

Moreover, the Baseline WIP appears to produce less nuanced probability outputs, with an OOB error rate of 4.37%, indicating that the model had higher certainty in wetland areas, resulting in an output with less gradation in the probability raster. This could have driven the lower correlation with VWC. Lastly, the Baseline WIP is 5-meter resolution, and the TW WIP is 3 meters; the differences in resolution can have a large impact on correlation to soil moisture. For example, TWI's correlation to soil moisture will differ depending on resolution and raster

generalization. Given the differences in resolution between these 2 models, they are not fully comparable and should be treated as such. Here, we explored the effectiveness of using WIP as a proxy for moisture, however fully investigating the driving factors behind WIP model performance differences falls outside the scope of this work and warrants future investigation.

The WIP model is a powerful wetland-detection tool that has enabled the mapping of historically under-represented “cryptic wetlands” (Creed et al., 2003; Halabisky et al., 2023). In addition to wetland detection, Stewart et al. (2024) demonstrated the feasibility of using the WIP to map soil organic carbon (Stewart et al., 2024). The results presented here build on this body of work and support the idea that the WIP model extends beyond the binary of wetland and upland, depicting the spatial distribution of the entire moisture gradient, highlighting a wider use for the WIP model than pure wetland detection.

The HIP model’s ability to act as a soil moisture proxy varied across resolutions and training data splits, with the highest-performing correlation with summer soil moisture observed in the 10m three-class model, with a delta AICc of 0 ($R^2 = 0.62$). The TW WIP had the second lowest (delta AICc score of 3.15). The 10-meter HIP fit with a binary set of data from 2000 to 2026 had the third lowest Delta AICc (Delta AICc = 4.6). The HIP model is built on the core assumption that plant occurrence is a function of the abiotic environmental factors in which plants reside, allowing us to infer wetness or the probability of hydrophytic plants from these observations. The HIP model outperformed the WIP model as a soil moisture proxy; the best HIP model explains 62% of the variation in summer soil moisture data, whereas the best WIP model explains 60%. The difference in soil moisture correlation is likely because WIP is trained on wetland presence/absence data that may include unmapped wetlands and mesic areas. The HIP model training dataset shows greater specificity in upland areas, as points are derived from

observations rather than a random sample outside wetlands, allowing the model to build stronger relationships in those areas.

The HIP model outputs were masked to exclude areas above 500 meters because the lack of training data in these higher elevations led to an overestimation of moisture (Figure 11). The 500-meter cut off excluded much of the Skykomish watershed, resulting in a significant drawback to the HIP framework. This highlights the inherent spatial bias of using community-sourced plant observations. Observations will be clustered in areas where people recreate and are easily accessible, resulting in gaps of data in more remote locations with less recreation. Given this, the WIP model is more reliable in higher-elevation or remote locations. Although this work did not collect any soil moisture observations in areas above 500m, future work should consider validating soil moisture proxies such as HIP and WIP in high-altitude remote areas.

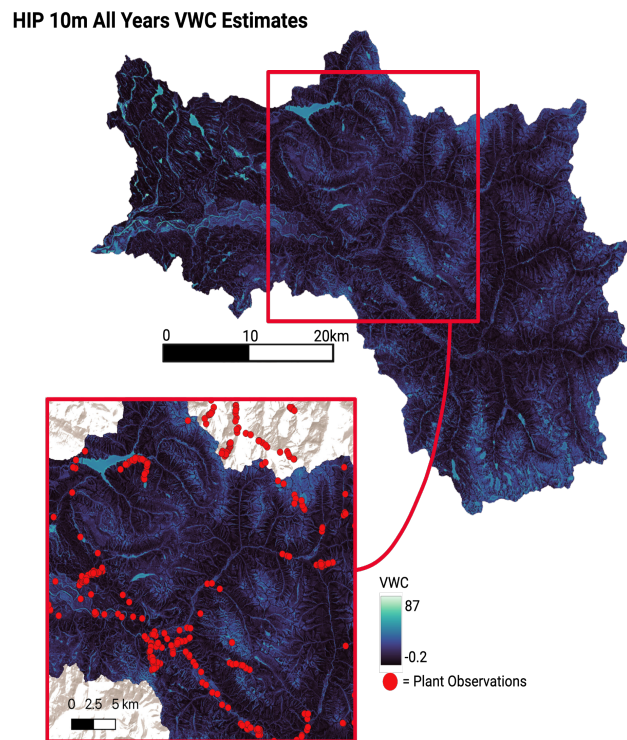


Figure 11. HIP 10m VWC Estimates

6.2 Where Do Models Fail to Act as a Soil Moisture Proxy?

From the residual analysis we determined that the HIP model has a tighter fit with normalized VWC in the lower moisture class, with higher normalized residuals in the wettest class. This indicates that VWC is higher than the model HIP probability in the wettest zones. This could be driven by an underrepresentation of specific wetland types in the training data. Including wetland training points from NWI alongside plant occurrence data could help reduce this underestimation. Moreover, it is noteworthy that the wettest group is defined as 30% VWC or higher, meaning 30% of the soil volume is water, resulting in a very wet soil. Depending on the modeling purpose, the HIP model may still be adequate. For example, if one wants to know where the soil is wet, this model performs relatively well, with predictive power beginning to decay only as soils approach saturation.

The residual analysis for both the baseline WIP and TW WIP revealed no systematic over- or under-prediction across moisture regimes, indicating that the models have consistent correlation with VWC across moisture zones. The positive correlation between clay percentage and the Baseline WIP residuals indicates that the model is underestimating in high-clay soils. This could be attributed to clay's high field capacity, leading to VWC measurements that exceed what topographic position alone can predict, because the model lacks soil textural information. This highlights the need for more spatially explicit and accurate soil physical property data to use as covariates in future soil moisture proxy models.

Moreover, areas of highly organic soil should also be met with more uncertainty across both the HIP and WIP models. Although there was no statistically significant correlation between LOI and model residual, the scatter plot indicates there is a downward (negative) trend. This is most evident in the 5m three-class model, suggesting that models are over-predicting in

these areas (Figure 12). This could be attributed to a known limitation of TDR, possibly leading to slightly underestimated VWC readings, resulting in high negative residuals in the highly organic soils.

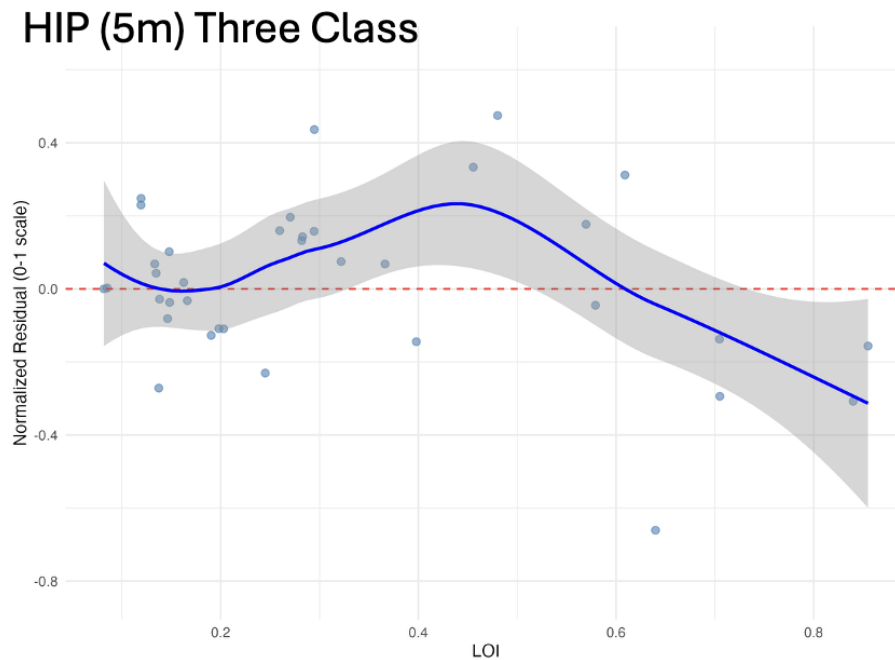


Figure 12. HIP (5m) Three Class Model Residuals and LOI

There are several key limitations of the TDR method one of which stems from the methodology used to convert the raw TDR waveform to VWC. Soil texture, bulk density, temperature, and soil organic matter all impact the relationship between VWC and dielectric permittivity a key variable in the empirical calibration function used in TDR (Jacobsen & Schjønning, 1993; Or & Wraith, 1999). In mineral soils, the dielectric permittivity (ϵ) is normally less than 40, varying by soil texture, whereas in soils with a higher water holding capacity, characterized by high amounts of pore space, like peat, the dielectric permittivity (ϵ) is often greater than 40. Therefore, the utilization of non-calibrated VWC risks the underestimation of VWC in peat and other porous soils (Nielsen & Thomsen, 2023). This suggests that perhaps

the model is not over-predicting, but rather the VWC measurements are under-predicting, given the high amount of organic matter in the soils. To combat this, future work should consider collecting the full waveform to perform texture-based/ percent organic TDR calibrations. Although this work used the standard setting on the Fieldscout 150 for consistency, VWC measurements in highly organic soil should be approached with more caution than their mineral counterparts.

In addition to TDR limitations, soil texture percentages are also subject to inherent laboratory error and limitations associated with the selected methodology. Here, we utilized a rapid texture assessment method outlined by Kettler et al. (2001). This method is advantageous for quick analysis because it allows for soil texture to be examined without removing organic matter, resulting in a fast soil texture estimation (Kettler et al., 2001). Although this method has some significant drawbacks. First, given that organic soils are not removed from the soil samples prior to performing texture analysis, it is possible that the sand fraction could be overestimated due to organics smaller than 2mm remaining in the sample and could possibly be counted in the sand fraction, although visible organic matter was removed to help mitigate this. Second, the separation of clay and silt particles may lead to an overestimation of clay content. The silt fraction is separated and removed by pouring off the suspended clay particles from a beaker. Although it's very challenging to remove all the silt fraction from the base of the beaker, there is some inherent loss during this step. The resulting loss will be counted as clay, given that clay is estimated as the remaining mass of the sample, and any accidental loss of sand will also compound this overestimation.

Given the limitations of the Kettler method, a sub-sample of soils was sent to Oregon State's Soil Health Lab for further analysis using the standard hydrometer texture analysis

method. The results confirm that the Kettler method systematically overestimates clay and underestimates silt. Although the average absolute difference across all samples was minimal, with only a 7% difference in sand, 11% in silt, and 9% in clay. With this in mind, residual results should be met with some caution for the soil particle percentages.

Moreover, alluvial areas or areas with coarser soil textures should be met with caution when interpreting the HIP and WIP outputs. No soil moisture data was taken in areas of alluvium, and we suspect that the model's ability to act as a soil moisture proxy degrades in these areas, given that these coarser textures have an inherently low field capacity. Similarly, areas near roadways should also be considered with caution, given the impact of roads on hydrological processes. Both drawbacks were also identified by Ågren et al. (2021); future work should consider sampling soil moisture along areas of alluvium to further explore the relationship between HIP/WIP and soil moisture.

It is also important to note that the small sample size is a real limitation, and all results from the residual analysis should be interpreted with this in mind. With only 36 total soil moisture observations, it is challenging to infer robust trends within moisture groups. For example, the wettest group contains only 9 observations, making it difficult to distinguish between outliers and actual trends within the data. Despite this limitation the analysis is still valuable in beginning to characterize where model accuracy decays along the moisture gradient. Future work should consider exploring how model residuals perform in relation to soil physical properties and moisture classes as demonstrated here, but with a larger dataset.

To our knowledge much of the current soil moisture modeling literature focuses on overall model accuracy and does not evaluate performance across drainage classes or soil physical properties; this represents an important gap for future investigation. Future work should

also consider exploring how plant community composition and ecosystem context impact model performance. Kopecký et al. (2021) explored TWI as a proxy for soil moisture and utilized the Bray-Curtis dissimilarity matrix to investigate TWI algorithm methods and explore where dissimilarity between plant communities could be explained by TWI. This is a novel approach to assessing TWI; similar methods utilizing the Bray-Curtis dissimilarity matrix to assess the WIP/HIP's ability to either predict plant community dissimilarity or evaluate whether model residuals degrade as a function of community dissimilarity would be a powerful addition to the literature (Kopecký et al., 2021).

Overall, both the HIP and WIP explain 50% or more of VWC variance in the summer months. The WIP has lower explanatory value overall, but residuals remain relatively consistent along the moisture gradient, indicating moderate explanatory power across normalized soil moisture observations throughout the full moisture gradient. The HIP model had higher explanatory value overall, with a clear pattern of underestimation of moisture in the wetter zone. This insight allows us to use the HIP with higher confidence in the driest zone, with lower confidence in the middle and upper moisture zones. The HIP is still an effective proxy for moisture and understanding where the model's ability to act as a proxy decays only increases its value as a tool for targeted soil moisture monitoring. Neither the WIP nor the HIP should be taken at face value as a replacement for soil moisture observations, but rather as a tool to better understand the spatial distribution of soil moisture for targeted monitoring efforts.

7. Conclusion

7.1 Management Implications and Future Work

This work has implications for both management and research. First, we demonstrated how the WIP model can serve as an effective proxy for soil moisture, illustrating the power of modeling wetlands as continuous metrics rather than discrete polygons. The knowledge that WIP does not simply detect wetlands but rather captures moisture variability across the full gradient opens the door to reimagining how wetlands are mapped. WIP can serve as a means to identify not only wetland areas, as demonstrated in Halabisky et al. (2023), but also the critical transition zones around them. Future researchers should investigate how frequently these transitional zones are inundated and how often these systems shift. Moreover, WIP can be used in tandem with wetland delineation to improve wetland mapping or serve as ancillary information alongside existing wetland maps. For soil moisture applications, researchers interested in fire, plant community convergence, carbon, and soil might consider using a similar framework to WIP to depict the spatial distribution of surface soil moisture in areas with sparse soil moisture observations, which hinder traditional soil moisture modeling.

