



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Water repellency prediction in high-organic Greenlandic soils

Comparing vis–NIRS to pedotransfer functions

Blaesbjerg, Natasha H.; Weber, Peter L.; de Jonge, Lis Wollesen; Moldrup, Per; Greve, Mogens H.; Arthur, Emmanuel; Knadel, Maria; Hermansen, Cecilie

Published in:
Soil Science Society of America Journal

DOI (link to publication from Publisher):
[10.1002/saj2.20407](https://doi.org/10.1002/saj2.20407)

Creative Commons License
CC BY 4.0

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Blaesbjerg, N. H., Weber, P. L., de Jonge, L. W., Moldrup, P., Greve, M. H., Arthur, E., Knadel, M., & Hermansen, C. (2022). Water repellency prediction in high-organic Greenlandic soils: Comparing vis–NIRS to pedotransfer functions. *Soil Science Society of America Journal*, 86(3), 643-657.
<https://doi.org/10.1002/saj2.20407>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

ASA, CSSA, and SSSA Virtual Issue Call for Papers: Advancing Resilient Agricultural Systems: Adapting to and Mitigating Climate Change

Content will focus on resilience to climate change in agricultural systems, exploring the latest research investigating strategies to adapt to and mitigate climate change. Innovation and imagination backed by good science, as well as diverse voices and perspectives are encouraged. Where are we now and how can we address those challenges? Abstracts must reflect original research, reviews and analyses, datasets, or issues and perspectives related to objectives in the topics below. Authors are expected to review papers in their subject area that are submitted to this virtual issue.

Topic Areas

- Emissions and Sequestration
 - » Strategies for reducing greenhouse gas emissions, sequestering carbon
- Water Management
 - » Evaporation, transpiration, and surface energy balance
- Cropping Systems Modeling
 - » Prediction of climate change impacts
 - » Physiological changes
- Soil Sustainability
 - » Threats to soil sustainability (salinization, contamination, degradation, etc.)
 - » Strategies for preventing erosion
- Strategies for Water and Nutrient Management
 - » Improved cropping systems
- Plant and Animal Stress
 - » Protecting germplasm and crop wild relatives
 - » Breeding for climate adaptations
 - » Increasing resilience
- Waste Management
 - » Reducing or repurposing waste
- Other
 - » Agroforestry
 - » Perennial crops
 - » Specialty crops
 - » Wetlands and forest soils



Deadlines

Abstract/Proposal Deadline: Ongoing
Submission deadline: 31 Dec. 2022








How to submit

Submit your proposal to
manuscripts@sciencesocieties.org

Please contact Jerry Hatfield at
jerryhatfield67@gmail.com with any questions.



Water repellency prediction in high-organic Greenlandic soils: Comparing vis–NIRS to pedotransfer functions

Natasha H. Blaesbjerg¹  | Peter L. Weber²  | Lis Wollesen de Jonge²  |
Per Moldrup¹ | Mogens H. Greve²  | Emmanuel Arthur²  | Maria Knadel²  |
Cecilie Hermansen² 

¹Dep. of the Built Environment, Faculty of Engineering and Science, Aalborg Univ., Thomas Manns Vej 23, Aalborg DK-9220, Denmark

²Dep. of Agroecology, Faculty of Technical Sciences, Aarhus Univ., Blichers Allé 20, Tjele DK-8830, Denmark

Correspondence

Cecilie Hermansen, Dep. of Agroecology, Faculty of Technical Sciences, Aarhus Univ., Blichers Allé 20, Tjele, DK-8830, Denmark.
Email: cecilie.hermansen@agro.au.dk

Assigned to Associate Editor Aaron Lee Motschenbacher Daigh.

Abstract

Soil water repellency (SWR) is a common phenomenon across agricultural soils of South Greenland that can negatively affect soil functions. Existing methods to measure SWR as a function of water content (w) are laborious. This study was conducted to compare the potential of visible–near-infrared spectroscopy (vis–NIRS) as an alternative method to pedotransfer functions (PTFs) for predicting four SWR indices in 143 agricultural soils from South Greenland (clay, 0.016–0.172 kg kg⁻¹; organic carbon (OC), 0.009–0.241 kg kg⁻¹). Pedotransfer functions were established by multiple linear regression based on OC, clay, and pH. Partial least squares regression (PLS-R) and interval PLS-R were applied to build vis–NIRS prediction models (vis–NIR range 400–2,500 nm). The area under the SWR– w curve (SWR_{area}) and the critical soil water content (w_{non}) were accurately predicted by PTFs ($R^2 = .90$; $R^2_{\text{adj}} = .91$) while the SWR after 60 °C pretreatment (SWR_{60}), and the integrative repellency dynamic index (IRDI) were predicted less accurately ($R^2_{\text{adj}} = .36$; $R^2_{\text{adj}} = .27$). Vis-NIRS models with variable selection performed at a better or close to the same level of accuracy as PTFs (SWR_{area} , $R^2 = .88$; w_{non} , $R^2 = .90$; SWR_{60} , $R^2 = .63$; IRDI, $R^2 = .54$). This study demonstrated vis–NIRS as a valuable alternative to PTFs for rapid assessment of SWR and as a tool for SWR mitigation for farmers in South Greenland. The results may well apply to other regions with similar texture and OC ranges, but further testing is required.

Abbreviations: iPLS-R, interval partial least squares regression; IRDI, integrative repellency dynamic index; LOI₅₅₀, loss of ignition at 550 °C; MED, molarity of an ethanol droplet; MLR, multiple linear regression; OC, organic carbon; PLS-R, partial least squares regression; PTF, pedotransfer function; R^2_{adj} , R^2 adjusted for number of independent variables; RMSE_{CV}, RMSE of the cross-validation; SRMSE, standardized root mean square error; SWR, soil water repellency; SWR_{60} , soil water repellency after 60 °C pretreatment; SWR_{area} , area under the SWR– w curve; vis–NIRS, visible–near-infrared spectroscopy; w , water content; w_{non} , critical soil water content.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Soil Science Society of America Journal* published by Wiley Periodicals LLC on behalf of Soil Science Society of America.

1 | INTRODUCTION

Agricultural soils in South Greenland exhibit a high prevalence in soil water repellency (SWR); 99 and 98% of the investigated soils in the study by Weber et al. (2021) were water repellent and extremely water repellent, respectively. Comparably, other studies across different scales, as well as climatic and geographic regions, found the prevalence of SWR to exceed 90% (Davari et al., 2022; Deurer et al., 2011; Hermansen et al., 2019b; Seaton et al., 2019).

Soil water repellency is a phenomenon that exists for some soils within a soil-specific range in water content (w) (de Jonge et al., 1999, 2007; Dekker & Ritsema, 1995; Dekker et al., 2001). Above a certain threshold in w (w_{non}), the soil remains hydrophilic regardless of further increases in moisture content. Below w_{non} , the severity in SWR varies nonlinearly with w , and the shape of the SWR vs. w (SWR- w) curve can either be unimodal or bimodal (de Jonge et al., 1999; Regalado et al., 2008). Soil water repellency affects soil functional properties and has been associated with increased overland flow, increased soil erosion, decreased infiltration, decreased plant yield, finger flow or preferential flow, enhanced nutrient loss, and enhanced pesticide leaching (Dekker & Ritsema, 1996a; b; Blackwell, 2000; Leighton-Boyce et al., 2007; de Jonge et al., 2009; Müller et al., 2018). As warmer temperatures are projected for South Greenland (Hanna et al., 2021), a subsequent increase in agricultural land use must be expected (Masson-Delmotte et al., 2012). This will require a better understanding of the physical parameters affecting the severe SWR in this region. Knowledge of both w_{non} and the severity in SWR across a gradient in soil water content (SWR_{area}) will be useful for SWR remediation purposes. The soil property with the greatest influence on SWR_{area} and w_{non} is the soil organic carbon (OC) content. Several studies have reported a high and often linear correlation between OC content and SWR_{area} and w_{non} (Kawamoto et al., 2007; Regalado et al., 2008; Hermansen et al., 2019b; Weber et al., 2021). Organic matter acting as a hydrophobic coating on soil particles is generally considered as the main cause of SWR in soils (Doerr et al., 2000), and it can be formed either from waxes produced by plants or by microbial activity (Giovannini et al., 1983).

There are several methods for assessing SWR. The most widely used methods are the molarity of an ethanol droplet (MED) test (Watson & Letey, 1970; King, 1981; Roy & McGill, 2002), the water drop penetration time (Letey et al., 2000), and the sessile drop method (Bachmann et al., 2000). While the degree of SWR is measured with the MED test, the persistence of SWR is measured with the water drop penetration time method. Further, the sessile drop method measures the contact angle between a water droplet and the soil. It is extremely time consuming and laborious to measure the full SWR- w curve because it requires SWR to

Core Ideas

- Soil water repellency (SWR) vs. w curves were measured for 143 South Greenlandic soil samples.
- The area under the SWR- w curve and the critical w were determined.
- PTFs based on basic soil properties predicted SWR_{area} and w_{non} accurately.
- Rapid spectral measurements (vis-NIRS) predicted SWR indices with similar accuracy.
- Variable selection within the vis-NIR spectrum (iPLS-R) improved vis-NIRS model performance.

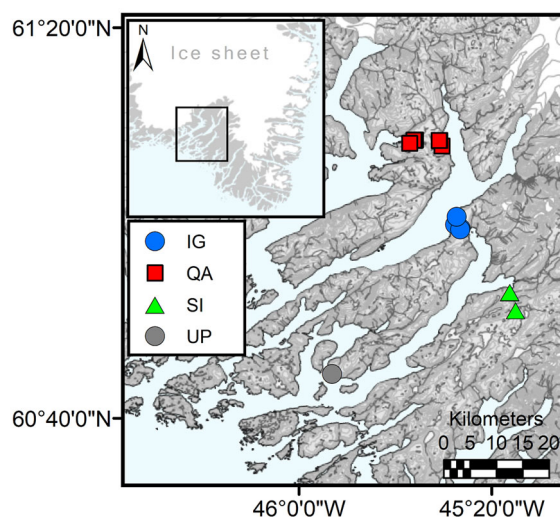


FIGURE 1 Map over the study area in South Greenland with the four field locations marked: Igaliku (IG), Qassiarsuk (QA), Sondre Igaliku (SI), and Upernaviarsuk (UP)

be measured across a range in w from completely dry to moist conditions. As an example, it requires a minimum of one full working day to measure a full SWR- w curve using the MED test when accumulating the time required to perform all steps involved. It would thus be highly beneficial with alternative and less time-consuming methods to assess SWR.

