Terrain is a stronger predictor of peat depth than airborne radiometrics in Norwegian landscapes

Julien Vollering[1], Naomi Gatis[2], Mette Kusk Gillespie[1], Karl-Kristian Muggerud[1], Sigurd Daniel Nerhus[1], Knut Rydgren[1], and Mikko Sparf[1]

[1]Department of Civil Engineering and Environmental Sciences, Western Norway University of Applied Sciences, Norway
[2]Department of Geography, University of Exeter, United Kingdom

Correspondence: Julien Vollering (julien.vollering@hvl.no)

Abstract. Peatlands are Earth's most carbon-dense terrestrial ecosystems and their carbon density varies with the depth of the peat layer. Accurate mapping of peat depth is crucial for carbon accounting and land management, yet existing maps lack the resolution and accuracy needed for these applications. This study evaluates whether digital soil mapping using remotely sensed data can improve existing maps of peat depth in western and southeastern Norway. Specifically, we assessed the predictive
5   value of LiDAR-derived terrain variables and airborne radiometric data across two, >10 km^2 sites. We measured peat depth by probing and ground-penetrating radar at 372 and 1878 locations at the two sites, respectively. Then we trained Random Forest models using radiometric and terrain variables, plus the national map of peat depth, to predict peat depth at 10 m resolution. The two best models achieved mean absolute errors of 60 and 56 cm, explaining one-third of the variation in peat depth. Terrain variables were better predictors than radiometric variables, with elevation and valley bottom flatness showing the strongest
10 relationships to depth. Radiometric variables showed inconsistent and weak predictive value – improving performance at one site while degrading it at the other. Our remote sensing models had better accuracy than the national map of peat depth, even when we calibrated the national map to the same depth data. Still, weak relationships with remotely sensed variables made peat depth hard to predict overall. Based on these findings, we conclude that digital soil mapping can improve the existing, national map of peat depth in Norway, but detailed local maps are best made from tailored field measurements. Together, these
15 pathways promise more accurate landscape-scale carbon stock assessments and better-informed land management policies.

1 Introduction

Peat soils are a terrestrial carbon stock of global importance. They store 450-650 Gt of carbon, or about 30 % of global soil carbon, despite covering only 2–3 % of Earth's land (Xu et al., 2018; Friedlingstein et al., 2020; UNEP, 2022). Peatlands (areas with peat soil) are more carbon dense per square meter than any other ecosystem (Temmink et al., 2022). This makes them
20 crucial to climate change mitigation. Intact peatlands sequester carbon and overall produce a negative temperature forcing (Joosten et al., 2016). When disturbed – often by conversion to another land use – they can produce large greenhouse gas emissions (Ma et al., 2022).

One of the keys to the areal density of peatland carbon stocks (and variation therein) lies in the third dimension: their depth. Peat soils range from zero to over ten meters deep (Widyastuti et al., 2024). Their depth depends on the rate and duration of organic matter accumulation, potentially over thousands of years (Loisel et al., 2014; Joosten et al., 2016). In the anoxic and acidic conditions created by a high water table, plant material decay is slightly outweighed by new growth, and the surplus carbon is laid down as peat.

Peatlands are most common at high latitudes, and in regions with high cover they are frequently converted to human land use (UNEP, 2022). They are attractive for agriculture and forestry because they are flat, treeless, and have developed soils, but other land uses also displace peatlands. In Norway – where flat lowland is relatively scarce but upward of 9 % of the land area is peatland (Bryn et al., 2018; Bakkestuen et al., 2023) – lawmakers have restricted peatland afforestation and cultivation in recent decades. Since then, a larger proportion of peatland loss is driven by construction (Flaget et al., 2024).

Variation in peat depth and its spatial distribution is often overlooked in land use planning and carbon accounting because peat depth is not mapped with sufficient coverage, resolution, or accuracy (Beilman et al., 2008; Hastie et al., 2022; UNEP, 2022). Maps are crucial because, unlike spatially aggregated estimates, they link high-level targets to specific management decisions (OECD, 2022). Maps of peat depth also make it possible to quantify the effect of specific management decisions and thereby understand how local outcomes contribute to regional and national outcomes (OECD, 2022).