The HIP method illustrates the utility of using plant occurrence data paired with wetland indicator status to map wetness, opening the doors for future advancements in moisture mapping in areas with sparse VWC data for model fitting. This method could also be effective for wetland mapping in areas where NWI is nonexistent or the quality of polygons is questionable. Moreover, this work supports the fundamental assumption that plants can act as soil moisture proxies; future work should consider expanding on this by including bryophyte species. The NWPL includes only vascular plants, but bryophytes, particularly *Sphagnum*, are characteristic

of peatlands. Including more bryophyte observations might improve model performance in areas with highly organic peat soils. This work attempted to explore this, but only 31 *Sphagnum* species observations were present in the plant occurrence dataset leveraged here, resulting in no meaningful change in model output due to the small sample size in the training data.

The HIP model framework could also be expanded upon to model specific plant ecosystems, such as meadows or forested wetlands. For example, a model could be fit with American skunk cabbage (*Lysichiton americanus*) and western hemlock (*Tsuga heterophylla*), western red cedar (*Thuja plicata*), and similar species to model possible forested wetland extent, given that these species are commonly associated with these wetland types. The HIP framework demonstrates that community science plant occurrence data can serve as a scalable and low-cost input for soil moisture proxy modeling in landscapes where in-situ observations remain sparse. This positions community science not simply as a supplementary data source, but as a foundational input for fine-scale environmental monitoring in heavily forested and data-limited landscapes.

The utilization of community science in environmental and landscape modeling is not novel. However, community science data comes with its own set of biases. Recent work by Moreno-Martinez et al. (2026) used GBIF plant occurrence data coupled with plant functional trait data and remote sensing to map plant functional traits on a global scale. This work effectively demonstrates the challenges in moving between plant occurrence and abundance data, given the uneven sampling design common in much community science data. The authors highlight the nonlinearity in the occurrence and abundance of species across ecosystems and present novel methods to account for it. For our soil moisture application, we rely solely on occurrence, since we are fitting a model based on topographic variables at fine scales. Although

future work aiming to model soil moisture at coarse scales might consider using an abundance framework, given that multiple observations would fall within a single pixel. With this in mind, future scaling efforts might consider adopting the novel methods presented by Moreno-Martínez et al. (2026) to adjust for non-linearity between occurrence, and abundance data.

The HIP and WIP frameworks not only provide qualitative moisture gradient proxies but, through VWC model calibration, can be used to estimate VWC across the watershed. This allows managers and researchers to derive spatially explicit VWC estimates from model outputs. However, given that we aimed to use the linear regression here as an explanatory tool for HIP and WIP's correlation with soil moisture, a holdout analysis was not conducted, and if extrapolating VWC beyond the scope of this work, a holdout analysis would be advised to further characterize the limitations of the VWC estimation as seen in Winzeler et al. (2022) and Halter et al. (2025).

Although the highest-performing models have a relatively strong fit with observed summer soil moisture, the VWC estimates should still be interpreted with caution, given that the relationship was developed from a small spatially constrained sample size of 36 observations. Moreover, the intercepts and slopes reported here are specific to the Skykomish watershed and the summer soil moisture conditions from which they were derived and should not be applied across differing watersheds or seasons without locally and climatically appropriate recalibration. Beyond static moisture estimation, the direct VWC estimates produced here offer a possible foundation for future soil moisture simulation modeling.

With a spatially explicit baseline VWC surface established for the Skykomish watershed, future researchers could apply projected shifts in climate to simulate fine-scale drought

vulnerability across the landscape. This is increasingly important as climate change is projected to drive subsequent hydrological regime shifts with cascading impacts on forest ecosystems (Crausbay et al., 2020). Further work is needed to investigate how these climatic shifts will impact soil moisture across the full moisture gradient, particularly given the likely non-linear relationships between climate drivers and soil moisture response.

7.2 Concluding Remarks

This research aimed to test the effectiveness of topographically based wetland probability models as proxies for surface soil moisture across the full moisture gradient in a heavily forested Pacific Northwest watershed. Based on our analysis, we demonstrated the effectiveness of using WIP not only as a proxy for soil moisture but also as a tool for directly estimating VWC, significantly expanding the application of the WIP model in both research and applied land management contexts. The WIP model explained 46–60% of the variance in summer soil moisture, with correlations varying across WIP modeling methods. These results support using WIP as a continuous moisture gradient tool rather than a purely binary wetland-detection model, opening new possibilities for land managers at the Washington DNR, stakeholders in the Skykomish watershed, and researchers working in data-limited forested systems.

Furthermore, **we investigated whether hydrophytic plant indicator models have predictive power over spatial patterns in soil moisture.** The HIP model showed a strong correlation with summer soil moisture, but its accuracy was limited to areas below 500 meters elevation due to sparse community science training data at higher elevations. The best-performing model explained 63% of the variance in the data, indicating the feasibility of utilizing

hydrophytic plant indicators as a scalable proxy for soil moisture in areas with sparse in-situ observations. However, when estimating VWC directly from HIP model outputs, predicted values were extrapolated outside the observed range, indicating that the calibration dataset does not adequately cover the full moisture gradient present in the model. The HIP model, therefore, illustrates general moisture trends effectively and can be used as a relative proxy for surface soil moisture in the lower 500 meters of the watershed, while the TW WIP model is recommended for applications requiring direct VWC estimation, given its more conservative and reliable predictions within the observed range.

To our knowledge, this is the first framework to couple community science wetland indicator observations with LiDAR-derived terrain metrics to produce a continuous soil moisture proxy. This approach opens the door to further research utilizing iNaturalist data for wetland and soil moisture applications and is especially meaningful for areas with limited or low-quality wetland inventory data, such as regions where the National Wetland Inventory (NWI) is absent or outdated. Moreover, the use of community science not only encourages broader community engagement in environmental monitoring but also offers an economically scalable approach to data collection. The density of observations that iNaturalist provides would be cost-prohibitive to replicate through traditional field campaigns.

Lastly, **we posed the question: where do these models show systematic over- or underprediction, and what soil properties drive model disagreement?** We found limited correlation between model residuals and soil physical properties across all variants of both HIP and WIP, indicating that soil texture, bulk density, and organic matter percentage have a limited impact on model performance as soil moisture proxies. The WIP normalized residuals showed no systematic correlation with VWC across the moisture gradient, whereas all HIP normalized

residuals showed a moderate positive correlation with VWC, indicating that the HIP model systematically underpredicts saturation in the wettest moisture classes. This underprediction likely reflects an underrepresentation of obligate wetland species in the community science training data. Future work should consider supplementing iNaturalist observations with NWI polygon training points in areas where the inventory is available and reliable, and relying on iNaturalist alone where it is not. We demonstrated a framework for estimating spatially explicit VWC from soil moisture proxy models trained on wetland occurrence data. Future work should aim to extend this into the temporal domain. HIP and WIP are static representations of the spatial distribution of surface moisture conditions, but a spatially explicit tool that evolves across seasons and years would substantially enhance researchers' and land managers' capacity to monitor and respond to dynamic soil moisture conditions.

Both the HIP and WIP models are effective proxies for the spatial distribution of surface soil moisture in the lower 500 meters of the Skykomish watershed. For direct VWC estimation, TW WIP should be considered the more conservative and reliable approach, given that the model displayed no systematic over- or underprediction. Soil moisture is a powerful but elusive variable that warrants spatially explicit mapping at management-relevant scales. Here, we demonstrated a reproducible, low-cost framework for mapping soil moisture at 10 meters or less using publicly available data providing a foundation for future work.

References

- Abrol, I. P., Khosla, B. K., & Bhumbla, D. R. (1968). Relationship of texture to some important soil moisture constants. *Geoderma*, 2(1), 33–39. [https://doi.org/10.1016/0016-7061\(68\)90004-9](https://doi.org/10.1016/0016-7061(68)90004-9)

- Ågren, A., Larson, J., Paul, S., Laudon, H., & Lidberg, W. (2021). Use of multiple LIDAR-derived digital terrain indices and machine learning for high-resolution national-scale soil moisture mapping of the Swedish forest landscape. *Geoderma*, 404, 115280. <https://doi.org/10.1016/j.geoderma.2021.115280>
- Ahanger, M. A., Siddique, K. H. M., & Ahmad, P. (2021). Understanding drought tolerance in plants. *Physiologia Plantarum*, 172(2), 286–288. <https://doi.org/10.1111/ppl.13442>
- Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer New York. https://doi.org/10.1007/978-1-4612-1694-0_15
- Amazirh, A., Merlin, O., Er-Raki, S., Gao, Q., Rivalland, V., Malbeteau, Y., Khabba, S., & Escorihuela, M. J. (2018). Retrieving surface soil moisture at high spatio-temporal resolution from a synergy between Sentinel-1 radar and Landsat thermal data: A study case over bare soil. *Remote Sensing of Environment*, 211, 321–337. <https://doi.org/10.1016/j.rse.2018.04.013>
- Asada, T., Warner, B. G., & Pojar, J. (2003). Environmental factors responsible for shaping an open peatland □ forest complex on the hypermaritime north coast of British Columbia. *Canadian Journal of Forest Research*, 33(12), 2380–2394. <https://doi.org/10.1139/x03-174>
- Ashok, A., Rani, H. P., & Jayakumar, K. V. (2021). Monitoring of dynamic wetland changes using NDVI and NDWI based landsat imagery. *Remote Sensing Applications: Society and Environment*, 23, 100547. <https://doi.org/10.1016/j.rsase.2021.100547>
- Beatty, S. W. (1984). Influence of Microtopography and Canopy Species on Spatial Patterns of Forest Understory Plants. *Ecology*, 65(5), 1406–1419. <https://doi.org/10.2307/1939121>
- Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie

- du bassin versant. *Hydrological Sciences Bulletin*, 24(1), 43–69.
<https://doi.org/10.1080/02626667909491834>
- Bisbing, S. M., Cooper, D. J., D'Amore, D. V., & Marshall, K. N. (2016). Determinants of conifer distributions across peatland to forest gradients in the coastal temperate rainforest of southeast Alaska. *Ecohydrology*, 9(2), 354–367. <https://doi.org/10.1002/eco.1640>
- Bonnell, R., Elder, K., McGrath, D., Marshall, H. P., Starr, B., Adebisi, N., Palomaki, R. T., & Hoppinen, Z. (2024). L-Band InSAR Snow Water Equivalent Retrieval Uncertainty Increases With Forest Cover Fraction. *Geophysical Research Letters*, 51(24), e2024GL111708. <https://doi.org/10.1029/2024GL111708>
- Booth, D. B., Haugerud, R. A., & Troost, K. G. (2003). 2. *The Geology of Puget Lowland Rivers*.
- Borden, R. W., Baillie, I. C., & Hallett, S. H. (2020). The East African contribution to the formalisation of the soil catena concept. *CATENA*, 185, 104291.
<https://doi.org/10.1016/j.catena.2019.104291>
- Buchanan, B. P., Fleming, M., Schneider, R. L., Richards, B. K., Archibald, J., Qiu, Z., & Walter, M. T. (2013). *Evaluating topographic wetness indices across central New York agricultural landscapes*. <https://doi.org/10.5194/hessd-10-14041-2013>
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33(2), 261–304.
<https://doi.org/10.1177/0049124104268644>
- Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M. H., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., ... Kerr, Y. (2016). Assessment of the SMAP Passive Soil Moisture Product. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8), 4994–5007.
<https://doi.org/10.1109/TGRS.2016.2561938>

- City of Everett. (2015). *Drinking water quality Report 2015*. City of Everett Public Works Department. <https://www.everettwa.gov/ArchiveCenter/ViewFile/Item/598>
- Clancy, M. A., Ruffing, B. E., Jarman, C., & Byrne, K. A. (2023). A novel method incorporating large rock fragments for improved soil bulk density and carbon stock estimation. *Soil Science Society of America Journal*, 87(5), 1029–1041.
<https://doi.org/10.1002/saj2.20562>
- Cowardin, L. M., Golet, F. C., & LaRoe, E. T. (1979). *Classification of wetlands and deepwater habitats of the United States*. U.S. Department of the Interior, Fish and Wildlife Service.
<https://www.fws.gov/wetlands/Documents/Classification-of-Wetlands-and-DeepwaterHabitats-of-the-United-States.pdf>
- Crausbay, S. D., Betancourt, J., Bradford, J., Cartwright, J., Dennison, W. C., Dunham, J., Enquist, C. A. F., Frazier, A. G., Hall, K. R., Littell, J. S., Luce, C. H., Palmer, R., Ramirez, A. R., Rangwala, I., Thompson, L., Walsh, B. M., & Carter, S. (2020). Unfamiliar Territory: Emerging Themes for Ecological Drought Research and Management. *One Earth*, 3(3), 337–353. <https://doi.org/10.1016/j.oneear.2020.08.019>
- Creed, I. F., Sanford, S. E., Beall, F. D., Molot, L. A., & Dillon, P. J. (2003). Cryptic wetlands: Integrating hidden wetlands in regression models of the export of dissolved organic carbon from forested landscapes. *Hydrological Processes*, 17(18), 3629–3648.
<https://doi.org/10.1002/hyp.1357>
- Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., & Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, 28(15), 2031–2064. <https://doi.org/10.1002/joc.1688>
- Dingman, L. (2015). Infiltration and Water Movement in Soils. In *Physical Hydrology* (Third). Waveland Pr Inc.