Throughout the literature, pedotransfer functions (PTFs) for SWR indices based on basic soil properties have been developed. For example, Weber et al. (2021) used clay and OC to predict w_{non} for Greenlandic soils, while Hermansen et al. (2019b) estimated SWR_{area} using OC and pH as predictors. Further, Karunarathna et al. (2010) developed a two-region model that related the degree of SWR to matric potential and, by using the van Genuchten soil water retention model, to water content. However, this method does not eliminate the time-consuming laboratory work required to obtain soil water retention data. Another method that has

proven successful for rapid (<2 min) estimation of a wide range of soil properties is visible–near-infrared spectroscopy (vis–NIRS) (Chang et al., 2001). Organic carbon, texture, and a wide range of soil functional properties can be predicted with this method (Ben-Dor & Banin, 1995; Stenberg et al., 2010; Katuwal et al., 2018). One of the advantages of this method is that several basic and functional soil properties can be derived from the same spectrum. Its broad applicability for soil property estimation can be attributed to overtone and combination bands of fundamental vibrations in the mid-infrared region originating from different molecular bonds (i.e., O–H, metal–OH, C–H, C–O, and N–H) and electronic transitions of iron oxides (Hunt, 1977; Ben-Dor, 2002; Stenberg et al., 2010). Some studies have already predicted SWR using vis–NIRS. The degree of SWR after heat pretreatments of 60 and 105 °C was predicted well for soils from a coarse-textured Danish agricultural field (OC, 0.014–0.025 kg kg⁻¹) (Knadel et al., 2016). Another study predicted both the degree and persistence of SWR after heat pretreatment at 65 °C on soils from the North Island of New Zealand (OC, 0.033–0.123 kg kg⁻¹) (Kim et al., 2014). Further, a study used partial least squares regression (PLS–R) with vis–NIRS to predict the SWR index (measured with the relative sorptivity method) and obtained acceptable results ($R^2 > .52$ and ratio of performance to interquartile distance > 2.27) for 100 soil samples from Iran (soil organic matter, 0.007–0.136 kg kg⁻¹) (Davari et al., 2022). Lastly, the SWR– w curve, as well as SWR_{area}, were predicted using vis–NIRS by Hermansen et al. (2019a) on soil samples from the South Island of New Zealand (OC, 0.021–0.147 kg kg⁻¹) using an approach where fitting parameters (maximum SWR, soil water content at maximum SWR and w_{non}) of a three-parameter moisture-dependent SWR model were predicted. The ratio between SWR_{area} and w_{non} , called integrative repellency dynamic index (IRDI), is a measure of the average SWR proposed by Regalado and Ritter (2005). The IRDI has not yet been predicted using vis–NIRS in any published work known to the authors. Thus, the limited literature on this topic indicates that some SWR indices can be predicted using vis–NIRS although the prediction accuracy varies. The arctic soils in South Greenland differ significantly from the soils used in previous datasets to develop vis–NIRS prediction models for SWR indices throughout literature. Comparatively, the soils in South Greenland are farmed across a range of altitudes, and the environment is unique with steep mountain slopes, fjords, and proximity (<60 km) to the inland ice sheet. These factors can affect the microclimatic conditions directly or indirectly, and further affect both the quantity and quality of organic matter present in these soils. Given the high prevalence of SWR in South Greenland, estimating key SWR indices using rapid (<2 min) vis–NIRS analysis to replace time-consuming laboratory work has the potential to ease the process of SWR remediation. For example, for the purpose of using irrigation

as an SWR amelioration technique (Wallis & Horne, 1992), knowledge of w_{non} is beneficial.

Thus, based on 143 soil samples, the objectives of this study were to assess the potential of vis–NIRS as an alternative method to PTFs for estimating SWR indices in high-organic agricultural soils in South Greenland by (a) developing PTFs for the four SWR indices: SWR₆₀, SWR_{area}, w_{non} , and IRDI based on basic soil properties (OC, clay, and pH); (b) developing full-curve and interval PLS–R vis–NIRS prediction models for the same four SWR indices and; (c) comparing model performances of vis–NIRS to PTFs. Further, the spectral intervals selected in the variable selection will be analyzed to gain knowledge of the spectral intervals and soil constituents involved in the prediction of SWR.

2 | MATERIALS AND METHODS

2.1 | Study area and soil sampling

A total of 143 samples were collected from four different areas comprising 16 fields in South Greenland as part of sampling campaigns during the summers of 2015, 2017, and 2018. Data from 127 samples originated from the paper by Weber et al. (2021), while SWR data from the remaining 16 samples were new in this study. The four areas—Qassarsuk (61°09' N, 45°30' W), Igaliku (61°00' N, 45°26' W), Søndre Igaliku (60°53' N, 45°16' W), and Upernavarsuk (60°44'57.3'' N, 45°53'24.4'' W) (Figure 1) are all part of the main agricultural region in South Greenland where sheep farming and pastures are most prevalent (Westergaard-Nielsen et al., 2015). This region is permafrost-free and stretches ~60 km from the inland ice sheet to the open ocean and is cut by several deep fjords. The climatic zone varies from oceanic in the areas closest to open seas to subcontinental closer to the inland ice sheet (Jacobsen, 1987). Mean annual temperatures ranged between –3 and 4.8 °C close to the ice sheet (Narsarsuaq) and between –2.9 and 4.0 °C close to the open ocean (Qarqortoq) in the period from 1961 to 1990, while mean annual precipitation in the same period ranged between 615.1 and 857.6 mm for the two measurement stations, respectively (Hanna & Cappelen, 2002).

The soils were sampled at ~15 cm depth in fields predominantly grown with perennial grass mixtures. A detailed description of soil sample collection and field conditions can be found in Weber et al. (2021).

2.2 | Soil water repellency

The MED test (King, 1981; de Jonge et al., 1999; Hermansen et al., 2019b) was used in this study to assess the degree of SWR, and the protocol described in Hermansen et al.

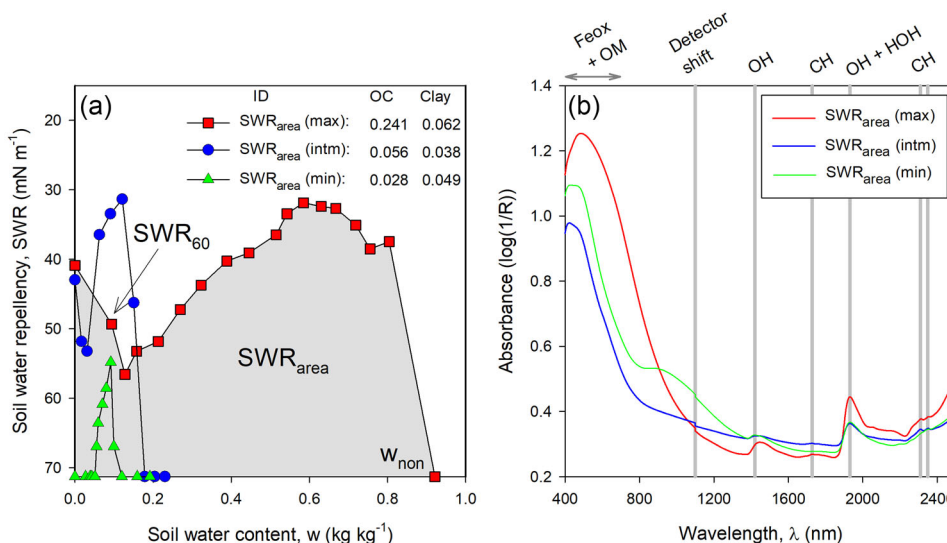


FIGURE 2 (a) Soil water repellency (SWR) as a function of soil water content for three samples representing the minimum, intermediate, and maximum trapezoidal integrated area underneath the SWR vs. soil water content curve (SWR_{area}). The maximum SWR_{area} is highlighted, as are the indices: SWR determined after heat pretreatment at 60 °C (SWR₆₀) and critical soil water content (w_{non}). Note the y axis is inverted. (b) Visible near-infrared spectra of the samples representing the minimum, intermediate, and maximum SWR_{area} and the corresponding spectrally active components

(2019b) was carefully followed. In short, droplets (60 μl) of varying ethanol solution concentrations (0–50 m³ m⁻³, 0.01 m³ m⁻³ steps) were pipetted onto a smoothed soil surface. The highest ethanol solution concentration that stayed on the surface >5 s corresponded to the degree of SWR.

Soil water repellency was measured across a gradient in soil water contents—from completely dry conditions until w_{non}, where water repellency ceased. Water contents of 0 kg kg⁻¹ were obtained by oven drying the soils at 105 °C for 24 h. After oven drying at 105 °C, the soils were kept in a desiccator until room temperature was reached. Water contents between 0 kg kg⁻¹ and air-dry conditions were obtained by first oven-drying the soils at 60 °C, after which the soils were equilibrated for 48 h at climate-controlled laboratory conditions (20 °C). To obtain water contents above air-dry conditions, increasing amounts of water were pipetted onto the soils, after which the samples were thoroughly mixed and refrigerated in zip-lock bags for at least 2 wk to ensure even distribution of water before SWR measurement. The soil water content was then measured on a subsample of the moist sample by oven drying at 105 °C, while MED measurement was performed on a second subsample. Using the equation provided by Roy and McGill (2002), ethanol solution concentration was converted to the unit of surface tension (mN m⁻¹). The four indices used throughout this paper all characterize the SWR–w curve (Figure 2a): SWR₆₀, SWR_{area}, w_{non}, and IRDI. The y axis in the SWR–w plot (Figure 2a) is inverted and crosses the x axis at 71.27 mN m⁻¹, which is the surface tension of water and the point at which the soil will become hydrophilic.

2.3 | Development of pedotransfer functions

Forward multiple linear regression was performed to establish PTFs describing the SWR indices (SWR₆₀, SWR_{area}, w_{non}, and IRDI) using the soil OC, clay content, and pH. All soil variables that contributed significantly ($p < .05$) were used in the expressions. The prediction accuracy was evaluated using the coefficient of determination adjusted for the number of independent variables used in the expression (R^2_{adj}), the root mean square error (RMSE), and the standardized RMSE (SRMSE) calculated as the RMSE divided by the range in the response variable [SRMSE = RMSE/(max – min)] (Arthur, 2017). The SRMSE enables model comparison across indices, and a relatively high error is reflected in a high SRMSE. In this paper, the PTFs for SWR_{area} and w_{non} with OC as predictor variable were comparable to PTFs originally presented in the study by Weber et al. (2021). However, these PTFs were updated to represent the specific dataset of this study, such that model performances of PTFs and vis–NIRS models could be directly compared. The remaining PTFs presented throughout this paper were specifically developed for this study.

2.4 | Visible near-infrared spectroscopy measurements

A benchtop DS2500 vis–NIR spectrometer (FOSS) was used to measure diffuse reflectance (R) within the visible (400–780 nm) and near-infrared (780–2,500 nm) regions

with a sampling interval of 0.5 nm. The protocol described in Hermansen et al. (2016) was closely followed. During vis-NIRS measurements, the laboratory temperature ranged from 23 to 24 °C and the relative humidity of the room ranged from 46 to 53%. Prior to measuring, the instrument was tested for stability with a white reference. A representative part (~50 g) of each sample (air dried and 2 mm sieved) was placed in a measuring cup fitted with a quartz window. The cup rotated while the instrument took measurements at seven different points of the sample surface facing the quartz window. The measurements were then averaged into one representative diffuse reflectance spectra and converted to absorbance (A) using the relation $A = \log(1/R)$.

2.5 | Multivariate data analysis

Multivariate data analysis was performed using the PLS Toolbox (Eigenvector Research Inc.) software. Calibration models were established by correlating the predictor variables (spectral data) with reference data for each response variable (OC, clay, SWR_{area} , w_{non} , SWR_{60} , and IRDI). Two approaches were used in this process: partial least squares regression (PLS-R) and interval PLS-R (iPLS-R), both using the SIMPLS algorithm (de Jong, 1993).

Partial least squares regression is a widely used method within chemometrics that relates two sets of data—predictor variables and response variables—by regression. The resulting calibration model can then be used to predict values of the response variables for new samples (Esbensen, 1994; Wold et al., 2001). It is a data compression technique where the spectral data gets reduced to a number of factors explaining the maximum covariance between the predictor and the response variable. Using too many factors will result in a very complex calibration model also explaining the variation caused by noise. The optimal number of factors was determined by inspecting the plot of the RMSE vs. the number of factors and selecting the number of factors where the lowest RMSE could be achieved without the distance between the RMSE of calibration and cross-validation increased.

Visible-near-infrared spectroscopy models for SWR indices, OC, and clay were established on the full vis-NIR range, hereafter referred to as ‘full-curve’, as well as index-specific spectral regions. An automated function in PLS Toolbox designed to perform forward iPLS-R was used for variable selection to identify important spectral regions and decrease the complexity of the models (Zou et al., 2010). The processes involved in this tool can be divided into several steps. In the first step, the spectra are divided into equally sized intervals, and a calibration model is developed based on each interval. Next, the interval with the lowest RMSE is selected, and the algorithm run again while only adding intervals that decrease the total RMSE value. Finally, the

output is a selection of intervals on which a calibration model is developed to predict new values of the variable in question. A separate iPLS-R variable selection was performed for each of the soil indices (OC, clay, SWR_{60} , SWR_{area} , w_{non} , and IRDI). The variable selection was tested with different interval sizes: 60, 80, 100, and 120, which corresponded to widths of 30, 40, 50, and 60 nm because the measurements were recorded every 0.5 nm.

Smoothing and derivatives of spectral data can be applied to remove additive and multiplicative effects and correct for baseline effects caused by nonchemical effects (Rinnan et al., 2009). Thus, multiple different pretreatment techniques were systematically tested on the full-curve spectral data for each calibration model. The tested pretreatments included the following: Savitzky–Golay smoothing, first and second derivatives (Savitzky & Golay, 1964), and Gap Segment smoothing, first and second derivatives (Norris, 2001). Based on this, the two best-performing pretreatments were used for iPLS-R, and the pretreatment that collectively performed best for both PLS-R and iPLS-R was chosen. Additionally, the predictor variables and the response variables were all mean centered, and both approaches were tested with mean centering only. The RMSE of the cross-validation ($RMSE_{CV}$), coefficient of determination (R^2), and SRMSE were used to assess the performance of each model. Generally, the best performing model in terms of RMSE and R^2 with the lowest number of factors was chosen.

Full cross-validation (leave-one-out) was carried out for models based on full-curve and specific spectral intervals as well. The full cross-validation served several functions, specifically, to assess the optimal complexity of the model (number of factors) and to attain better comparability in the results with multiple linear regression (MLR) PTFs.

3 | RESULTS AND DISCUSSION

3.1 | Basic soil properties

The 143 soil samples ranged from 0.016 to 0.172 kg kg⁻¹ in clay content with a mean of 0.053 kg kg⁻¹ and the OC content spanned widely from 0.009 to 0.241 kg kg⁻¹ (Table 1). The soil samples covered four USDA soil textural classes (sand, loamy sand, sandy loam, and loam) as described in the Weber et al. (2021) study that reported on 127 of the samples also included in this study.

The relationship between clay content and OC content is plotted in Figure 3a. The distribution of the samples relative to the Dexter n (clay/OC = 10) line showed an overweight of OC to clay content. A clay/OC ratio of 10 has been suggested as the upper limit for clay complexation with OC (Dexter et al., 2008), indicating that the OC in the Greenland soils was present in both complexed and noncomplexed form.

TABLE 1 Descriptive statistics ($n = 143$) for soil texture, organic carbon (OC), loss-on-ignition at 550 °C (LOI_{550}), pH, and soil water repellency (SWR) indices

Property	Min	Max	Mean	Median	Q_1	Q_3
Clay, kg kg^{-1}	0.016	0.172	0.053	0.047	0.038	0.063
Silt, kg kg^{-1}	0.068	0.435	0.278	0.297	0.223	0.323
Sand, kg kg^{-1}	0.237	0.876	0.543	0.527	0.486	0.606
OC, kg kg^{-1}	0.009	0.241	0.063	0.059	0.039	0.075
LOI_{550} , kg kg^{-1}	0.023	0.649	0.121	0.109	0.082	0.132
pH	4.5	7.5	5.2	5.1	4.9	5.5
SWR_{60} , mN m^{-1}	35.54	71.27	56.59	54.80	49.34	63.58
SWR_{area} , $\text{mN m}^{-1} \text{kg kg}^{-1}$	0.54	25.89	6.17	5.04	3.82	7.23
w_{non} , kg kg^{-1}	0.052	0.921	0.227	0.197	0.159	0.254
IRDI	4.52	36.25	26.17	26.65	23.66	29.18

Note. Min, minimum; Max, maximum; Q_1 , first quartile of data set; Q_3 , third quartile of data set; SWR_{60} , SWR after 60 °C pretreatment; SWR_{area} , the area under SWR– w curve; w_{non} , critical soil water content; IRDI, the ratio between SWR_{area} and w_{non} .

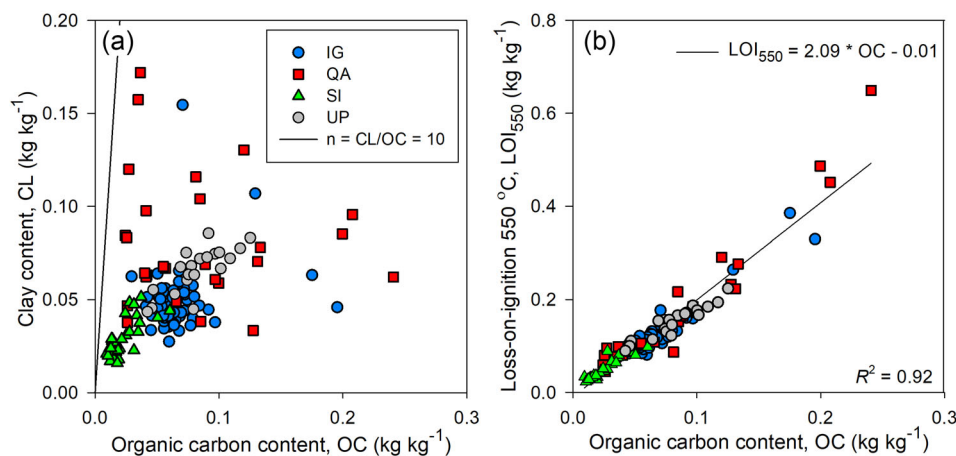


FIGURE 3 (a) The clay content (CL) with Dexter n ($n = \text{CL}/\text{OC} = 10$) and (b) loss-on-ignition at 550 °C (LOI_{550}) as a function of the organic carbon content (OC)

The ratio between loss of ignition at 550 °C (LOI_{550}) and OC was 2.09 for the sampled soils (Figure 3b). This aligns well with Pribyl (2010) who argued that using a conversion factor between $\text{OC}/\text{SOM} = 2$, corresponding to a 50% OC content in SOM, is accurate in most cases.

3.2 | Soil water repellency curves

All 143 samples used in this study exhibited water repellency (Figure 2a). The majority of the measured SWR– w curves were illustrated by Weber et al. (2021).

The SWR– w curves of the soils with the minimum, intermediate, and maximum SWR_{area} are plotted in Figure 2a to show the range in different curve shapes and sizes. The index SWR_{area} spanned widely from 0.54 to 25.89 mN m^{-1}

kg kg^{-1} (Table 1). The curve exhibiting the lowest SWR_{area} was characterized by a unimodal curve shape, whereas the curves exhibiting the intermediate and highest SWR_{area} were characterized by a bimodal curve shape (de Jonge et al., 1999; Regalado et al., 2008). A high SWR_{area} can be attributed to a high OC content (de Jonge et al., 1999; Regalado et al., 2008; Hermansen et al., 2019b; Weber et al., 2021). This was also the case for the sample with the maximum SWR_{area} , which had the highest OC content, while the opposite was true for the sample with the minimum SWR_{area} . Clay content was highest in the sample with the maximum SWR_{area} , however, the sample with the lowest clay content did not exhibit the minimum SWR_{area} , demonstrating the lack of a clear influence of clay on soil water repellency.