- Dong, J., Akbar, R., Short Gianotti, D. J., Feldman, A. F., Crow, W. T., & Entekhabi, D. (2022). Can Surface Soil Moisture Information Identify Evapotranspiration Regime Transitions? *Geophysical Research Letters*, 49(7), e2021GL097697.
<https://doi.org/10.1029/2021GL097697>
- Ellenberg, H. (1974). Zeigerwerte der Gefäßpflanzen Mitteleuropas (Indicator values of vascular plants in Central Europe). *Scripta Geobotanica*, 9, 1–97.
- 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>
- Gardner, W., & Kirkham, D. (1952). DETERMINATION OF SOIL MOISTURE BY NEUTRON SCATTERING. *Soil Science*, 73(5), 391.
- Gaskin, G. J., & Miller, J. D. (1996). Measurement of Soil Water Content Using a Simplified Impedance Measuring Technique. *Journal of Agricultural Engineering Research*, 63(2), 153–159. <https://doi.org/10.1006/jaer.1996.0017>
- GBIF.org. (2026). *GBIF Occurrence Download* GBIF.org [Dataset]. iNaturalist Research-grade Observations. <https://doi.org/https://doi.org/10.15468/dl.tfrn55>
- Gleick, P. (1993). *Water in Crisis: A Guide to the World's Fresh Water Resources*. Oxford University Press, New York.
- Guo, F., Huang, Z., Su, X., Li, Y., Luo, L., Ba, Y., Zhang, Z., & Yao, Y. (2024). Soil Moisture Content Inversion Model on the Basis of Sentinel Multispectral and Radar Satellite Remote Sensing Data. *Journal of Soil Science and Plant Nutrition*.
<https://doi.org/10.1007/s42729-024-02087-z>
- Hajek, O. L., & Knapp, A. K. (2022). Shifting seasonal patterns of water availability: Ecosystem responses to an unappreciated dimension of climate change. *New Phytologist*, 233(1), 119–125. <https://doi.org/10.1111/nph.17728>

- Halabisky, M., Miller, D., Stewart, A. J., Yahnke, A., Lorigan, D., Brasel, T., & Moskal, L. M. (2023). The Wetland Intrinsic Potential tool: Mapping wetland intrinsic potential through machine learning of multi-scale remote sensing proxies of wetland indicators. *Hydrology and Earth System Sciences*, 27(20), 3687–3699. <https://doi.org/10.5194/hess-27-3687-2023>
- Halter, T., Lehmann, P., Wicki, A., Aaron, J., & Stähli, M. (2025). Optimising landslide initiation modelling with high-resolution saturation prediction based on soil moisture monitoring data. *Landslides*, 22(9), 2897–2914. <https://doi.org/10.1007/s10346-024-02304-x>
- Hirschi, M., Mueller, B., Dorigo, W., & Seneviratne, S. I. (2014). Using remotely sensed soil moisture for land–atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability. *Remote Sensing of Environment*, 154, 246–252. <https://doi.org/10.1016/j.rse.2014.08.030>
- Huisman, J. A., Hubbard, S. S., Redman, J. D., & Annan, A. P. (2003). Measuring Soil Water Content with Ground Penetrating Radar: A Review. *Vadose Zone Journal*, 2(4), 476–491. <https://doi.org/10.2136/vzj2003.4760>
- Hutchinson, T. F., Boerner, R. E. J., Iverson, L. R., Sutherland, S., & Sutherland, E. K. (1999). Landscape patterns of understory composition and richness across a moisture and nitrogen mineralization gradient in Ohio (U.S.A.) Quercus forests. *Plant Ecology*, 144(2), 177–189. <https://doi.org/10.1023/A:1009804020976>
- Jacobsen, O. H., & Schjønning, P. (1993). A laboratory calibration of time domain reflectometry for soil water measurement including effects of bulk density and texture. *Journal of Hydrology*, 151(2–4), 147–157. [https://doi.org/10.1016/0022-1694\(93\)90233-Y](https://doi.org/10.1016/0022-1694(93)90233-Y)
- Jenny, H. (1941). *Factors of Soil Formation, A System of Quantitative Pedology*. Dover Publications.
<https://books.google.com/books?id=p94TjEnzCTcC&dq=Factors+of+Soil+Formation:+A+System+of+Quantitative+Pedology&printsec=frontcover&source=bn&hl=en&ei=>

- XfXZS_udG5OINtr74Ug&sa=X&oi=book_result&ct=result&resnum=4&ved=0CBkQ6AEwAw#v=onepage&q&f=false (Original work published 1941)
- Keddy, P. A. (2017). *Plant Ecology: Origins, Processes, Consequences*. Cambridge University Press.
- Kemppinen, J., Niittynen, P., Riihimäki, H., & Luoto, M. (2018). Modelling soil moisture in a high-latitude landscape using LiDAR and soil data. *Earth Surface Processes and Landforms*, 43(5), 1019–1031. <https://doi.org/10.1002/esp.4301>
- Kettler, T. A., Doran, J. W., & Gilbert, T. L. (2001). Simplified Method for Soil Particle-Size Determination to Accompany Soil-Quality Analyses. *Soil Science Society of America Journal*, 65(3), 849–852. <https://doi.org/10.2136/sssaj2001.653849x>
- Kopecký, M., Macek, M., & Wild, J. (2021). Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition. *Science of The Total Environment*, 757, 143785. <https://doi.org/10.1016/j.scitotenv.2020.143785>
- Li, W., Migliavacca, M., Forkel, M., Denissen, J. M. C., Reichstein, M., Yang, H., Duveiller, G., Weber, U., & Orth, R. (2022). Widespread increasing vegetation sensitivity to soil moisture. *Nature Communications*, 13(1), 3959. <https://doi.org/10.1038/s41467-022-31667-9>
- Liaw, A., & Wiener, M. (2022). *Classification and Regression by randomForest* (Version v4.7-1.2) [Computer software]. <https://CRAN.R-project.org/doc/Rnews/>.
- Lichvar, R. W., & Gillrich, J. J. (2011). *Final Protocol for Assigning Wetland Indicator Status Ratings during National Wetland Plant List Update*. US Army Corp of Engineers. https://wetland-plants.sec.usace.army.mil/static/references/NWPL/pubs/2011v3_Lichvar_Gillrich.pdf
- Lidberg, W., Nilsson, M., & Ågren, A. (2020). Using machine learning to generate high-resolution wet area maps for planning forest management: A study in a boreal forest landscape. *Ambio*, 49(2), 475–486. <https://doi.org/10.1007/s13280-019-01196-9>

- McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1–2), 3–52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Mesaglio, T., Shepherd, K. A., Wege, J. A., Barrett, R. L., Sauquet, H., & Cornwell, W. K. (2025). Expert identification blitz: A rapid high value approach for assessing and improving iNaturalist identification accuracy and data precision and confidence. *PLANTS, PEOPLE, PLANET*, 7(5), 1469–1484. <https://doi.org/10.1002/ppp3.70005>
- Milne, G. (1947). A Soil Reconnaissance Journey Through Parts of Tanganyika Territory December 1935 to February 1936. *Journal of Ecology*, 35(1/2), 192–265. <https://doi.org/10.2307/2256508>
- Moreira, A. C., Portezan Filho, O., Cavalcante, F. H. M., Coimbra, M. M., & Appoloni, C. R. (2007). Gamma ray transmission for hydraulic conductivity measurement of undisturbed soil columns. *Brazilian Archives of Biology and Technology*, 50(2), 321–328. <https://doi.org/10.1590/S1516-89132007000200017>
- Moreno-Martínez, Á., Muñoz-Marí, J., Adsua, J. E., Dechant, B., Kattge, J., Kattenborn, T., Sabatini, F. M., Schneider, F. D., Duveiller, G., Bodegom, P. M. V., Butler, E. E., Mahecha, M. D., Izquierdo-Verdiguier, E., Peñuelas, J., Boenisch, G., Pavlick, R., Townsend, P. A., Carvalhais, N., Lusk, D., & Camps-Valls, G. (2026). Leveraging remote sensing and crowd-sourced biodiversity data for enhanced plant functional trait mapping. *Nature Communications*. <https://doi.org/10.1038/s41467-026-72111-6>
- Munsell Color. (2012). *Munsell soil-color charts: With genuine Munsell color chips*. Grand Rapids, MI : Munsell Color, [2012].
- Murphy, P. N. C., Ogilvie, J., & Arp, P. (2009). Topographic modelling of soil moisture conditions: A comparison and verification of two models. *European Journal of Soil Science*, 60(1), 94–109. <https://doi.org/10.1111/j.1365-2389.2008.01094.x>

- Murphy, P. N. C., Ogilvie, J., Castonguay, M., Zhang, C., Meng, F.-R., & Arp, P. A. (2008). Improving forest operations planning through high-resolution flow-channel and wet-areas mapping. *The Forestry Chronicle*, 84(4), 568–574. <https://doi.org/10.5558/tfc84568-4>
- Murphy, P. N. C., Ogilvie, J., Connor, K., & Arp, P. A. (2007). Mapping wetlands: A comparison of two different approaches for New Brunswick, Canada. *Wetlands*, 27(4), 846–854. [https://doi.org/10.1672/0277-5212\(2007\)27%5B846:MWACOT%5D2.0.CO;2](https://doi.org/10.1672/0277-5212(2007)27%5B846:MWACOT%5D2.0.CO;2)
- Nielsen, C. K., & Thomsen, A. G. (2023). Local Calibration of TDR Measurements for Determining Water and Organic Carbon Contents of Peaty Soils. *Soil Systems*, 7(1), 10. <https://doi.org/10.3390/soilsystems7010010>
- Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., Waterloo, M., & Saleska, S. (2011). Height Above the Nearest Drainage – a hydrologically relevant new terrain model. *Journal of Hydrology*, 404(1–2), 13–29. <https://doi.org/10.1016/j.jhydrol.2011.03.051>
- Northwest Climate Hub. (n.d.). *Northwest Forests and Woodlands*. USDA. Retrieved April 21, 2025, from <https://www.climatehubs.usda.gov/hubs/northwest/topic/northwest-forests-and-woodlands#:~:text=Over%2085%20percent%20of%20forests,tree%20in%20all%20three%20states.>
- Or, D., & Wraith, J. M. (1999). Temperature effects on soil bulk dielectric permittivity measured by time domain reflectometry: A physical model. *Water Resources Research*, 35(2), 371–383. <https://doi.org/10.1029/1998WR900008>
- Orlandini, S., Moretti, G., Franchini, M., Aldighieri, B., & Testa, B. (2003). Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevation models. *Water Resources Research*, 39(6), 2002WR001639. <https://doi.org/10.1029/2002WR001639>

- Paxton, D. T. (2024). *Evaluation of Soil Moisture Regimes in the Inland Pacific Northwest, US* [M.S., University of Idaho].
<https://www.proquest.com/docview/3093327354/abstract/DDCCE9EF8784DCFPQ/1>
- Pedregosa, F., Pedregosa, F., Varoquaux, G., Varoquaux, G., Org, N., Gramfort, A., Gramfort, A., Michel, V., Michel, V., Fr, L., Thirion, B., Thirion, B., Grisel, O., Grisel, O., Blondel, M., Prettenhofer, P., Prettenhofer, P., Weiss, R., Dubourg, V., ... Cournapeau, D. (2011). Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.
- Peng, J., Albergel, C., Balenzano, A., Brocca, L., Cartus, O., Cosh, M. H., Crow, W. T., Dabrowska-Zielinska, K., Dadson, S., Davidson, M. W. J., De Rosnay, P., Dorigo, W., Gruber, A., Hagemann, S., Hirschi, M., Kerr, Y. H., Lovergine, F., Mahecha, M. D., Marzahn, P., ... Loew, A. (2021). A roadmap for high-resolution satellite soil moisture applications – confronting product characteristics with user requirements. *Remote Sensing of Environment*, 252, 112162. <https://doi.org/10.1016/j.rse.2020.112162>
- Peterson, A. M., Helgason, W. H., & Ireson, A. M. (2019). How Spatial Patterns of Soil Moisture Dynamics Can Explain Field-Scale Soil Moisture Variability: Observations From a Sodic Landscape. *Water Resources Research*, 55(5), 4410–4426.
<https://doi.org/10.1029/2018WR023329>
- Porporato, A., Daly, E., & Rodriguez-Iturbe, I. (2004). Soil Water Balance and Ecosystem Response to Climate Change. *The American Naturalist*, 164(5), 625–632.
<https://doi.org/10.1086/424970>
- Radmehr, R., Brorsen, B. W., Krueger, E. S., & Ochsner, T. E. (2026). Understanding the potential uses of soil moisture information in the United States. *Vadose Zone Journal*, 25(2), e70087. <https://doi.org/10.1002/vzj2.70087>
- Raduła, M. W., Szymura, T. H., & Szymura, M. (2018). Topographic wetness index explains soil moisture better than bioindication with Ellenberg's indicator values. *Ecological Indicators*, 85, 172–179. <https://doi.org/10.1016/j.ecolind.2017.10.011>

- Rasheed, M. W., Tang, J., Sarwar, A., Shah, S., Saddique, N., Khan, M. U., Imran Khan, M., Nawaz, S., Shamshiri, R. R., Aziz, M., & Sultan, M. (2022). Soil Moisture Measuring Techniques and Factors Affecting the Moisture Dynamics: A Comprehensive Review. *Sustainability*, 14(18), Article 18. <https://doi.org/10.3390/su141811538>
- Ratliff, L., Ritchie, J., & Cassel, D. (1983). Field-Measured Limits of Soil Water Availability as Related to Laboratory-Measured Properties¹. *Soil Science Society of America Journal - SSSAJ*, 47. <https://doi.org/10.2136/sssaj1983.03615995004700040032x>
- Riihimäki, H., Kemppinen, J., Kopecký, M., & Luoto, M. (2021). Topographic Wetness Index as a Proxy for Soil Moisture: The Importance of Flow-Routing Algorithm and Grid Resolution. *Water Resources Research*, 57(10), e2021WR029871. <https://doi.org/10.1029/2021WR029871>
- Robinson, D. A., Campbell, C. S., Hopmans, J. W., Hornbuckle, B. K., Jones, S. B., Knight, R., Ogden, F., Selker, J., & Wendroth, O. (2008). Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review. *Vadose Zone Journal*, 7(1), 358–389. <https://doi.org/10.2136/vzj2007.0143>
- Rodriguez, A. (2023). *Washington Soil Moisture Monitoring* (Version 1.0) [Map]. Seattle, Washington: Office of the Washington State Climatologist. <https://www.arcgis.com/apps/instant/interactivelegend/index.html?appid=49d2833971e04e488aff2263a27a7da5>
- Rowland, P. I., & Lovelock, C. E. (2024). Global impacts of introduced ungulates on wetland carbon and biodiversity: A review. *Biological Conservation*, 290, 110432. <https://doi.org/10.1016/j.biocon.2023.110432>
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010a). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3), 125–161. <https://doi.org/10.1016/j.earscirev.2010.02.004>

- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010b). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3–4), 125–161.
<https://doi.org/10.1016/j.earscirev.2010.02.004>
- Soil Science Division Staff. (2017). Soil Survey Manual 2017. *United States Department of Agriculture*, (No. 18). <https://www.nrcs.usda.gov/sites/default/files/2022-09/The-Soil-Survey-Manual.pdf>
- Soil Survey Staff. (n.d.). *Web Soil Survey* [Dataset]. Natural Resources Conservation Service, United States Department of Agriculture. Retrieved <https://websoilsurvey.nrcs.usda.gov/>. Accessed 2025.
- Soille, P. (2004). Morphological carving. *Pattern Recognition Letters*, 25(5), 543–550.
<https://doi.org/10.1016/j.patrec.2003.12.007>
- Soille, P., Vogt, J., & Colombo, R. (2003). Carving and adaptive drainage enforcement of grid digital elevation models. *Water Resources Research*, 39(12), 2002WR001879.
<https://doi.org/10.1029/2002WR001879>
- Southee, F. M., Treitz, P. M., & Scott, N. A. (2012). Application of Lidar Terrain Surfaces for Soil Moisture Modeling. *Photogrammetric Engineering & Remote Sensing*, 78(12), 1241–1251. <https://doi.org/10.14358/PERS.78.11.1241>
- Spectrum Technologies, Inc. (n.d.). *FieldScout® TDR 350 Soil Moisture Meter Product Manual* [Manual]. www.specmeters.com. Retrieved https://www.specmeters.com/core/media/media.nl?id=10572110&c=3703055&h=vbVYSZGW3UsoUXZVvP2H26tro9q52xCqgjANCwkfwneT7xHE&_xt=.pdf
- Sreeram, R., Adithya Krishna, S., Kumar, A. S., Remya, S., & Cho, Y. Y. (2026). Soil moisture monitoring technologies in smart agriculture: A comprehensive review. *Farming System*, 4(2), 100189. <https://doi.org/10.1016/j.farsys.2025.100189>

- Stewart, A. J., Halabisky, M., Babcock, C., Butman, D. E., D'Amore, D. V., & Moskal, L. M. (2024). Revealing the hidden carbon in forested wetland soils. *Nature Communications*, 15(1), 726. <https://doi.org/10.1038/s41467-024-44888-x>
- Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections: Further analysis of the data by akaike' s. *Communications in Statistics - Theory and Methods*, 7(1), 13–26. <https://doi.org/10.1080/03610927808827599>
- Tavakol, A., McDonough, K. R., Rahmani, V., Hutchinson, S. L., & Hutchinson, J. M. S. (2021). The soil moisture data bank: The ground-based, model-based, and satellite-based soil moisture data. *Remote Sensing Applications: Society and Environment*, 24, 100649. <https://doi.org/10.1016/j.rsase.2021.100649>
- Tichý, L., Axmanová, I., Dengler, J., Guarino, R., Jansen, F., Midolo, G., Nobis, M. P., Van Meerbeek, K., Ačić, S., Attorre, F., Bergmeier, E., Biurrun, I., Bonari, G., Bruelheide, H., Campos, J. A., Čarni, A., Chiarucci, A., Ćuk, M., Ćušterevska, R., ... Chytrý, M. (2023). Ellenberg-type indicator values for European vascular plant species. *Journal of Vegetation Science*, 34(1), e13168. <https://doi.org/10.1111/jvs.13168>
- 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>
- Tyystjärvi, V., Kempainen, J., Luoto, M., Aalto, T., Markkanen, T., Launiainen, S., Kieloaho, A.-J., & Aalto, J. (2022). Modelling spatio-temporal soil moisture dynamics in mountain tundra. *Hydrological Processes*, 36(1), e14450. <https://doi.org/10.1002/hyp.14450>
- Vereecken, H., Amelung, W., Bauke, S. L., Bogen, H., Brüggemann, N., Montzka, C., Vanderborght, J., Bechtold, M., Blöschl, G., Carminati, A., Javaux, M., Konings, A. G., Kusche, J., Neuweiler, I., Or, D., Steele-Dunne, S., Verhoef, A., Young, M., & Zhang, Y. (2022). Soil hydrology in the Earth system. *Nature Reviews Earth & Environment*, 3(9), 573–587. <https://doi.org/10.1038/s43017-022-00324-6>

- Vivian-Smith, G. (1997). Microtopographic Heterogeneity and Floristic Diversity in Experimental Wetland Communities. *Journal of Ecology*, 85(1), 71–82.
<https://doi.org/10.2307/2960628>
- Wang, C., Fu, B., Zhang, L., & Xu, Z. (2019). Soil moisture–plant interactions: An ecohydrological review. *Journal of Soils and Sediments*, 19(1), 1–9.
<https://doi.org/10.1007/s11368-018-2167-0>
- Washington State Department of Natural Resources. (2025). *RS-FRIS* (Version 5.2) [GIS dataset]. Washington State Department of Natural Resources GIS Open Data.
<https://data-wadnr.opendata.arcgis.com/datasets/wadnr::polygon-rs-fris/about>
- Whiting, M. L., Li, L., & Ustin, S. L. (2004). Predicting water content using Gaussian model on soil spectra. *Remote Sensing of Environment*, 89(4), 535–552.
<https://doi.org/10.1016/j.rse.2003.11.009>
- Winzeler, H. E., Owens, P. R., Read, Q. D., Libohova, Z., Ashworth, A., & Sauer, T. (2022). Topographic Wetness Index as a Proxy for Soil Moisture in a Hillslope Catena: Flow Algorithms and Map Generalization. *Land*, 11(11), 2018.
<https://doi.org/10.3390/land11112018>
- Zeng, Y., Hao, D., Huete, A., Dechant, B., Berry, J., Chen, J. M., Joiner, J., Frankenberg, C., Bond-Lamberty, B., Ryu, Y., Xiao, J., Asrar, G. R., & Chen, M. (2022). Optical vegetation indices for monitoring terrestrial ecosystems globally. *Nature Reviews Earth & Environment*, 3(7), 477–493. <https://doi.org/10.1038/s43017-022-00298-5>
- Zhang, K., Ali, A., Antonarakis, A., Moghaddam, M., Saatchi, S., Tabatabaeenejad, A., Chen, R., Jaruwatanadilok, S., Cuenca, R., Crow, W. T., & Moorcroft, P. (2019). The Sensitivity of North American Terrestrial Carbon Fluxes to Spatial and Temporal Variation in Soil Moisture: An Analysis Using Radar-Derived Estimates of Root-Zone Soil Moisture. *Journal of Geophysical Research: Biogeosciences*, 124(11), 3208–3231.
<https://doi.org/10.1029/2018JG004589>

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>

Appendix A. Loss on Ignition for Mineral and Organic Soils

Standard operating procedure

A 1.1. Mineral Soil LOI

Adapted from the USGS Mercury Research Laboratory LOI methods (2016)

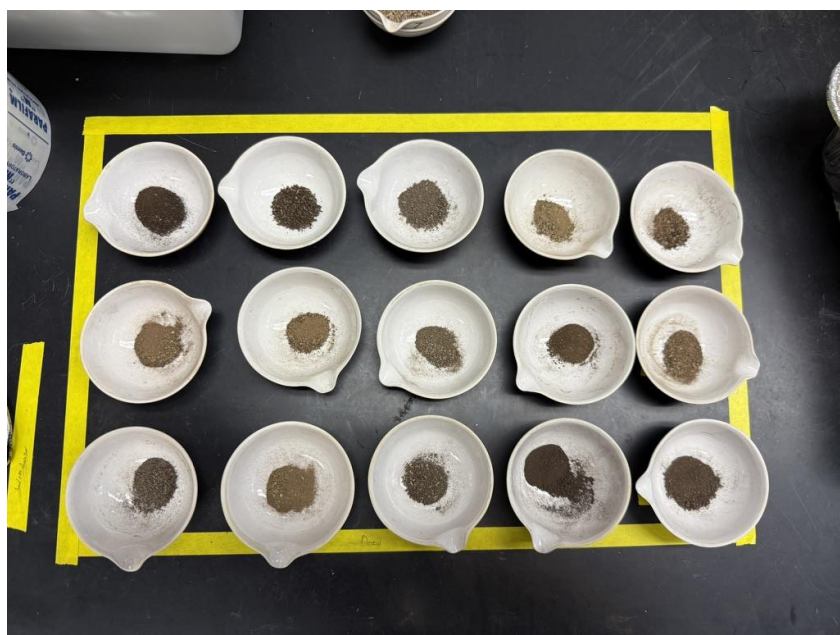
Description: Loss on Ignition (LOI) estimates the percentage of organic matter in a soil or sediment sample by burning off all organic matter in a muffle furnace. The samples are first dried at 105°C for 2 hours (or to a consistent weight) to remove any excess moisture. The samples are weighed to 4 decimal places on an analytical scale, and the dry weight is recorded. Next, the organic material is ignited by placing a sample in a furnace set at 550°C for 2 hours. After the samples have cooled enough to be removed from the furnace, they are stored in a desiccator until they are cool and can be weighed, and the LOI can be calculated.

Purpose and Scope: The LOI will yield the percentage of loss of mass on ignition of the sample, hence the name. This can be used to estimate soil organic matter percentage, given the assumption that all organic matter will ignite during the 2-hour period at 550°C. The percentage of organic matter in a soil can also be used to estimate soil organic carbon stocks and can be interpolated through the profile using bulk density.

Materials needed: Tin foil, Tape, Mortar and pestle, Scale, Muffle furnace, Desiccator, porcelain crucible

Procedure steps: these methods were written to be conducted on sets of 10-15 samples at a time. Each sample should be between 100mg to 500 mg (0.1g- 0.5 g). For our purposes, we used 0.5g for all samples.

Sample preparation: All samples should be oven-dried, ground in a mortar and pestle, and sieved to 2mm prior to performing LOI. This should be done before weighing out samples and placing them in the porcelain crucibles. All the crucibles should be numbered and weighed to get the “tin weight”. During this step, it's important to create a tin number key to keep track of all sample IDs when baking. Labels written in pen will burn off in the furnace, so this step is key, see example below.



In the picture above, you can see the key for all my samples and their corresponding positions. The bottom yellow tap reads Door, and the right-hand side reads window. This helps to orient where each sample is located in the furnace. Under each sample, there is a number 1 in the lower right corner and 15 in the top left corner.

Note: Boats made from tin foil will lose integrity at 550, and the bottoms might fall out; it's best to use premade aluminum boats or porcelain crucibles (see image below for example).



Procedure Steps

Step 1: Initial drying. This is not needed if the samples have been stored in a desiccator since preprocessing. Place all samples in a muffle furnace or drying oven set to 105 °C, and dry for 2h to drive off the residual hygroscopic moisture. Then store the sample in a desiccator till cooled, then take the mass using a scale accurate to the 4th decimal place. Record the dry weight.

Step 2: If not already recorded, weigh out all tin weights and record, then weigh out .1 to .5 grams of dried sample stored in the desiccator. Place in the crucible and place on the key in the correct location, or directly in the furnace. Take care to note the sample's plot ID and ID number for position in the furnace.

Step 3: After all samples are weighed, recorded, and placed in the furnace. Set the furnace to 550 °C, wait until it reaches temp, and set a 2-hour timer. After 2 hours have elapsed, begin cooling the furnace by setting the temp to 0. This can take some time and is best left overnight to cool. Note that it's also advised to crack a window during the 550 runs, and it's normal for the furnace to smoke a bit, especially around 200 °C as it comes to temp. If you have highly organic samples, it's best to space them out to reduce smoke. (Ideally, this would be done with the furnace inside a fume hood)

Step 4: Remove all samples and place them in the Desiccator. Make sure to label the samples before placing them in the Desiccator. I use foil to cover the sample and label it with a Sharpie.

Step 5: After the sample has cooled (which can take an hour or so), weigh it on a scale and record the final weight. Now, LOI and % LOI can be calculated using the equation below. I weighed all my LOIs 3 times and averaged the weights to ensure accuracy.

Tips for extracting the sample: All samples for this project were pulled from the plot sample bag. These samples were taken as a cross-section using a sharpshooter. The samples were then dried and stored in gallon freezer bags. The samples for LOI were taken from these bags; each bag was mixed, then a small sample of about 5 grams was removed and ground using the mortar and pestle (for about 1 minute) prior to sieving to 2mm.

Limitations: LOI at 550°C may overestimate organic matter content in high-clay soils because structural water within clay minerals is lost during combustion (Sun et al., 2009). Although the literature suggests that LOI is fairly accurate for wetland soils and soils with high SOM content, the structural water loss represents a smaller proportion of total mass loss (Sun et al., 2009). Thus, the high-clay samples in this study should be treated with more uncertainty. Additionally, single measurements per sample (rather than triplicate analyses) may reduce precision and the ability to detect subtle differences between plots.

Example LOI Equation set up

$$\text{Loss on Ignition} = \frac{\left(\left(\text{Mass of Tin} \right) - \left(\text{Tin Tare} \right) \right) - \left(\left(\text{Mass of Tin} \right) - \left(\text{Tin Tare} \right) \right)}{\left(\text{Mass of Tin} \right) - \left(\text{Tin Tare} \right)}$$

USGS Mercury Research Laboratory, 2016

A 1.1. Organic Soil LOI

Adapted from the USDA Kellogg Soil Survey Laboratory Methods Manual (Soil Survey Staff, 2014)

Description: The methods presented below measure soil mineral content and soil organic matter by combusting all non-mineral content. These methods differ from the LOI methods presented above in 2 distinct ways. 1. These methods use a lower temperature of 400 °C. 2. These methods require samples to be combusted for a much longer time than 16 hours, rather than the 4 hours used above.

Purpose and Scope: Loss on ignition can be used to measure the mineral content within a given sample. This value can then be subtracted to estimate the percentage of organic material and to estimate soil organic carbon (SOC) using the Van Bemmelen factor. The Van Bemmelen factor assumes that about 58% of all SOM is SOC; therefore, SOM values derived from LOI can be scaled by 0.58 to estimate the sample's SOC percentage.

Steps 1 and 2 remain the same as the above methods. Please refer to those steps for sample prep. The following steps will assume all samples are weighed and recorded

Step 3: All samples should be placed in the furnace at the appropriate location, with their IDs recorded on the key for easy identification after the LOI is complete. Once all samples are in place, set the furnace to 400 C and allow it to run overnight (16 hours)

Steps 4 and 5 remain the same as the methods above. Please refer to steps 4 and 5 above.

References:

Mercury Research Laboratory. "Analysis of Soils and Sediments for Loss on Ignition." Standard

Operating Procedure. U.S. Geological Survey, 2016.

<https://www.usgs.gov/media/files/analysis-soils-and-sediments-loss-ignition>.

Sun, H, M Nelson, F Chen, and J Husch. “Soil Mineral Structural Water Loss during Loss on Ignition Analyses.” *Canadian Journal of Soil Science* 89, no. 5 (2009): 603–10. <https://doi.org/10.4141/CJSS09007>.

Soil Survey Staff. 2014. Kellogg Soil Survey Laboratory Methods Manual. Soil Survey Investigations Report No. 42, Version 5.0. R. Burt and Soil Survey Staff (eds.). U.S. Department of Agriculture, Natural Resources Conservation Service.

Appendix B. Rapid Soil Particle-Size Analysis

Standard Operating Procedure

Description: These methods can be used to estimate the percentage of soil particles within a soil sample. The method is a wet sieve method. Using a dispersant agent and agitation prior to sizing soils to separate sand from silt and clay. The clay and silt are then separated using a sedimentation method. The sand and silt fraction is then dried to a consistent weight, and the clay is back-calculated from the original sample mass.

Purpose and Scope: Soil texture is a key parameter for understanding soil moisture because it influences field capacity, infiltration rates, bulk density, and overall soil structure. Here, we used a rapid soil texture analysis method rather than the standard hydrometer-and-pipette method outlined by Gee and Bauder (1986). (Gee and Bauder, 1986). Although this method is widely used and recognized for its accuracy, we will use a quick and less resource-intensive particle analysis method. First introduced in 2001 and outlined in Kettler et al (2001), this method was developed for Particulate organic matter (POM) and as a standalone texture analysis. However, here we will use only the texture analysis workflow. This method is particularly appealing because it does not require the removal of organic matter in the lab.

Materials Needed:

- Balance accurate to 0.1 g
 - 0.053-mm mesh sieve (No. 270)
 - Oven for sample drying (105°C)
 - Hexametaphosphate (HMP) solution (3% by weight)
 - Reciprocating shaker
 - 600-800 mL beaker
 - Container for shaking (ideally with lids, we used baby plastic Nalgene bottles)
 - 15 grams of soil sample

Procedure:

Step 1) Prepare the HMP

- 3% constriction of HMP by weight
- Measure out 30 grams of HMP in a glass beaker (or jug with a lid)
- Add in 1 liter of DI water (1,000 grams)
- Shake or stir until the HMP is dissolved
- Store in a tightly sealed container.

Step 2) Prepare the soil sample

- Collect 15 g of air-dried soil that has been sieved to <2 mm

Step 3) Chemical dispersion

- Place a 15 g soil sample in a container
- Add HMP solution in a 3:1 ratio (45 mL HMP to 15 g soil)
- Shake on the reciprocating shaker at 120 reciprocations per minute for 2 hours
 - We used 180 osc/min* (on low for our table)

Step 4) Sieve the sand fraction

- Pour the dispersed soil slurry through a 0.053-mm sieve (# 270)
 - Rinse sand particles remaining on the sieve until the water runs clear, take care to collect all the sample that passes through the sieve.
 - Note: There will likely be a lot of samples for our samples. We ended up with about five 600ml beakers for each sample.
 - Collect sand (>0.053 mm) and dry at 105°C to constant weight

Step 5) Collect the silt and clay solution

- Collect the solution passing through the sieve in a 600-800 mL beaker
- This solution contains silt and clay particles

Step 6) Sedimentation

- Stir the silt-clay solution thoroughly to suspend all particles
- Allow to settle undisturbed at room temperature (18-24°C) for at least 90 minutes but less than 6 hours

Step 7) Separate the clay from the silt

- After sedimentation, carefully decant the suspended clay fraction and discard
- The settled material at the bottom is the silt fraction

Step 8) Dry the silt fraction

- Dry the settled silt fraction in the beaker at 105°C to constant weight.
 - It is much easier to collect all the silt if you let it air dry for a day or so before trying to remove it from the beaker.

Calculate the soil texture

$$\text{Sand\%} = \left(\frac{\text{oven dry sand mass}}{\text{original sample mass}} \right) \times 100\%$$

$$\text{Silt\%} = \left(\frac{\text{oven dry silt mass}}{\text{original sample mass}} \right) \times 100\%$$

The clay% is determined by calculating the difference of 100% minus the sum of the Sand% and Silt%,

$$\text{Clay\%} = 100 - (\text{Sand\%} + \text{Silt\%})$$

(Kettler et al., 2001).

Gee, G. W., and J. W. Bauder. "Particle-Size Analysis." In *Methods of Soil Analysis*, 383–411. John Wiley & Sons, Ltd, 1986. <https://doi.org/10.2136/sssabookser5.1.2ed.c15>.

Kettler, T.A., J.W. Doran, and T.L. Gilbert. "Simplified Method for Soil Particle-Size Determination to Accompany Soil-Quality Analyses." *Soil Science Society of America Journal* 65, no. 3 (May 2001): 849–52. <https://doi.org/10.2136/sssaj2001.653849x>.

Appendix C. Bulk Density Methods

Purpose and scope: Bulk density is the weight of soil per unit volume and is used to measure compaction and can serve as an indicator of soil health (South Dakota State University Extension, 2023). In general, the higher the density, the less pore space available for water movement, root growth, root penetration, and seedling germination (USDA Agricultural Research Service & Natural Resources Conservation Service, 1999). Bulk density is calculated using an excavation method and varies depending on soil structural properties, and is heavily influenced by packing processes (Blake et al., 1986).

Description: To calculate bulk density, you must start with a known volume. For this project, we used a cylindrical cut of a PCP pipe with a volume of 165.88 cm³. Bulk Density will be expressed in grams per cubic centimeter (g/ cm³). Bulk density will be calculated for the entire soil sample, along with an additional bulk density for the fine-earth fraction, following the methods outlined in Clancy et al. (2023).

Step 1) Oven-dry all field moisture samples to a constant weight (our samples were dried at 70 °C for 7 days)

Step 2) Sort the dry sample by fragment size, removing all large rocks and roots greater than 2mm. Clancy et al. recommend sorting by hand first, then brushing each sample and using the 2mm sieve as the final step

Step 3) Obtain the mass of the larger than 2mm fraction of the sample.

Step 4) Obtain the volume of the larger fraction using either the water-displacement method or a consent for rock density (see Equation for large fragment volume using particle density).

Equation for large fragment volume using particle density

$$V_c = M \div PD$$

Where V is the volume, and M is the mass. PD is the particle density; here, 2.65 g cm⁻³ was used as a consent for the particle density (rock density), which is a common convention in soil sciences (Skopp, 2002; Rühlmann et al., 2006).

Equations for Bulk Density (BD)

$$BD_{total} = M_s \div V_s$$

$$BD_{fine} = M_f \div V_f$$

$$BD_{hybrid} = M_f \div V_s$$

Where M_f is the mass of the fine earth, and M_s is the mass of the total sample. V_s is the volume of the entire sample, and V_f is the volume of the fine earth samples.

Volume of the fine earth sample

$$V_f = V_s - V_c$$

Where V_s is the total volume, and V_c is the volume of the coarse fragments calculated above.

References

- Blake, G. R., & Hartge, K. H. (1986). Bulk density. In A. Klute (Ed.), *Methods of soil analysis: Part 1, Physical and mineralogical methods* (2nd ed., pp. 363-375). American Society of Agronomy. <https://doi.org/10.2136/sssabookser5.1.2ed.c13>
- Clancy, Michael A., Blair E. Ruffing, Caren Jarman, and Kenneth A. Byrne. "A Novel Method Incorporating Large Rock Fragments for Improved Soil Bulk Density and Carbon Stock Estimation." *Soil Science Society of America Journal* 87, no. 5 (2023): 1029–41. <https://doi.org/10.1002/saj2.20562>.
- USDA Agricultural Research Service & Natural Resources Conservation Service. (1999). "Soil

- quality test kit guide.” Soil Quality Institute.
https://efotg.sc.egov.usda.gov/references/public/WI/Soil_Quality_Test_Kit_Guide.pdf
- Rühlmann, Jörg, Martin Körschens, and Jan Graefe. “A New Approach to Calculate the Particle Density of Soils Considering Properties of the Soil Organic Matter and the Mineral Matrix.” *Geoderma* 130, nos. 3–4 (2006): 272–83.
<https://doi.org/10.1016/j.geoderma.2005.01.024>.
- South Dakota State University Extension. (2023, September 7). “Bulk density is an indicator of soil health.” USDA Natural Resources Conservation Service.
<https://extension.sdstate.edu/bulk-density-indicator-soil-health>
- Skopp, J. M. (2002). Physical properties of primary particles. In A. W. Warrick (Ed.), *Soil Physics Companion* (pp. 13–28). CRC Press. <https://doi.org/10.1201/9781420041651-4>

Appendix D. SoilRx Field Protocol

Site Assessment and Setup

This protocol aimed to monitor soil moisture along a soil catena, from a wetland upslope towards the uplands and drier soil moisture classes. The sample design will consist of 6 study sites; at each site, 6 plots will be conducted along the moisture gradient. The entire study area will be referred to as the site, and the sub-section for each soil moisture class will be referred to as the plot. The plot location should fall within the same general azimuth as all other plots, but the specific location is flexible. Plots should aim to be representative of the vegetation and moisture conditions of the part of the soil catena on which they are located. Sites will be predetermined based on their WIP variability, but it's possible that some sites will have less variability than previously thought; in this event, it's okay to establish more than one plot in a similar landscape. Under this framework, there will be a total of 6 sites with 36 subplots. These sites will be sampled during a 6-day sampling period in June, July, and August, resulting in VWC measurements from each site three times throughout the summer months.

Arrive on Site

Locate plot centers (the point)

- Assess site and plot criteria (accept or reject).

- The entire site may be rejected if there is a safety concern or if vegetation and moisture are uniform. In the event of a site rejection, a new backup site will be added at the end of the 6-day sampling period.

Site Assessment: All sites must fit the following criteria:

- Samples (VWC) must be taken from the soil
 - No bare rock.
 - No human-made surfaces
- At least 10 meters from a road
- If there is running water or a channel through the plot, that's ok, but VWC cannot be sampled in water, and soil samples should also be taken in soils following the same criteria as VWC points. Sites should be moved to the nearest appropriate location if they do not fit the criteria, and the adjustment should be noted in the site description. Additionally, circumstances such as active management or open clearcuts may be encountered. In the case of a clear-cut, the plot should be adjusted to a less disturbed area.
- **A site is rejected only if it cannot be adjusted to meet the criteria within 100 meters of the original sample point.**

Plot Set up: After the site is assessed and the first plot is placed, the remaining plots will be established along a 300m digital transect running in the direction of the highest moisture variability. This is delineated by selecting the azimuth with the highest Wetland Intrinsic Potential (WIP) variability. This is done visually using a WIP layer available on the iPad. Note that high variability should be within a 300-meter radius surrounding the site. After the direction of highest variability is selected, a 300-meter line can be drawn on the iPad (digital transect) to help guide the crew as they walk towards where plot 2 will be established. The following rules are used to establish new plots along this digital transect.

- A new plot should be placed when there is a change in moisture regime. Look for the following indicators:
 - Visible shift in hydrology: change in visible saturation, change in the presence or absence of scour lines from previous saturation, change in surface water
 - Shift in vegetation communities: this can be shifts in species composition and abundance within a plot, meaning that the shift can be in abundance, not just species
 - Change in landscape position: change in slope, change in surface curvature, change in relative position to a water body

As you are walking, look for any of the above indicators. Plots should be established when any of the above indicators are met, and you are at least 20 meters from the previous plot. Continue

walking the digital transect until 6 plots are set up or 300 meters are reached. If 6 plots were not delineated, consider revisiting the transect, paying attention to subtle changes. If there is really no change, even in slope, place the plots about 50 meters apart.

Where to place the plot along the transect? To avoid bias in plot placement, plots will be established alternating along the left and right side of the transect, facing upslope or away from plot 1.

Plot Center Measurements

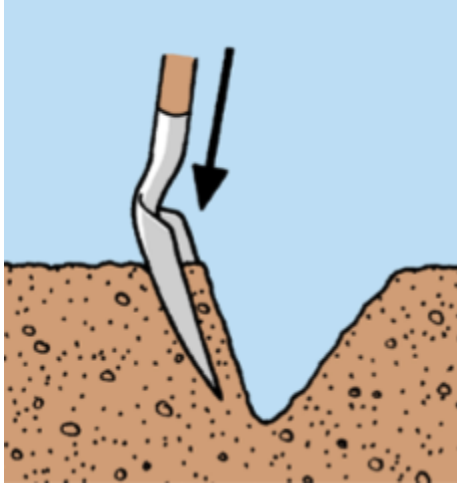
- GPS Coordinates
 - Record precise location coordinates
 - Record using the iPad
 - Record using a handheld GPS unit
 - Using the Javad: take the Javad out and place it on a stand, turn on the Javad, and let it warm up for at least 2 minutes before recording the point. Start recording and leave undisturbed for at least 30 minutes. Site IDs can be recorded using the Javad R app..
- Soil Moisture
 - Measure Volumetric Water Content
 - Using the Field Scout 150, collect a soil moisture measurement. Ensure the meter is fully pressed into the soil and that the 20cm rods are attached. Take the measurements as close to the plot center as possible. Make sure to remove the pin flag before taking the reading, as the metal can interfere with the TDR readings. Take the first plot center measurement, then 20cm to the right, take a second measurement, followed by the third, 20 cm to the left. These should be taken horizontally to the transect line.
 - Take care not to walk over or near the plot to minimize compaction and disturbance.