The critical water content also varied considerably ranging from 0.052 to 0.921 kg kg^{-1} (Table 1). For 71 New Zealand

TABLE 2 Pedotransfer functions using linear regression (LR) and multiple linear regression (MLR) for soil water repellency (SWR) indices—SWR after 60 °C pretreatment (SWR_{60}), the area under SWR– w curve (SWR_{area}), critical soil water content (w_{non}) and the ratio between SWR_{area} and w_{non} (IRDI)—based on organic carbon (OC) in combination with clay content (CL) and pH for 143 samples

Property	Regression	Equation	r	R^2	R^2_{adj}	RMSE	SRMSE
SWR_{60}	LR	$-106.82 \text{ OC} + 63.29$	0.47	.22	.21	7.82	0.22
	MLR	$-128.59 \text{ OC} + 87.46 \text{ CL} + 5.39 \text{ pH} + 31.72$	0.61	.38	.36	6.98	0.20
SWR_{area}	LR	$106.14 \text{ OC} - 0.49^a$	0.95	.90	.90	1.36	0.05
w_{non}	LR	$3.30 \text{ OC} + 0.02^a$	0.94	.88	.88	0.05	0.06
	MLR	$3.10 \text{ OC} + 0.82 \text{ CL} + 0.025 \text{ pH} - 0.14$	0.96	.92	.91	0.04	0.05
IRDI	LR	$51.88 \text{ OC} + 22.92$	0.44	.19	.18	4.17	0.13
	MLR	$66.30 \text{ OC} - 58.05 \text{ CL} + 25.11$	0.53	.28	.27	3.93	0.12

Note. r , coefficient of determination; R^2_{adj} , coefficient of determination adjusted for number of variables; SRMSE, standardized RMSE.

^aModified from Weber et al. (2021).

soil samples, w_{non} was found to range between 0.050 and 0.438 kg kg⁻¹ and SWR_{area} to range between 0.12 and 11.57 mN m⁻¹ kg kg⁻¹ (Hermansen et al., 2019a). Similarly, de Jonge et al. (2007) found values of w_{non} from 0.03 to 0.095 kg kg⁻¹ but much lower values of SWR_{area} (0–2.9 mN m⁻¹ kg kg⁻¹) on Danish soils with different crop types. Thus comparatively, the subarctic soils in this study had a relatively high severity in SWR_{area} and w_{non} compared with the New Zealand soils, while the severity was much more evident when comparing SWR_{area} with Danish soils. The critical water content found in this study, however, was quite similar to the Danish soils.

3.3 | Performance of pedotransfer functions

The OC content contributed significantly to explaining the variance in all four SWR indices (SWR_{60} , SWR_{area} , w_{non} , and IRDI; $p < .001$). However, an improvement in prediction accuracy was achieved when OC was used in combination with clay for IRDI and clay and pH for SWR_{60} and w_{non} , thereby establishing PTFs using multiple linear regression (Table 2). The SWR_{area} was best described using the linear regression model based on OC alone with an $R^2 = .90$ and RMSE = 1.36 mN m⁻¹ kg kg⁻¹, because neither clay nor pH were significant ($p > .05$).

Figure 4a–d shows the best-performing PTFs established for the four SWR indices. A strong positive correlation was demonstrated between SWR_{area} and OC ($R^2_{adj} = .90$) and w_{non} and OC ($R^2_{adj} = .91$) as was also found by Weber et al. (2021) using parts of the same data set. This is consistent with previous studies (de Jonge et al., 1999; Regalado & Ritter, 2005; Wijewardana et al., 2016; Hermansen et al., 2019b). The average SWR (IRDI) exhibited a weak correlation with OC ($R^2_{adj} = .27$), as did the SWR_{60} ($R^2_{adj} = .36$). The weak correlation between OC and SWR_{60} might indicate that SWR_{60} was more dependent on the type of OC and degree of decomposi-

tion than the amount as opposed to SWR_{area} (de Jonge et al., 1999; Hermansen et al., 2019b). The PTFs for estimating w_{non} and IRDI did not resolve the variation for the Sondre Igaliku samples well. Most likely, this was due to the low variation and range in OC and clay content relative to the other regions included in the PTFs (Figure 3a). This emphasizes the importance of including a diverse range of locations, as it will improve the versatility of the resulting PTFs.

3.4 | Visible near-infrared spectra

Figure 2b shows the unprocessed vis–NIR absorbance spectra of the samples with the maximum, intermediate, and minimum SWR_{area} . The sample with the maximum SWR_{area} and highest OC content had the highest absorbance in the visible range (400–700 nm). However, the sample with the intermediate SWR_{area} had a lower absorbance than the sample with the minimum SWR_{area} in this region. Absorption in the visible range is mainly caused by overlapping responses from iron oxides and OC (Hunt, 1977; Scheinost et al., 1998). Thus, the high absorbance in the visible range may be due to the presence of a high OC content for the sample with the maximum SWR_{area} , while the medium-high absorbance in the sample with the minimum SWR_{area} may be due to the presence of iron oxides because it had the lowest OC content. The shoulder around 800–900 nm present only in this sample was interpreted as a possible response from hematite or other iron oxides (Scheinost et al., 1998). A pronounced absorbance feature around 1,930 nm, which was present for all three samples, was associated with O–H groups adsorbed onto the particle surfaces in the mineral lattice (Ben-Dor, 2002; Stenberg et al., 2010). A lesser prominent feature around 1,420 nm, also present for all three samples, was attributed to the presence of O–H functional groups. Features around 1,730 nm and 2,310–2,350 nm may be attributed to vibrations of C–H stretches associated with possible OC components

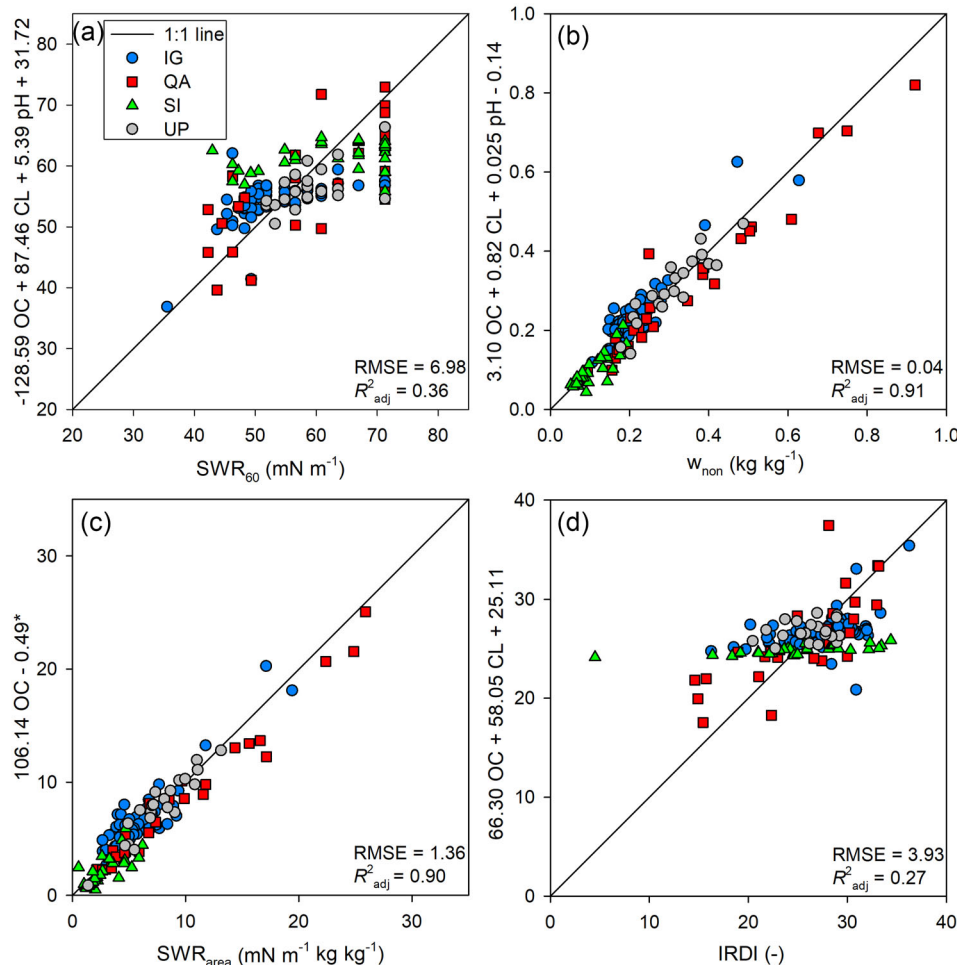


FIGURE 4 (a, b, d) Pedotransfer functions using multiple variables for soil water repellency (SWR) indices; SWR after 60 °C pretreatment (SWR_{60}), critical soil water content (w_{non}) and the ratio between the area under SWR– w curve (SWR_{area}) and w_{non} (IRDI) based on organic carbon content (OC), clay content (CL), and pH. (c) SWR_{area} is described with a single variable pedotransfer function based on OC alone because neither clay content nor pH was significant. *Function is modified from Weber et al. (2021) to include all 143 samples. Locations: IG, Igaliku; QA, Qassarsuk; SI, Sondre Igaliku; UP, Upernavisuk

such as humic acids, waxes, starches, or cellulose (Ben-Dor et al., 1997).

3.5 | Vis–NIRS prediction of soil water repellency

Partial least squares regression models were established to predict OC, clay content, and the four SWR indices (SWR_{60} , SWR_{area} , w_{non} , and IRDI) using full-curve (400–2,500 nm) vis–NIR spectral measurements. The vis–NIR spectra were pretreated using the first derivative gap-segment (gap, 9; segment, 9) for OC, SWR_{area} , and w_{non} and the second derivative gap-segment (gap, 19; segment, 25) for SWR_{60} . A first derivative Savitzky–Golay (filter width, 25) pretreatment was used for IRDI, and finally, second derivative Savitzky–Golay (filter width, 39) was used for clay content.