Soil Pit Data

Soil Sample Collection (this only needs to happen once at each plot)

- Dig a 20 cm hole, using a trowel and a sharpshooter
- Use the sharpshooter to cut a chunk of the side wall, so the sample is representative of the entire profile. (see diagram below)
- Collect a sample in a labeled bag (the sample should fill about ½ gallon freezer bag)
- Follow standard label nomenclature
 - Example plot id_D_M_Y
 - Please also add any notes for the sample to the data sheet

Figure 1



(King Conservation District)

Soil Characteristics Assessment

- Using the Munsell soil color book, record:
 - Soil color
 - Soil color should be taken in direct natural light, on moist soils.
 - Presence of redoximorphic features
 - Presence of rocks, cobble, and gravel
- If redox features are present, record their color separately, and the percent for the top 20cm
 - Document general soil notes:
 - Root presence
 - Any distinct smells
 - Organic matter on the surface
 - Leaf litter depth/composition
 - Determine soil hand texture using the texture dichotomous key (available in the Appendix)

Bulk Density (BD) Sample

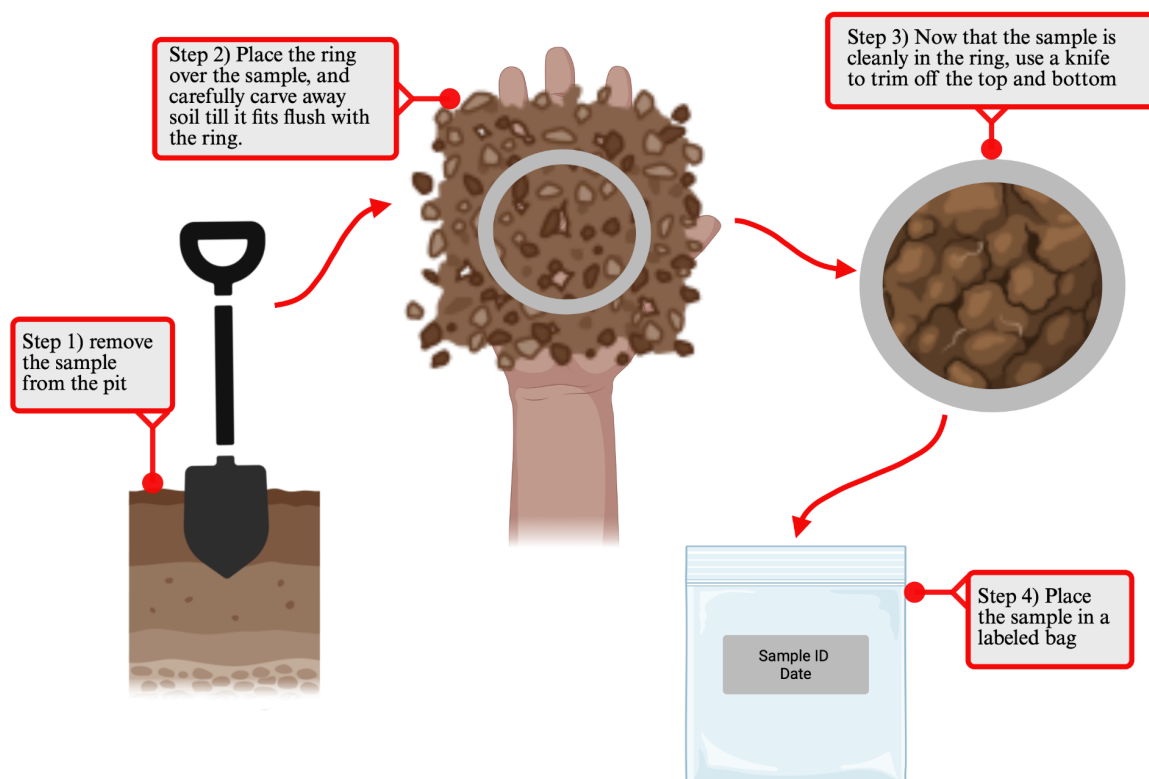
From the same 20cm hole dug for the soil pit sample, we will collect 2 bulk density (BD) samples using a bulk density core.

- Using the sharp shooter or trowel, extract a side cut of the soil pit, as done for the plot sample. Place the core on a representative part of the sample. Now, we will try to carve out the sample without compacting it.
- The most important thing here is not to compact the soil at all; it may take several tries to get a good sample.

- Using the flat knife, slice off the top and bottom excesses, so the soil is flush with the core (Figure 2)
- Deposit the sample into a plastic zip-lock back
- Labeled as follows
 - Example plotid_D_M_Y_BD
 - BD stands for bulk density here
- Weigh the sample using the field scale, ideally at the soi pit, but this can also be done in the car.
- Record the weight in the data log sheet.

Figure 2

Soil bulk density sampling procedure using a ring excavation method.



Site Photography

1. Required Photos

1. Take a site overview photo from approximately 5 feet above the ground
- c. Plot center photo
- e. Ensure all photos are properly labeled with site ID and date

Site Description

- Please note the presence of hydrology on the site, for example, scour lines, standing water, flowing water, etc..
- Take a slope measurement for the site

Botany

- Assess site dominant species composition, including canopy cover species
 - Using the acceptable keys (we mainly used Flora of the Pacific Northwest, along with the Washington Wildflower Search App), key out all unknown dominant species
 - Collected unknown species
- USNVC alliances, we used draft keys provided to us by the WA DNR; we used a wetland key and an upland key.

How to take a soil sample

Please refer to the King County conservation methods linked here:

<https://kingcd.org/wp-content/uploads/2025/03/Soil-Testing-2025.pdf>

Although our soil sample will be up to 20cm in order to match the depth of the soil moisture probes.

Soil Hand Texture

The NRCS key is linked here: <https://www.nrcs.usda.gov/sites/default/files/2022-11/texture-by-feel.pdf>

*****The above information will also be available as an offline resource on each iPad .

References

- Southee, Florence Margaret, Paul M. Treitz, and Neal A. Scott. "Application of Lidar Terrain Surfaces for Soil Moisture Modeling." *Photogrammetric Engineering & Remote Sensing* 78, no. 12 (December 1, 2012): 1241–51. <https://doi.org/10.14358/PERS.78.11.1241>.
- Stewart, Anthony J., Meghan Halabisky, Chad Babcock, David E. Butman, David V., D'Amore, and L. Monika Moskal. "Revealing the Hidden Carbon in Forested Wetland Soils." *Nature Communications* 15, no. 1 (January 25, 2024):726. <https://doi.org/10.1038/s41467-024-44888-x>.
- Kemppinen, Julia, Pekka Niittynen, Henri Riihimäki, and Miska Luoto. "Modelling Soil Moisture in a High-Latitude Landscape Using LiDAR and Soil Data." *Earth Surface Processes and Landforms* 43, no. 5 (2018): 1019–31. <https://doi.org/10.1002/esp.4301>.
- Riihimäki, H., Kemppinen, J., Kopecký, M., & Luoto, M. (2021). Topographic Wetness Index as a proxy for soil moisture: The importance of flow-routing algorithm and grid resolution. *Water Resources Research*, 57, e2021WR029871. <https://doi.org/10.1029/2021WR029871>

Appendix E. Linear Mixed Effects Model

Introduction and Methods

Over the summer of 2025, 36 VWC data were collected across 6 study sites; all plots were sampled 3 times, yielding a total of 108 plot VWC measurements. Summer VWC was then calculated as the average for each plot, yielding 36 average VWC measurements. We explored the relationships among VWC, site characteristics, and seasonal patterns using a linear model to identify dependencies within our dataset prior to running predictive VWC models. The model can be expressed as follows:

```
lm_model <- lm(Plot.VWC ~ Month * Site_ID, data = soil_data)
```

The linear model revealed differing intercepts between sites, indicating that VWC baselines (or baseline moisture conditions) differ between sites. Moreover, the differing slopes for each site show that sites dry at slightly different rates, ranging from -5.09% to -2.11% VWC per month between June and August. However, all p-values were greater than 0.67, indicating that these differences in temporal trends are not statistically significant. Suggesting that sites have differing baseline moisture conditions, with no statistically significant differences in drying rates. A larger sample size with a higher frequency of VWC measurements is necessary to fully understand the differences in drying rates between sites, but for the 36 sites presented here, there is no statistical evidence of differing drying rates among sites.

Given the differing intercepts between sites, we tested using a linear mixed effects model to predict VWC across the watershed. The linear mixed-effects model was selected because it allows for site-level attributes. This is helpful because the model will have separate baselines (intercepts) for each site. For example, it knows that Site 2 is baseline wetter than Site 1, while maintaining the same relationship between topography and moisture across all sites. In the linear

model with only fixed variables, the coefficients for each independent variable risk becoming inflated or deflated by sites with dramatically higher or lower baseline moisture conditions. The random effect separates these site-level differences, allowing the model to estimate more accurate topographic relationships that apply consistently across sites. The LMM model framework works well for the nested structure of the VWC data collected in the Skykomish. Summer 2025 VWC data were used to fit the LMM with site ID as a random effect. The model covariates were selected based on the correlation between topography and hydrological indices used in the 5-meter WIP and HIP models. Only 5 covariates were selected, given the relatively small number of VWC observations used to fit the model: DEV 100m, PROF 100m, DTW, and TWI. The LMM model was fit using the *lmer* function from the lme4 library in R (Bates et al., 2015).

$$\text{VWC} \sim \text{TWI} + \text{DEV} + \text{Prof} + \text{DTW} + (1|\text{Site ID})$$

Given that the model knows the site-level variations, the predictions are close to the observations, resulting in an almost perfect relationship and inflated accuracy metrics. To mitigate this and accurately assess the model's predictive power, a holdout cross-validation (CV) was performed. The CV was performed at both the site level and plot level. For the site-level CV, the model was fit on 30 plots, with one site held back for testing per fold, for a total of 6 folds. The plot-level CV comprised 36 folds, with each model fit on 35 VWC observations.

Results

For the site-level CV, the model was fit on 30 plots, with one site held back for testing per fold. A total of 6 models were fit, and the total R-squared across all 6 folds was 0.33, and the Mean Absolute Error (MAE) was 13.54. Looking at site-specific results, the site with the highest R-squared was CSB, with an R-squared of 0.95 and a MAE of 8.99, and the lowest R-squared was at Echo Lake, at 0.19 and a MAE of 15.51. This range is notable and indicates that Echo Lake had distinct characteristics compared to the other sites the model was fit to. This is not surprising given that the plots at Echo Lake fall in and around a peatland, with drastically different botany and hydrology from other sites.

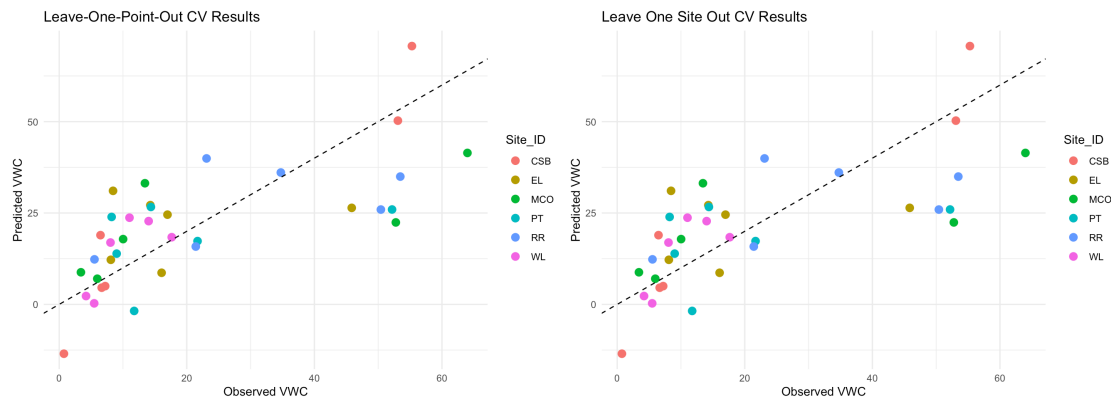
The plot-level CV used 36 folds, with one plot held out per fold for testing. The overall R-squared was 0.48, and the MAE was 11.08. Looking at the site-level R-squared, CSB still had the highest value at 0.90, with a MAE of 8.20. The lowest site level R-squared remained at Echo Lake; the R-squared was lower at 0.04, and the MAE was 12.33. The plot-level CV likely has a higher overall R-squared value because when only one plot from each site was held back, the model still learned site-level characteristics from the other 5 plots at that same site. This made it easier to predict the held-out plot because the model could lean on the site baseline established by that site's random effect.

Discussion

Both these cross-validations help deepen our understanding of the relationship between soil moisture and topographic indices. The results indicate that the site and plot level characteristics outside of the topographic variables used in this model have a strong predictive power over soil moisture, and thus, this modeling framework is missing key insights that might

help to more accurately predict soil moisture. Although other variables that likely impact plot-level soil moisture are challenging to obtain on a spatially continuous scale, for example, soil texture and percent organic matter likely have a strong impact on soil moisture, but are challenging to measure on spatially continuous scales, given their heterogeneity.