Table 3 shows the PLS-R vis–NIRS calibration and cross-validation results using R^2 , RMSE and SRMSE to assess model accuracy. The OC content was estimated well yielding an $R^2 = .86$ and an $RMSE_{CV} = 0.015 \text{ kg kg}^{-1}$ using nine factors. Clay was estimated with a low accuracy ($R^2 = .53$ and an $RMSE_{CV} = 0.018 \text{ kg kg}^{-1}$) using four factors with particularly one location (Qassarsuk) displaying a large scatter around the 1:1 line (Figure 5d). The estimations of SWR_{60} and IRDI were both quite poor with performance results of $R^2 = .44$ and an $RMSE_{CV} = 6.69 \text{ mN m}^{-1}$ and $R^2 = .41$ and an $RMSE_{CV} = 3.62$, respectively. The SWR_{area} and w_{non} however, were both well estimated ($R^2 = .79$ and $RMSE_{CV} = 1.99 \text{ mN m}^{-1} \text{ kg kg}^{-1}$ and $R^2 = .84$ and $RMSE_{CV} = 0.054 \text{ kg kg}^{-1}$, accordingly) both using nine factors. The model accuracy of SWR_{60} and IRDI was improved while it decreased slightly for SWR_{area} and w_{non} in comparison with results from MLR PTFs.

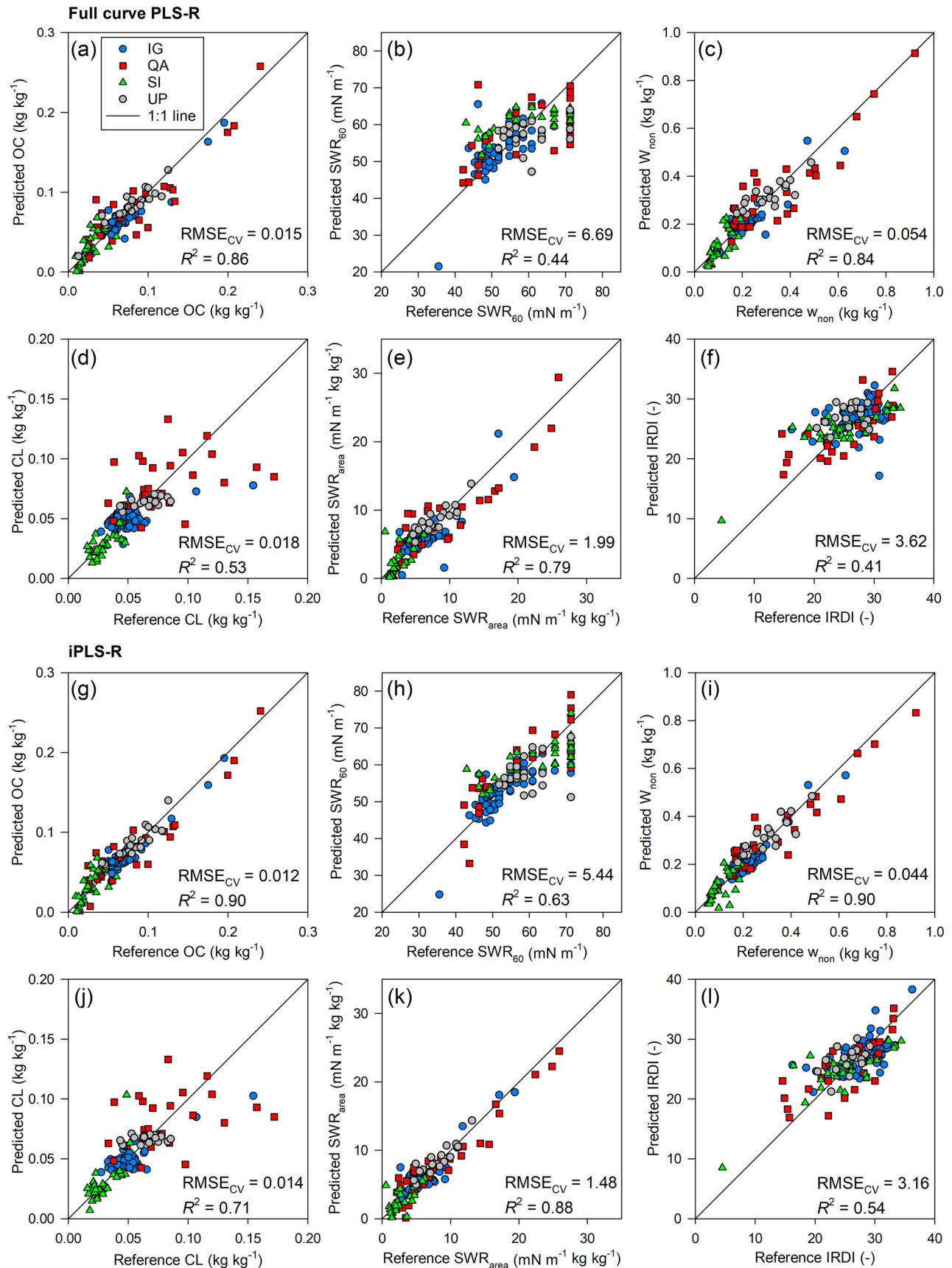


FIGURE 5 Reference and vis-NIRS-predicted (cross-validation) values for (a) organic carbon content (OC), (b) soil water repellency (SWR) after 60 °C heat pretreatment (SWR_{60}), (c) critical soil water content (w_{non}), (d) clay content (CL), (e) the area under SWR- w curve (SWR_{area}), and (f) the ratio between SWR_{area} and w_{non} (IRDI) using full-curve partial least squares regression (PLS-R), and (g-l) equivalent values using interval PLS-R (iPLS-R)

TABLE 3 Full-curve partial least squares regression (PLS-R) and interval PLS-R (iPLS-R) calibration and cross-validation results for organic carbon content (OC), clay content, and four soil water repellency (SWR) indices—SWR after 60 °C pretreatment (SWR_{60}), the area under SWR– w curve (SWR_{area}), critical soil water content (w_{non}), and the ratio between SWR_{area} and w_{non} (IRDI)—for 143 samples using visible–near-infrared spectral measurements

Property	Mean	NF	R^2		RMSE		SRMSE
			Calibration	Cross-validation	Calibration	Cross-validation	
Full-curve PLS-R							
OC, kg kg ⁻¹	0.06	9	.91	.86	0.011	0.015	0.06
Clay, kg kg ⁻¹	0.05	4	.67	.53	0.015	0.018	0.12
SWR_{60} , mN m ⁻¹	56.59	7	.57	.44	5.79	6.69	0.19
SWR_{area} , mN m ⁻¹ kg kg ⁻¹	6.17	9	.89	.79	1.42	1.99	0.08
w_{non} , kg kg ⁻¹	0.23	9	.91	.84	0.041	0.054	0.06
IRDI	26.17	9	.59	.41	2.96	3.62	0.11
iPLS-R							
OC, kg kg ⁻¹	0.06	11	.93	.90	0.010	0.012	0.05
Clay, kg kg ⁻¹	0.05	5	.76	.71	0.013	0.014	0.09
SWR_{60} , mN m ⁻¹	56.59	10	.69	.63	4.96	5.44	0.15
SWR_{area} , mN m ⁻¹ kg kg ⁻¹	6.17	12	.93	.88	1.15	1.48	0.06
w_{non} , kg kg ⁻¹	0.23	9	.92	.90	0.037	0.044	0.05
IRDI	26.17	9	.62	.54	2.84	3.16	0.10

Note. NF, number of factors; R^2 , coefficient of determination of calibration and cross-validation; SRMSE, standardized RMSE.

TABLE 4 Wavelength intervals for prediction of organic carbon content (OC), clay content and four soil water repellency (SWR) indices—SWR after 60 °C pretreatment (SWR_{60}), the area under SWR– w curve (SWR_{area}), critical soil water content (w_{non}), and the ratio between SWR_{area} and w_{non} (IRDI)—selected using interval partial least squares regression (iPLS-R)

Property	Wavelength intervals nm
OC	1,120–1,159.5, 1,520–1,559.5, 1,640–1,679.5, 1,720–1,759.5, 2,000–2,039.5, 2,160–2,199.5
Clay	720–759.5, 1,160–1,199.5, 1,520–1,559.5, 1,640–1,649.5, 2,080–2,119.5, 2,160–2,199.5
SWR_{60}	610–639.5, 1,090–1,149.5, 1,210–1,239.5, 1,450–1,479.5, 1,660–1,689.5, 2,050–2,139.5, 2,320–2,349.5, 2,410–2,439.5, 2,470–2,499.5
SWR_{area}	1,560–1,599.5, 1,720–1,799.5, 1,840–1,879.5, 2,040–2,199.5
w_{non}	1,450–1,549.5, 1,700–1,749.5, 2,050–2,099.5, 2,150–2,199.5
IRDI	700–759.5, 1,840–1,899.5, 2,380–2,439.5

Besides using full-curve vis–NIR spectra (400–2,500 nm) for modeling, variable selection by iPLS-R was additionally used to establish models based on specific spectral intervals. The performance of the iPLS-R models can also be found in Table 3. When using the selected intervals (Table 4), the estimation of OC improved slightly ($R^2 = .90$ and $RMSE_{CV} = 0.012$ kg kg⁻¹), while the estimation of clay improved significantly ($R^2 = .71$ and $RMSE_{CV} = 0.014$ kg kg⁻¹). The estimation of SWR_{60} now showed acceptable results ($R^2 = .63$ and $RMSE_{CV} = 5.44$ mN m⁻¹), and IRDI was estimated better than when using the entire spectral range, although still with a low model accuracy ($R^2 = .54$ and $RMSE_{CV} = 3.16$). The SWR_{area} and w_{non} were both still well estimated with slightly better results than when using the entire spectral range

($R^2 = .88$ and $RMSE_{CV} = 1.48$ mN m⁻¹ kg kg⁻¹ and $R^2 = .90$ and $RMSE_{CV} = 0.044$ kg kg⁻¹).

There was a decrease in SRMSE for all response variables when iPLS-R was used instead of full-curve PLS-R (Table 3). The models that showed the lowest SRMSE was for OC content (SRMSE = 0.06 and SRMSE = 0.05) and for w_{non} (SRMSE = 0.06 and SRMSE = 0.05) using full-curve and iPLS-R, respectively. The SWR_{area} was estimated with a similar low error using iPLS-R (SRMSE = 0.06). The model for IRDI had a relatively high error using full-curve PLS-R and iPLS-R, respectively (SRMSE = 0.11 and 0.10). The models for both clay content and SWR_{60} showed the greatest improvement in SRMSE from full-curve PLS-R to iPLS-R. However, the model for SWR_{60} still showed the

highest SRMSE for both approaches (SRMSE = 0.19 and 0.15). Clay, SWR₆₀, and IRDI all showed a high density of data around the mean with a relatively low range between maximum and minimum points (Figure 5b,d,f), which could cause the low model accuracy for these three indices.