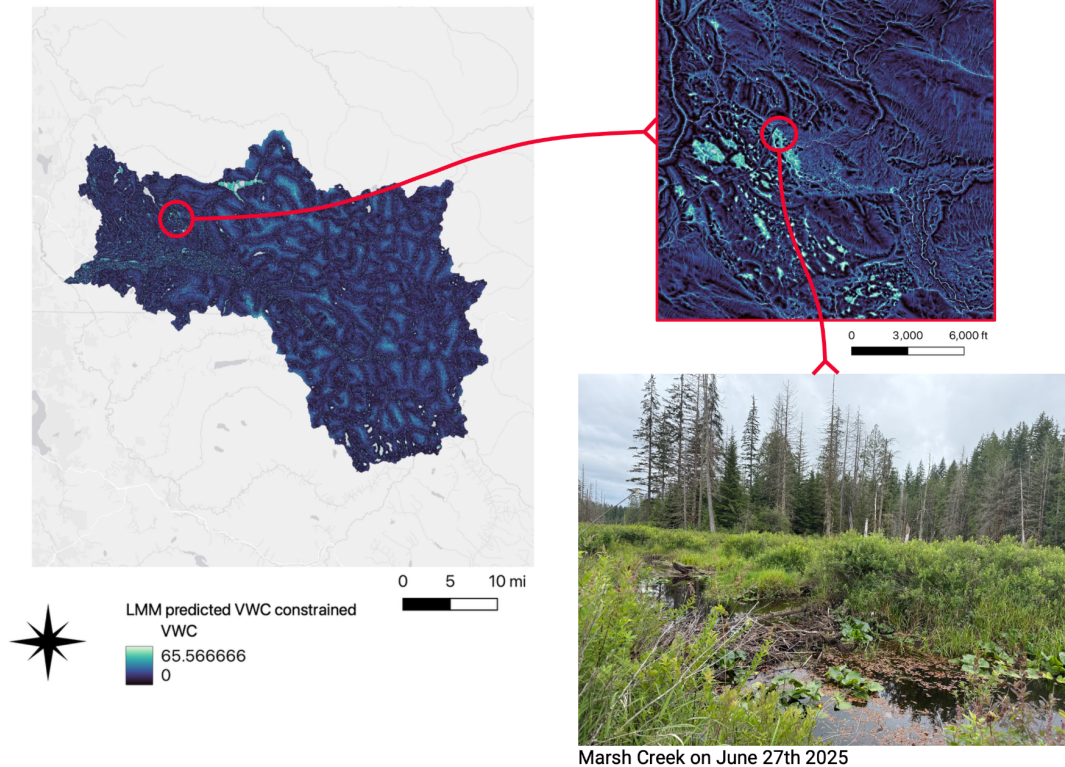
Scatter Plots from Both Cross-Validations



Note: Both models had a tighter fit in the dryer portion of the moisture gradient. There is no visible difference between the models; however, the plot-level CV (left side) had slightly higher R-squared, indicating a stronger fit.

Given that we want to use these models to project over a larger spatial extent, we used the model with the Plot level CV, since we will be predicting in areas with unknown site characteristics, and we are not confident that our field data covers all possible site types in the watershed. When we run the prediction across the entire raster, we get very extrapolated values well outside the expected VWC range (-50 to >100). This indicates that the model is encountering relationships between covariates beyond those observed in the training dataset. We used a clamping function to force values into the expected VWC range in order to visualize where the model was extrapolating. It is clear that the model struggles in elevations above 500 meters, given that the models were fit with data constrained to the lower western extent of the watershed. Given these results, the LMM was not used moving forward. The LMM application should be revised to use a larger training dataset that encompasses a more representative sample of site types and hydrological regimes to achieve accurate, watershed-wide VWC predictions.

Linear Mixed Effects Model VWC Estimates



Note: The VWC estimates are from the Linneax mixed effects model, expressed as $VWC \sim TWI + DEV + Prof + DTW + (1|Site ID)$. The output was constrained using a clamping function from -50 to >100 to 0 to 65 for interpretability. On the left-hand side, the entire Skykomish output is shown; however, areas above 500 meters are not reliable, as noted above. On the right, Marsh Creek is a wet shrubland characterized by *Salix* spp. - *Spiraea douglasii* / *Carex* (aquatilis var. dives, obnupta, utriculata).

References

Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48.
doi:10.18637/jss.v067.i01.

Appendix F. Topographic Wetness Index (TWI) in the Skykomish

TWI is a well-established proxy for soil moisture, although the literature suggests that the methodology used in calculating TWI has a large impact on the index's predictive power over

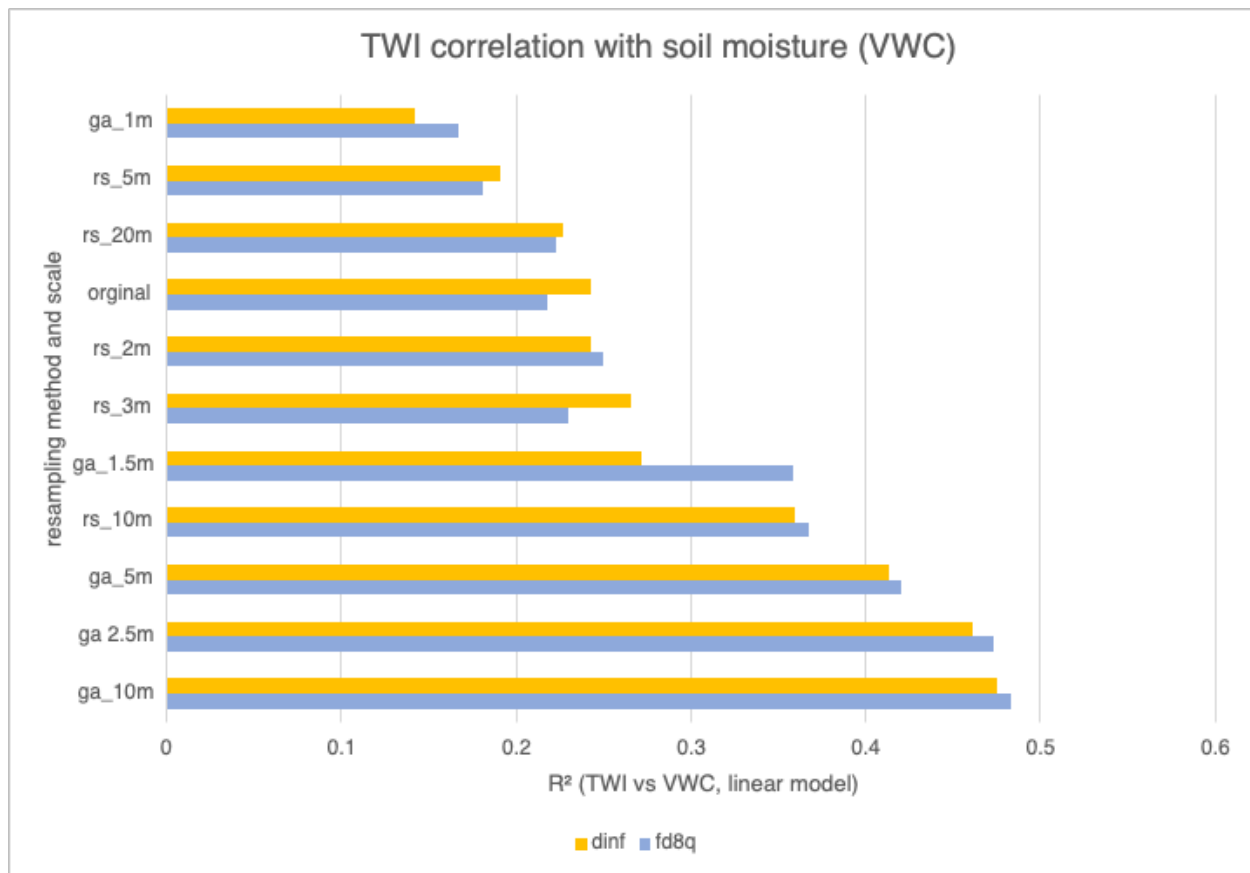
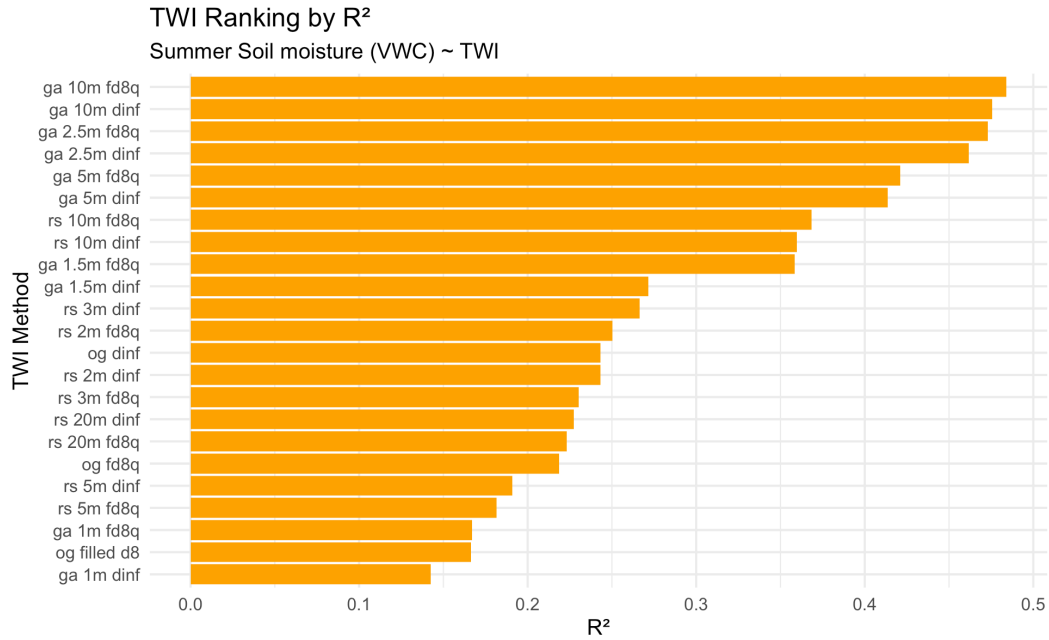
soil moisture (Sørensen et al., 2005, Grabs et al., 2009, Southee et al., 2012, Buchanan et al., 2013, Kopecký et al., 2021, Meles et al., 2020, Riihimäki et al., 2021, and Winzeler et al., 2022). Here, we tested several TWI calculation methods, focusing on raster generalization, resolution, and the choice of flow accumulation algorithm. This workflow was executed in Python and leveraged Whitebox tools. TWI testing was constrained to the Skykomish watershed, and a 1-meter lidar-derived DEM was used as the baseline for all analysis.

The raster generalization methods tested were resample (rs) (WBT resample function), a Gaussian filter (GA) (WBT gaussian_filter), and original (og), meaning no raster generalization was applied, and TWI was calculated on the native resolution of the DEM. The Gaussian filter was tested at scales of 1, 1.5, 2.5, 5, and 10 meters, and the resampling at scales of 2, 3, 5, 10, and 20 meters. Hydrological filling and breaching methods were also tested. For the breach, the WBT BreachDepressionsLeastCost function was used with a set distance of 20 for all (Lindsay and Dhun, 2015). The filling was applied using the WBT's FillBurn function (Saunders, 1999, and Lindsay et al., 2016).

The flow accumulation algorithm testing focused on the Multiple-flow Quinn (FD8f) and Deterministic infinity (dinf), both implemented using WBTs. The Deterministic infinity, first described in Tarboton (1997), is a multi-flow direction algorithm in which flow entering each cell is rotated to one or two cells. In WBTs, this is the d_inf_flow_accumulation function. The Multiple-flow Quinn introduced in Quinn et al. (1991), is also classified as a multi-flow direction algorithm. Lastly, naive TWI was also generated using the original raster resolution, filled for smoothing, and the deterministic single-flow- d8 (D8) flow accumulation algorithm. The D8 differs from the other algorithms tested in this study because it is a single flow algorithm, meaning that flow diversion is not permitted (O'Callaghan and Mark, 1984). After the above parameters were tested, a total of 23 TWI raster's were generated. All scripts and TWI calculations were executed by Tealwaters.

We explored whether summer soil moisture in the Skykomish watershed can be explained as a function of TWI, using a linear model (LM (vwc~ twi)). The results indicate that the raster generalization method and the choice of flow accumulation algorithm affect TWI's correlation with soil moisture. The TWI outputs explained 14% to 48% of the variance in summer soil moisture data, depending on the flow accumulation algorithm, raster generalization methods, and spatial resolution. The Gaussian filter at 10m using the Multi-Flow Quinn (fd8q) had the highest R-squared (0.48 and $p < 0.001$), and the lowest AICc (295.49). The second-highest scoring output was also the Gaussian filter at 10m, but with the Deterministic Infinity (dinf), although this output was not dramatically behind the fd8q TWI with an R-squared of 0.48 and an AICc of 296.10. Looking at $\Delta AICc$, we can use the substantial support threshold of 2 to compare top-performing TWIs (Burnham and Anderson, 2002). This indicates that there is no statistical difference among the top 4 performing models in their correlation with summer soil moisture in the Skykomish watershed. This suggests that the Gaussian filter at 10 or 2.5 meters with either fd8q or dinf are statistically equivalent in terms of model AICc score. The three lowest-ranking TWI were all at the highest spatial resolution tested (1 meter), with the Gaussian filter at 1 meter using the dinf ranking lowest ($R^2 = 0.14$, $P = 0.023$).

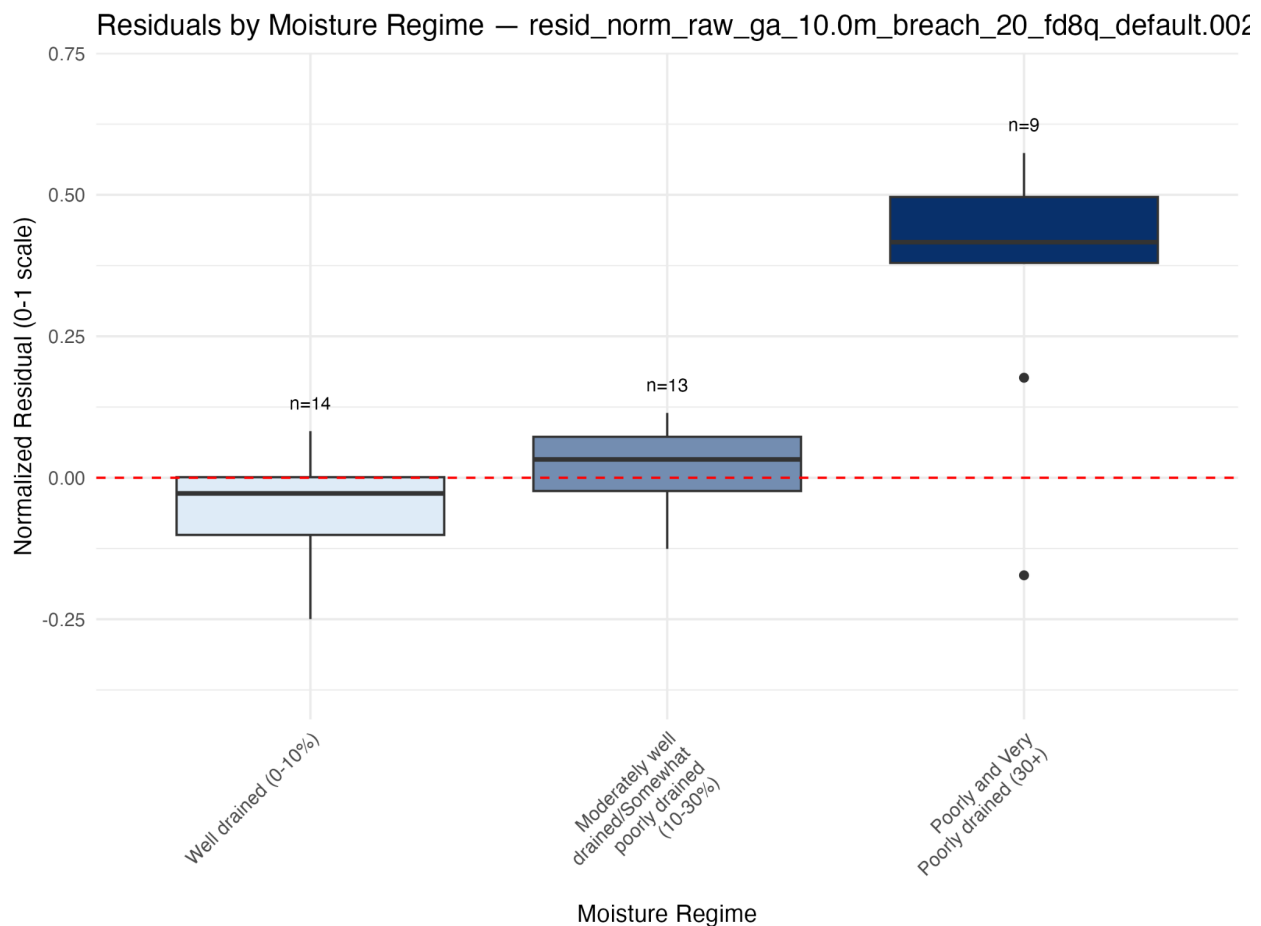
TWI Ranked by R-Squared



TWI Residual Analysis Results

Across the top three performing TWI models (ga 10m fd8q, ga 10m dinf, ga 2.5m fd8q), normalized residuals show a strong and consistent positive correlation with observed VWC ($r = 0.78, 0.78, \text{ and } 0.73$, respectively, all $p < 0.001$), indicating systematic underprediction in wetter plots. Residuals are also significantly positively correlated with LOI across all three models ($r = 0.51, 0.52, 0.49$, all $p \leq 0.002$), and significantly negatively correlated with bulk density ($r = -0.51, -0.50, -0.53$, all $p \leq 0.002$). Silt shows a consistent negative correlation across all three models ($r = -0.37, -0.40, -0.36$, all $p \leq 0.031$), while sand and clay show no significant relationships.

TWI Normalized Residuals by Drainage Class



Note: The residual from a normalized TWI raster created using a Gaussian filter at 10m using the fd8q show systematic under prediction of VWC in the highest end of the moisture spectrum.

References

- Liaw A, Wiener M (2002). "Classification and Regression by randomForest." R News, 2(3), 18-22. <https://CRAN.R-project.org/doc/Rnews/>.
- Lindsay J, Dhun K. 2015. Modelling surface drainage patterns in altered landscapes using

- LiDAR. *International Journal of Geographical Information Science*, 29: 1-15. DOI: 10.1080/13658816.2014.975715
- Lindsay JB. 2016. The practice of DEM stream burning revisited. *Earth Surface Processes and Landforms*, 41(5): 658-668. DOI: 10.1002/esp.3888
- O'Callaghan JF, Mark DM. 1984. The extraction of drainage networks from digital elevation data. *Computer Vision, Graphics, and Image Processing*, 28(3), 323-344.
- Quinn PF, Beven KJ, Lamb R. 1995. The in ($a/\tan\beta$) index: How to calculate it and how to use it within the top model framework. *Hydrological Processes* 9(2): 161-182.
- Saunders W. 1999. Preparation of DEMs for use in environmental modeling analysis, in: *ESRI User Conference*. pp. 24-30.
- Tarboton DG. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2), 309-319.

Appendix G. SAR (L-band) for Soil Moisture

Introduction

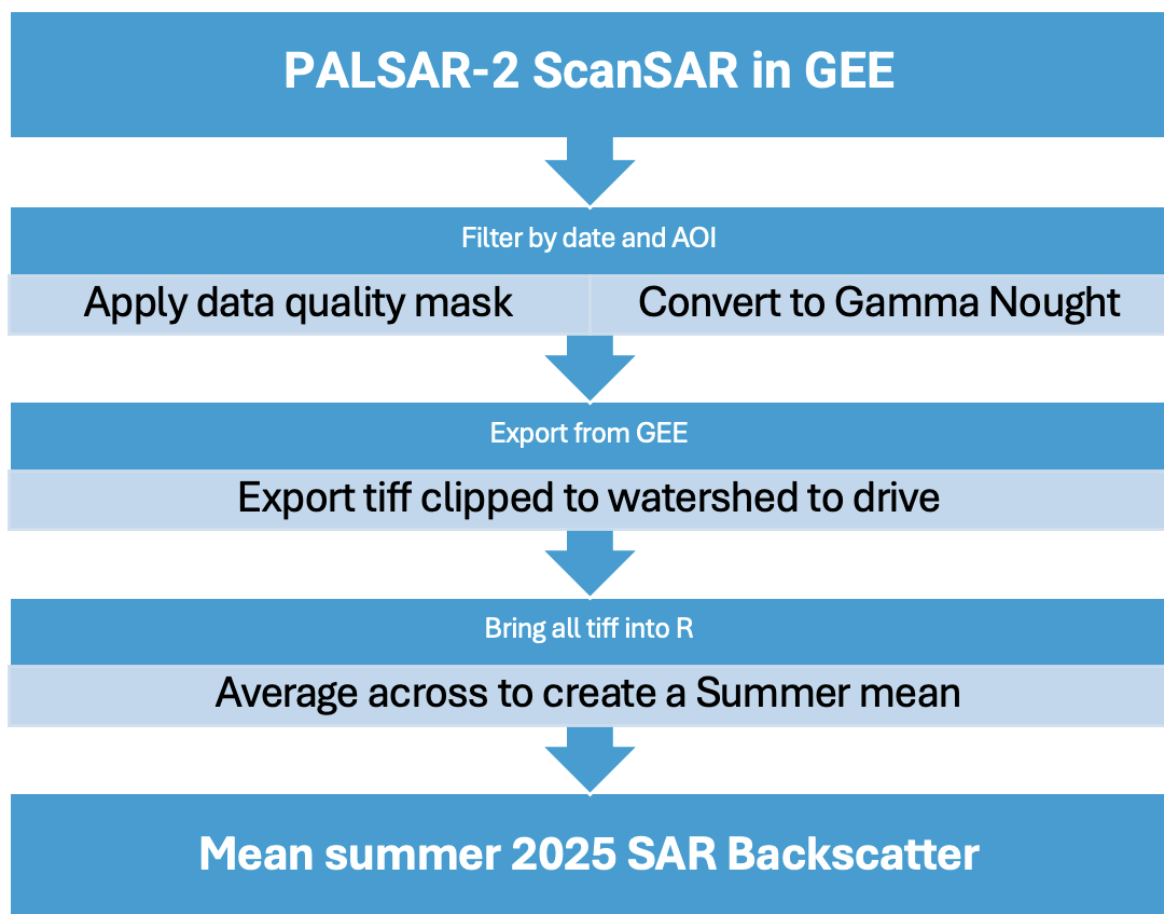
Radar remote sensing has great promise for soil moisture mapping and has been widely adopted. NASA's soil moisture active passive (SMAP) is one of the most widely used global soil moisture products. It was launched in 2015 and equipped with an L-band Synthetic Aperture Radar (SAR) and a radiometer. Unfortunately, the SAR sensor failed within the first year of the mission, leaving only the L-band radiometer (Chan et al., 2015). The radiometer remains a valuable soil moisture sensor and has been used to generate global soil moisture maps and support hydrological modeling (Chan et al., 2015; ARSET, 2024; Nayak et al., 2025). Although these data are at coarse scales on the order of kilometers rather than meters, SMAP data range from 3 to 36 km with a temporal resolution of 1-2 days since 2015 (Peng et al., 2021). These resolutions are effective for larger sites, such as global or even state-level observation, but are less effective for more local-scale monitoring.

Synthetic Aperture Radar (SAR) has great potential for soil moisture monitoring, with spatial resolutions of 2–100 meters (Mohanty et al., 2017). SAR data is a form of spaceborne active remote sensing; SAR is analogous to ultrasound and operates in the longer portion (>3.8 cm) of the electromagnetic spectrum (Aukema et al., 2019; ARSET, 2024). However, these products require a higher level of processing and expertise to interpret, as both spectral and polarimetric resolutions heavily influence their application. Not only do the radar parameters such as wavelength, polarization, and backscatter coefficients impact these data's ability to inform soil moisture monitoring, but also the characteristics of the observed surface. In densely forested landscapes or areas with complex topography, SAR can be less effective, underscoring the need for further research in these environments (Bonnell et al., 2024).

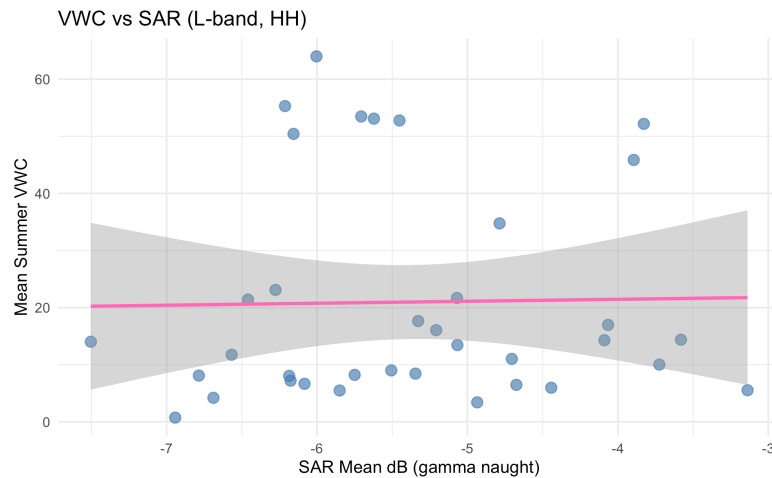
Methods

For the purposes of this work, we will be utilizing L-band data from JAXA's ALOS PALSAR-2. Higher resolution Level 2.2 scans are available in GEE and are ideal for this project, given their coverage of the study area, level of processing, and wavelength. There are 4 images covering the Skykomish during the summer of 2025, all images are in HH polarization, are descending, and at 25-meter resolution. The data is stored as digital numbers (DN) and requires a calibration equation to convert to backscatter values. GEE offers the following equation to convert DNs to Gamma nought ($\gamma^0 = 10 \times \log_{10}(\text{DN}^2) - 83.0 \text{ dB}$). This backscatter coefficient is ideal given that Gamma nought is more effective in complex terrain. The Gamma nought backscatter coefficient uses a different reference area in the equation than sigma nought. Sigma nought uses the ground range area ("locally tangent to an ellipsoidal model of the ground surface"), whereas Gamma nought uses the illuminated area accounting for the local incidence angle (Small, 2011).

SAR data processing workflow:



SAR and VWC relationship:



A linear model testing the relationship between mean summer soil moisture and mean SAR gamma naught backscatter revealed no evidence of a correlation between L-band SAR backscatter and VWC in the study area ($R^2 = 0.0003787$, $P = 0.9103$). These results highlight the limitations of SAR, even L-band SAR, in densely forested mountainous terrain.

References:

- ARSET - An Introduction to Synthetic Aperture Radar (SAR) and its Applications. NASA Applied Remote Sensing Training Program (ARSET).
<https://appliedsciences.nasa.gov/get-involved/training/english/arset-introduction-synthetic-aperture-radar-sar-and-its-applications>
- Aukema J, Wilson S, Thapa R, and Cherrington E. Synthetic Aperture Radar (SAR) Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation. 2019.
<https://doi.org/10.25966/NR2C-S697>
- Bonnell R, Elder K, McGrath D, et al. "L-Band InSAR Snow Water Equivalent Retrieval Uncertainty Increases With Forest Cover Fraction." *Geophysical Research Letters* 51, no. 24 (2024): e2024GL111708. <https://doi.org/10.1029/2024GL111708>
- Chan S, Bindlish R, O'Neill P, Jackson T, Njoku E, Dunbar S, Chaubell J, Piepmeier J, Yueh S, Entekhabi D. 2018. Development and assessment of the SMAP enhanced passive soil moisture product. *Remote Sensing of Environment* 204: 931–941.
<https://doi.org/10.1016/j.rse.2017.08.025>
- Mohanty BP, Cosh MH, Lakshmi V, and Montzka C. "Soil Moisture Remote Sensing: State-of-the-Science." *Vadose Zone Journal* 16, no. 1 (2017): vzj2016.10.0105.
<https://doi.org/10.2136/vzj2016.10.0105>
- Nayak AK, Xu X, Frey SK, et al. "Quantifying the Potential of Using Soil Moisture Active Passive (SMAP) Soil Moisture Variability to Predict Subsurface Water Dynamics." *Hydrology and Earth System Sciences* 29, no. 1 (2025): 215–44.
<https://doi.org/10.5194/hess-29-215-2025>

- Peng J, Albergel C, Balenzano A, et al. "A Roadmap for High-Resolution Satellite Soil Moisture Applications – Confronting Product Characteristics with User Requirements." Remote Sensing of Environment 252 (January 2021): 112162.
<https://doi.org/10.1016/j.rse.2020.112162>
- Small D. "Flattening Gamma: Radiometric Terrain Correction for SAR Imagery." IEEE Transactions on Geoscience and Remote Sensing 49, no. 8 (2011): 3081–93.
<https://doi.org/10.1109/TGRS.2011.2120616>