This study showed that full-curve PLS-R and iPLS-R are reliable methods to predict OC content with high accuracy. This agrees with results on OC vis-NIRS predictions from Ogrič et al. (2019), who obtained high accuracy predictions for area-specific PLS-R models in South Greenland, and Hermansen et al. (2016), who included both Danish and Greenlandic fields. The good performance of the OC models in this study might be ascribed to the large range in OC content (mean, 6.3%; first quartile, 3.9%; third quartile, 7.5%) (Table 1) found in the sampled fields. This study also found a quite pronounced division between the predictive abilities of the highly OC-related SWR indices SWR_{area} and w_{non} with high-performance results and the two remaining SWR indices SWR₆₀ and IRDI with poor performance. Knadel et al. (2016) found a high prediction ability of SWR₆₀ ($R^2 = .85$ and $\text{RMSE}_{\text{CV}} = 2.52 \text{ mN m}^{-1}$) and soil water repellency after 105 °C heat pretreatment ($R^2 = .85$ and $\text{RMSE}_{\text{CV}} = 1.93 \text{ mN m}^{-1}$). However, as opposed to our study, they did not find a good predictive ability of OC because of low variability (OC, 0.014–0.025 kg kg⁻¹) and could therefore not ascribe the good performance of the SWR models to a good prediction of OC. Despite this, the good predictions of SWR_{area} and w_{non} show that using vis-NIRS to predict some SWR indices can be highly successful, and the accomplishment can probably be attributed to the strong correlation between these, OC, and the spectrally active components of soil organic matter.

3.6 | Spectral signatures

Regression coefficients indicate which wavelengths carry the most information for describing variation in the vis-NIR-predicted variables and can thus be used to analyze which spectrally active constituents may be important for the prediction of OC, clay, and SWR indices (Figure 6a–f). Based on Figure 6a, the most essential spectral regions for predicting OC were 400–700, 1,400–1,500, 1,700, and 2,200–2,400 nm. Other studies have found important absorption bands for OC around wavelengths of 480, 600–650, 1,724, and 2,300–2,350 nm (Hermansen et al., 2016); 2,275, 2,307–2,496, 2,137, and 2,381 nm (Viscarra Rossel & Behrens, 2010); and 1,700 and 2,000 nm (Stenberg, 2010). The region above 1,900 nm was specifically emphasized for predicting OC by Vohland and Emmerling (2011). The two well-predicted SWR indices—SWR_{area} and w_{non} —both showed almost identical regression coefficients compared with the regression coefficient of OC (Figure 6a–c), indicating that the same wavelengths were important in describing their variation. This could be a derived effect of the strong correlation found between OC

and SWR_{area} and OC and w_{non} in this study, and the role of OC in determining the severity of soil water repellency (de Jonge et al., 2009; Hermansen et al., 2019b).

The spectral intervals selected in the iPLS-R models are indicated with grey–blue areas in Figure 6a–f and listed in Table 4. From the iPLS-R intervals, we could conclude that including intervals in the visible region did not increase the prediction accuracy for either OC, SWR_{area}, or w_{non} . The OC fraction is spectrally active throughout the entire near-infrared range, where overtones and combination bands of OC compounds can be found (Ben-Dor & Banin, 1995; Kim et al., 2014). Six intervals between 1,120–2,200 nm were included in the best performing iPLS-R prediction of OC. Organic carbon, SWR_{area}, and w_{non} showed overlapping iPLS-R intervals around 1,550, 1,720–1,750, 2,040, and 2,160–2,200 nm. An interval around 1550 nm was selected for all three indices despite none of them showing pronounced peaks in the regression coefficient. Hermansen et al. (2019b) found intervals that increased predictive abilities using iPLS for SWR_{area} at 1,440–1,505 and 1,765–1,830 nm and for w_{non} at 1,445–1,500, 1,885–1,995, and 2,050–2,105 nm. The important spectral regions thus only overlap with those found in literature around 1,750 and 2,100–2,200 nm, suggesting a unique relation between relevant spectral intervals and SWR indices for the soils from South Greenland. Other studies have found important spectral regions for predicting SWR around 517–620, 750–880, 1,060–1,490, 1,765, 1,920, and 2,050–2,320 nm (Kim et al., 2014; Knadel et al., 2016; Davari et al., 2022). Absorption bands around 1,524 and 1,582 nm have been associated with amides (C=O) and O-H in water (Ben-Dor et al., 1997), respectively. A significant positive and negative peak in the regression coefficient coincides with the 1,720–1,750 nm interval selected for all three indices. Features around 1730 nm may be ascribed to vibrations of C-H stretches (Ben-Dor et al., 1997). These origin from aliphatic hydrocarbon (C-H) chains and hydrocarbons bonded with functional groups, which are thought to cause SWR as they act as a hydrophobic coating on particles (Doerr et al., 2000; McKissock et al., 2002; Knadel et al., 2016). Finally, the overlapping intervals between 2,040 and 2,200 nm showed some response in terms of regression coefficients that may be related to different organic compounds (Ben-Dor et al., 1997) as well as combinations of Al-OH bends with O-H stretches associated with clay minerals (Hunt, 1977).

The important spectral regions found in the regression coefficients for clay, SWR₆₀, and IRDI (Figure 6b,c,f) were spread out through the entire spectral range and did not appear to be intercorrelated in the same manner. This suggests that a more extensive use of the spectral data was necessary for prediction of these variables. While pronounced peaks can be found in the regression coefficients for clay at 1,400, 1,900, and 2,200 nm, similar to the findings of Hermansen et al. (2016) for Danish and Greenland soils, neither of the peaks

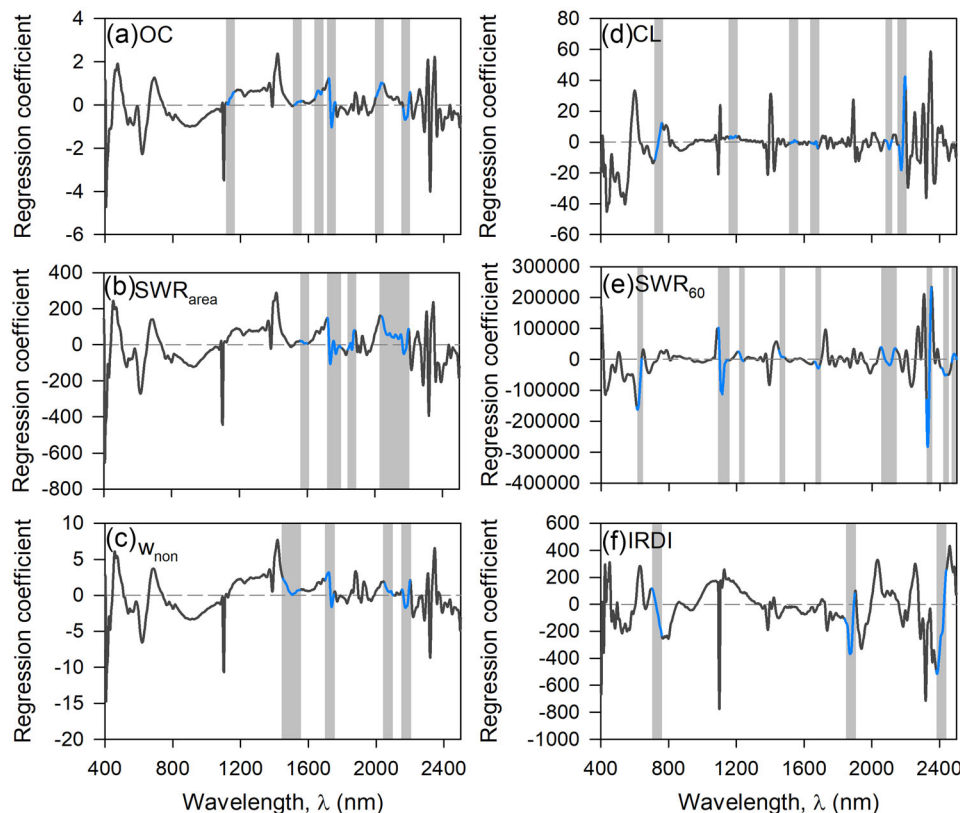


FIGURE 6 Regression coefficients for the full-curve partial least squares regression (PLS-R) models with wavelength intervals used in the prediction of each variable selected using interval PLS-R (iPLS-R) marked with grey and blue. From top left: (a) organic carbon (OC), (b) the area under SWR–w curve (SWR_{area}), (c) critical soil water content (w_{non}), (d) clay content (CL), (e) soil water repellency after 60 °C heat pretreatment (SWR_{60}), and (f) the ratio between SWR_{area} and w_{non} (IRDI)

were included in the six intervals selected for predicting clay (Table 4). These absorption bands are normally associated with structural water (1,400 and 1,900 nm) and aluminum content in clay minerals (2,200 nm) (Hunt, 1977) and usually comprise the main spectrally active regions found in clay. Nine intervals were shown to improve the predictive accuracy of SWR_{60} spanning almost the entire vis–NIR spectral range (610–2,500 nm). Note that one interval covers the detector shift at 1,100 nm. The IRDI required the least number of intervals.

The presented models performed well across the four included fields located in the main agricultural region of South Greenland. Several studies across different climates and soil types have observed a correlation between SWR_{area} and OC, suggesting a universal relation (de Jonge et al., 1999; Regalado & Ritter, 2005; Kawamoto et al., 2007; Wijewardana et al., 2016). Further, Weber et al. (2021) showed that samples from New Zealand and South Greenland followed the same linear trend in a SWR_{area} vs. OC plot, thus demonstrating that the effect of OC on SWR occurrence is comparable across climatically diverse locations. Analysis of the regression coefficients in this study showed that vis–NIRS predictions of SWR indices mainly relied on the information

from organic constituents. Because OC is a soil component, from which we see direct spectral responses throughout the vis–NIR spectral range, it is probably the main driver of the ability of vis–NIRS to predict SWR in this study. Given the suggested universal effect of OC on SWR_{area} , the presented vis–NIRS models in this study might be applicable for SWR prediction outside the study area. However, a prerequisite would be that OC can be vis–NIRS-predicted and further that OC and SWR_{area} correlate well. When a larger data set is available, an independent validation should be conducted to cover the full influence of a larger range in physical properties.

4 | CONCLUSIONS

This study showed that applying iPLS-R on vis–NIR spectra was found to perform either better or close to the same level as PTFs and better than full-curve PLS-R at obtaining prediction models for SWR indices— SWR_{60} , SWR_{area} , w_{non} , and IRDI—across 143 soil samples from the main agricultural region in South Greenland. This method could also provide accurate models for basic soil properties (OC and clay content).

The best results for estimating SWR_{area} was achieved when using an OC-based linear regression PTF, while w_{non} was best estimated using OC, clay, and pH in a MLR PTF. The high performance is partly credited to the strong correlation with OC. The non-OC related indices, SWR_{60} and IRDI, yielded poor PTF estimates. Although, with slightly lower accuracy than the PTFs, the full curve PLS-R vis–NIRS models generated estimates of SWR_{area} and w_{non} with a high precision, which was improved further when applying iPLS-R. Both SWR_{60} and IRDI were poorly estimated using full-curve PLS-R but were estimated with acceptable accuracy using iPLS-R. Consequently, vis–NIRS prediction models could estimate SWR indices with sufficient accuracy for the method to be a valuable addition to SWR mitigation for farmers in the region.

The spectral intervals containing most information of the SWR_{area} and w_{non} indices overlapped with iPLS-R intervals for OC prediction. Few specific spectral regions were found to carry important information for accurately estimating SWR_{area} and w_{non} while a more extensive use of the spectral range was needed for the estimation of SWR_{60} and IRDI.

In perspective, the suggested methodology for predicting SWR should be validated and further tested for soils representing an even broader range of soil textures and organic matter types and contents representing different climate zones.

AUTHOR CONTRIBUTIONS

Natasha H. Blaesbjerg: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing – original draft; Writing – review & editing. Peter L. Weber: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Writing – review & editing. Lis Wollesen de Jonge: Conceptualization; Data curation; Formal analysis; Funding acquisition; Methodology; Project administration; Resources; Supervision; Writing – review & editing. Per Moldrup: Conceptualization; Formal analysis; Funding acquisition; Methodology; Project administration; Resources; Supervision; Writing – review & editing. Mogens H. Greve: Conceptualization; Data curation; Investigation; Methodology; Project administration; Resources; Supervision; Writing – review & editing. Emmanuel Arthur: Conceptualization; Formal analysis; Methodology; Writing – review & editing. Maria Knadel: Conceptualization; Formal analysis; Methodology; Software; Supervision; Validation; Writing – review & editing. Cecilie Hermansen: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing.

ACKNOWLEDGMENTS

The authors would like to thank the farmers in South Greenland and the Greenland Agricultural Consulting Services for their important contribution during field surveys. This project

is funded by the Danish Council for Independent Research: Technology and Production Sciences, as part of the project: Newland: Glacial flour as a new, climate-positive technology for sustainable agriculture in Greenland (grant number: 022–00184B).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ORCID

Natasha H. Blaesbjerg  <https://orcid.org/0000-0001-8340-3590>

Peter L. Weber  <https://orcid.org/0000-0001-9249-0796>

Lis Wollesen de Jonge  <https://orcid.org/0000-0003-2874-0644>

Mogens H. Greve  <https://orcid.org/0000-0001-9099-8940>

Emmanuel Arthur  <https://orcid.org/0000-0002-0788-0712>

Maria Knadel  <https://orcid.org/0000-0001-7539-6191>

Cecilie Hermansen  <https://orcid.org/0000-0002-1925-3732>

REFERENCES

- Arthur, E. (2017). Rapid estimation of cation exchange capacity from soil water content. *European Journal of Soil Science*, 68, 365–373. <https://doi.org/10.1111/ejss.12418>
- Bachmann, J., Horton, R., van der Ploeg, R. R., & Woche, S. (2000). Modified sessile drop method for assessing initial soil-water contact angle of sandy soil. *Soil Science Society of America Journal*, 64, 564–567. <https://doi.org/10.2136/sssaj2000.642564x>
- Ben-Dor, E. (2002). Quantitative remote sensing of soil properties. *Advances in Agronomy*, 75, 173–243. [https://doi.org/10.1016/S0065-2113\(02\)75005-0](https://doi.org/10.1016/S0065-2113(02)75005-0)
- Ben-Dor, E., & Banin, A. (1995). Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Science Society of America Journal*, 59, 364–372. <https://doi.org/10.2136/sssaj1995.03615995005900020014x>
- Ben-Dor, E., Inbar, Y., & Chen, Y. (1997). The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sensing of Environment*, 61, 1–15. [https://doi.org/10.1016/S0034-4257\(96\)00120-4](https://doi.org/10.1016/S0034-4257(96)00120-4)
- Blackwell, P. S. (2000). Management of water repellency in Australia, and risks associated with preferential flow, pesticide concentration and leaching. *Journal of Hydrology*, 231–232, 384–395. [https://doi.org/10.1016/S0022-1694\(00\)00210-9](https://doi.org/10.1016/S0022-1694(00)00210-9)
- Chang, C.-W., Laird, D. A., Mausbach, M. J., & Hurburgh, C. R. (2001). Near-infrared reflectance spectroscopy - Principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65, 480–490. <https://doi.org/10.2136/sssaj2001.652480x>
- Davari, M., Fahmideh, S., & Mosaddeghi, M. R. (2022). Rapid assessment of soil water repellency indices using Vis-NIR spectroscopy and pedo-transfer functions. *Geoderma*, 406, 115486. <https://doi.org/10.1016/j.geoderma.2021.115486>
- DeJong, S. (1993). SIMPLS: An alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 18, 251–263. [https://doi.org/10.1016/0169-7439\(93\)85002-X](https://doi.org/10.1016/0169-7439(93)85002-X)

- de Jonge, L. W., Jacobsen, O. H., & Moldrup, P. (1999). Soil water repellency: Effects of water content, temperature, and particle size. *Soil Science Society of America Journal*, *63*, 437–442. <https://doi.org/10.2136/sssaj1999.03615995006300030003x>
- de Jonge, L. W., Moldrup, P., & Jacobsen, O. H. (2007). Soil-water content dependency of water repellency in soils. *Soil Science*, *172*, 577–588. <https://doi.org/10.1097/SS.0b013e318065c090>
- de Jonge, L. W., Moldrup, P., & Schjønning, P. (2009). Soil infrastructure, interfaces & translocation processes in inner space (“Soil-it-is”): Towards a road map for the constraints and crossroads of soil architecture and biophysical processes. *Hydrology and Earth System Sciences*, *13*, 1485–1502. <https://doi.org/10.5194/hess-13-1485-2009>
- Dekker, L. W., Doerr, S. H., Oostindie, K., Ziogas, A. K., & Ritsema, C. J. (2001). Water repellency and critical soil water content in a dune sand. *Soil Science Society of America Journal*, *65*, 1667–1674. <https://doi.org/10.2136/sssaj2001.1667>
- Dekker, L. W., & Ritsema, C. J. (1995). Fingerlike wetting patterns in two water-repellent loam soils. *Journal of Environmental Quality*, *24*, 324–333. <https://doi.org/10.2134/jeq1995.00472425002400020016x>
- Dekker, L. W., & Ritsema, C. J. (1996a). Preferential flow paths in a water repellent clay soil with grass cover. *Water Resources Research*, *32*, 1239–1249. <https://doi.org/10.1097/00010694-199702000-00001>
- Dekker, L. W., & Ritsema, C. J. (1996b). Variation in water content and wetting patterns in Dutch water repellent peaty clay and clayey peat soils. *Catena*, *28*, 89–105. [https://doi.org/10.1016/S0341-8162\(96\)00047-1](https://doi.org/10.1016/S0341-8162(96)00047-1)
- Deurer, M., Müller, K., Van Den Dijssel, C., Mason, K., Carter, J., & Clothier, B. E. (2011). Is soil water repellency a function of soil order and proneness to drought? A survey of soils under pasture in the North Island of New Zealand. *European Journal of Soil Science*, *62*, 765–779. <https://doi.org/10.1111/j.1365-2389.2011.01392.x>
- Dexter, A. R., Richard, G., Arrouays, D., Czyż, E. A., Jolivet, C., & Duval, O. (2008). Complexed organic matter controls soil physical properties. *Geoderma*, *144*, 620–627. <https://doi.org/10.1016/j.geoderma.2008.01.022>
- Doerr, S. H., Shakesby, R. A., & Walsh, R. P. D. (2000). Soil water repellency: Its causes, characteristics and hydro-geomorphological significance. *Earth-Science Reviews*, *51*, 33–65. [https://doi.org/10.1016/S0012-8252\(00\)00011-8](https://doi.org/10.1016/S0012-8252(00)00011-8)
- Esbensen, K. (1994). *Multivariate data analysis—In Practice: An introduction to multivariate data analysis and experimental design*. 5th ed. CAMO, Ålborg University.
- Giovannini, G., Lucchesi, S., & Cervelli, S. (1983). Water-repellent substances and aggregate stability in hydrophobic soil. *Soil Science*, *135*, 110–113.
- Hanna, E., & Cappelen, J. (2002). Recent climate of southern Greenland. *Weather*, *57*, 320–328. <https://doi.org/10.1256/00431650260283497>
- Hanna, E., Cappelen, J., Fettweis, X., Mernild, S. H., Mote, T. L., Mottram, R., Steffen, K., Ballinger, T. J., & Hall, R. J. (2021). Greenland surface air temperature changes from 1981 to 2019 and implications for ice-sheet melt and mass-balance change. *International Journal of Climatology*, *41*, 1336–1352. <https://doi.org/10.1002/joc.6771>
- Hermansen, C., Knadel, M., Moldrup, P., Greve, M. H., Gislum, R., & de Jonge, L. W. (2016). Visible-near-infrared spectroscopy can predict the clay/organic carbon and mineral fines/organic carbon ratios. *Soil Science Society of America Journal*, *80*, 1486–1495. <https://doi.org/10.2136/sssaj2016.05.0159>
- Hermansen, C., Moldrup, P., Müller, K., Knadel, M., & Jonge, L. W. (2019a). The relation between soil water repellency and water content can be predicted by vis-NIR spectroscopy. *Soil Science Society of America Journal*, *83*, 1616–1627. <https://doi.org/10.2136/sssaj2019.03.0092>
- Hermansen, C., Moldrup, P., Müller, K., Weber, P., van den Dijssel, C., Jeyakumar, P., & de Jonge, L. W. (2019b). Organic carbon content controls the severity of water repellency and the critical moisture level across New Zealand pasture soils. *Geoderma*, *338*, 281–290. <https://doi.org/10.1016/j.geoderma.2018.12.007>
- Hunt, G. R. (1977). Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics*, *42*, 502–513. <https://doi.org/10.1190/1.1440721>
- Jacobsen, N. K. (1987). Studies on soils and potential for soil erosion in the sheep farming area of South Greenland. *Arctic and Alpine Research*, *19*, 498–507. <https://doi.org/10.2307/1551416>
- Karunaratna, A. K., Moldrup, P., Kawamoto, K., de Jonge, L. W., & Komatsu, T. (2010). Two-region model for soil water repellency as a function of matric potential and water content. *Vadose Zone Journal*, *9*, 719–730. <https://doi.org/10.2136/vzj2009.0124>
- Katuwal, S., Hermansen, C., Knadel, M., Moldrup, P., Greve, M. H., & de Jonge, L. W. (2018). Combining x-ray computed tomography and visible near-infrared spectroscopy for prediction of soil structural properties. *Vadose Zone Journal*, *17*, 1–13. <https://doi.org/10.2136/vzj2016.06.0054>
- Kawamoto, K., Moldrup, P., Komatsu, T., de Jonge, L. W., & Oda, M. (2007). Water repellency of aggregate size fractions of a volcanic ash soil. *Soil Science Society of America Journal*, *71*, 1658–1666. <https://doi.org/10.2136/sssaj2006.0284>
- Kim, I., Pullanagari, R. R., Deurer, M., Singh, R., Huh, K. Y., & Clothier, B. E. (2014). The use of visible and near-infrared spectroscopy for the analysis of soil water repellency: Vis-NIR spectroscopy analysis of soil water repellency. *European Journal of Soil Science*, *65*, 360–368. <https://doi.org/10.1111/ejss.12138>
- King, P. (1981). Comparison of methods for measuring severity of water repellence of sandy soils and assessment of some factors that affect its measurement. *Soil Research*, *19*, 275–285. <https://doi.org/10.1071/SR9810275>
- Knadel, M., Masís-Meléndez, F., de Jonge, L. W., Moldrup, P., Arthur, E., & Greve, M. H. (2016). Assessing soil water repellency of a sandy field with visible near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, *24*, 215–224. <https://doi.org/10.1255/jnirs.1188>
- Leighton-Boyce, G., Doerr, S. H., Shakesby, R. A., & Walsh, R. P. D. (2007). Quantifying the impact of soil water repellency on overland flow generation and erosion: A new approach using rainfall simulation and wetting agent on in situ soil. *Hydrological Processes*, *21*, 2337–2345. <https://doi.org/10.1002/hyp.6744>
- Letej, J., Carrillo, M. L. K., & Pang, X. P. (2000). Approaches to characterize the degree of water repellency. *Journal of Hydrology*, *231*–232, 61–65. [https://doi.org/10.1016/S0022-1694\(00\)00183-9](https://doi.org/10.1016/S0022-1694(00)00183-9)
- Masson-Delmotte, V., Swingedouw, D., Landais, A., Seidenkrantz, M.-S., Gauthier, E., Bichet, V., Massa, C., Perren, B., Jomelli, V., Adalgeirsdottir, G., Hesselbjerg Christensen, J., Arneborg, J., Bhatt, U., Walker, D. A., Elberling, B., Gillet-Chaulet, F., Ritz, C., Gallée, H., & van den Broeke, M. (2012). Greenland climate change: From the past to the future. *WIREs Climate Change*, *3*, 427–449. <https://doi.org/10.1002/wcc.186>
- McKissock, I., Gilkes, R. J., & Walker, E. L. (2002). The reduction of water repellency by added clay is influenced by clay and soil prop-

- erties. *Applied Clay Science*, 20, 225–241. [https://doi.org/10.1016/S0169-1317\(01\)00074-6](https://doi.org/10.1016/S0169-1317(01)00074-6)
- Müller, K., Mason, K., Strozzi, A. G., Simpson, R., Komatsu, T., Kawamoto, K., & Clothier, B. (2018). Runoff and nutrient loss from a water-repellent soil. *Geoderma*, 322, 28–37. <https://doi.org/10.1016/j.geoderma.2018.02.019>
- Norris, K. H. (2001). Understanding and correcting the factors which affect diffuse transmittance spectra. *NIR News*, 12, 6–9. <https://doi.org/10.1255/nirn.613>
- Ogrič, M., Knadel, M., Kristiansen, S. M., Peng, Y., De Jonge, L. W., Adhikari, K., & Greve, M. H. (2019). Soil organic carbon predictions in Subarctic Greenland by visible–near infrared spectroscopy. *Arctic, Antarctic, and Alpine Research*, 51, 490–505. <https://doi.org/10.1080/15230430.2019.1679939>
- Pribyl, D. W. (2010). A critical review of the conventional SOC to SOM conversion factor. *Geoderma*, 156, 75–83. <https://doi.org/10.1016/j.geoderma.2010.02.003>
- Regalado, C. M., & Ritter, A. (2005). Characterizing water dependent soil repellency with minimal parameter requirement. *Soil Science Society of America Journal*, 69, 1955–1966. <https://doi.org/10.2136/sssaj2005.0060>
- Regalado, C. M., Ritter, A., de Jonge, L. W., Kawamoto, K., Komatsu, T., & Moldrup, P. (2008). Useful soil-water repellency indices: Linear correlations. *Soil Science*, 173, 747–757. <https://doi.org/10.1097/SS.0b013e31818d4163>
- Rinnan, Å., van den Berg, F., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28, 1201–1222. <https://doi.org/10.1016/j.trac.2009.07.007>
- Roy, J. L., & McGill, W. B. (2002). Assessing soil water repellency using the molarity of ethanol droplet (MED) test. *Soil Science*, 167, 83–97. <https://doi.org/10.1097/00010694-200202000-00001>
- Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36, 1627–1639. <https://doi.org/10.1021/ac60214a047>
- Scheinost, A. C., Chavernas, A., Barrón, V., & Torrent, J. (1998). Use and limitations of second-derivative diffuse reflectance spectroscopy in the visible to near-infrared range to identify and quantify Fe oxide minerals in soils. *Clays and Clay Minerals*, 46, 528–536. <https://doi.org/10.1346/CCMN.1998.0460506>
- Seaton, F. M., Jones, D. L., Creer, S., George, P. B. L., Smart, S. M., Lebron, I., Barrett, G., Emmett, B. A., & Robinson, D. A. (2019). Plant and soil communities are associated with the response of soil water repellency to environmental stress. *Science of the Total Environment*, 687, 929–938. <https://doi.org/10.1016/j.scitotenv.2019.06.052>
- Stenberg, B. (2010). Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on vis-NIR predictions of clay and soil organic carbon. *Geoderma*, 158, 15–22. <https://doi.org/10.1016/j.geoderma.2010.04.008>
- Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., & Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*, 107, 163–215. [https://doi.org/10.1016/S0065-2113\(10\)07005-7](https://doi.org/10.1016/S0065-2113(10)07005-7)
- Viscarra Rossel, R. A., & Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158, 46–54. <https://doi.org/10.1016/j.geoderma.2009.12.025>
- Vohland, M., & Emmerling, C. (2011). Determination of total soil organic C and hot water-extractable C from VIS-NIR soil reflectance with partial least squares regression and spectral feature selection techniques. *European Journal of Soil Science*, 62, 598–606. <https://doi.org/10.1111/j.1365-2389.2011.01369.x>
- Wallis, M. G., & Horne, D. J. (1992). Soil water repellency. *Advances in Soil Science*, 20, 91–144.
- Watson, C. L., & Letey, J. (1970). Indices for characterizing soil-water repellency based upon contact angle-surface tension relationships. *Soil Science Society of America Journal*, 34, 841–844. <https://doi.org/10.2136/sssaj1970.03615995003400060011x>
- Weber, P. L., Hermansen, C., Norgaard, T., Pesch, C., Moldrup, P., Greve, M. H., Müller, K., Arthur, E., & de Jonge, L. W. (2021). Moisture-dependent water repellency of Greenlandic cultivated soils. *Geoderma*, 402, 115189. <https://doi.org/10.1016/j.geoderma.2021.115189>
- Westergaard-Nielsen, A., Björnsson, A. B., Jepsen, M. R., Stendel, M., Hansen, B. U., & Elberling, B. (2015). Greenlandic sheep farming controlled by vegetation response today and at the end of the 21st century. *Science of the Total Environment*, 512–513, 672–681. <https://doi.org/10.1016/j.scitotenv.2015.01.039>
- Wijewardana, N. S., Müller, K., Moldrup, P., Clothier, B., Komatsu, T., Hiradate, S., de Jonge, L. W., & Kawamoto, K. (2016). Soil-water repellency characteristic curves for soil profiles with organic carbon gradients. *Geoderma*, 264, 150–159. <https://doi.org/10.1016/j.geoderma.2015.10.020>
- Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58, 109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1)
- Zou, X., Zhao, J., Povey, M. J. W., Holmes, M., & Mao, H. (2010). Variables selection methods in near-infrared spectroscopy. *Analytica Chimica Acta*, 667, 14–32. <https://doi.org/10.1016/j.aca.2010.03.048>

How to cite this article: Blaesbjerg, N. H., Weber, P. L., de Jonge, L. W., Moldrup, P., Greve, M. H., Arthur, E., Knadel, M., & Hermansen, C. (2022). Water repellency prediction in high-organic Greenlandic soils: Comparing vis-NIRS to pedotransfer functions. *Soil Science Society of America Journal*, 86, 643–657. <https://doi.org/10.1002/saj2.20407>