Measuring peat depth on the ground is straightforward, and a field survey can map a small area at low cost. However, surveying large areas is impractical when depth varies widely over short distances (e.g., 10 m) – as in many peatlands (Torppa, 2011; Proulx-McInnis et al., 2013; Henrion et al., 2024). A complementary approach from soil science is digital soil mapping (DSM). DSM scales up field measurements from a set of locations to a wider area, by relating the measured values to other variables mapped over the area of interest, through a statistical model. This approach has grown in importance with the availability of remotely sensed data and the advancement of machine learning methods (Minasny et al., 2019; Wadoux et al., 2020).

The crux for DSM of peat depth is the relationship between peat depth and the other, mapped predictors. For DSM to be effective, these relationships must be strong and consistent over the area of interest. They can be purely correlative rather than causal, but mechanistic relationships are stronger and more consistent than non-causal ones. The *scorpan* framework for DSM suggests seven predictor classes to explore: other soil properties, climate, organisms, relief (topography), parent material, age, and spatial position (McBratney et al., 2003).

For peat depth, the most practical and widespread *scorpan* factors are relief and spatial position. Spatial position is unique because it is always known (with varying accuracy). However, the short range of spatial autocorrelation in peat depth limits its mapping value (Hengl et al., 2004). Relief, or topography, is widely and accurately mapped in digital terrain models (DTMs). For example, most of mainland Norway has at least one elevation measurement per square meter from airborne Light Detection and Ranging (LiDAR) surveys. Moreover, mechanisms of peat formation are linked to topography.

Studies have consistently shown relationships between peat depth and topography, though the specific patterns vary. One of the most robust relationships is a negative correlation between depth and terrain slope (e.g., Holden and Connolly, 2011; Parry et al., 2012; Gatis et al., 2019). Peat depth also changes with elevation in many contexts, but with inconsistent directionality

(e.g., Holden and Connolly, 2011; Parry et al., 2012; Rudiyanto et al., 2016, 2018; Koganti et al., 2023; Li et al., 2024). Also more complex derivations of topography, such as the Topographic Wetness Index and the Multi-Resolution Valley Bottom Flatness index, have shown associations with peat depth (e.g., Rudiyanto et al., 2018; Koganti et al., 2023; Li et al., 2024; Pohjankukka et al., 2025). Some of the variation between studies in quantifying these relationships is undoubtedly attributable to issues of spatial scale – both the scaling of the topographic predictors and the resolution of the peat depth analysis.

Another set of predictors related to peat depth are measurements of natural radioactivity from the ground surface. Gamma-ray spectrometry can survey the activity (decay counts per second) from radioactive isotopes in bedrock and mineral soils: potassium-40, uranium-238, and thorium-232 (Beamish, 2014; Reinhardt and Herrmann, 2019). Although survey measurements are commonly reported as ground concentrations (linearly scaled from decay counts per second), in peatland environments these predictors do not reflect the concentration of radionuclides near the ground surface, but rather the radiation intensity after attenuation by the peat overburden. Deep soil with high water content will attenuate radiation most (Beamish, 2013; Reinhardt and Herrmann, 2019). Thus, reported ground concentrations can be statistically informative about peat soils, even if not physically correct, with respect to two *scorpan* factors: other soil properties and parent material (McBratney et al., 2003).

Although theory suggests that one meter of peat may fully attenuate radiometric signal (Beamish, 2013; Reinhardt and Herrmann, 2019), empirical investigations show that the association between peat depth and radioactivity can extend beyond the first meter (Keaney et al., 2013; Gatis et al., 2019; Koganti et al., 2023). Radiometric data are increasingly available over large areas (Minasny et al., 2019; Baranwal and Rønning, 2020), and – insofar as they are predictive of peat depth – this presents an opportunity for DSM.

The effectiveness of DSM depends not only on the methods and data, but also on the characteristics of the mapping area. Norway may be instructive in this respect because its peatlands vary widely across climates and topographies. Norway has 22 peatland mesotopes, with different hydrology, formation, and development – from topogeneous or soligeneous fens to blanket or raised bogs (Joosten and Clarke, 2002; Lyngstad et al., 2023). These mesotopes have fundamentally different geomorphology, which suggests that relationships between peat depth and terrain or radiometric predictors may vary by landscape.

The only systematic mapping of peat depth at the national scale in Norway comes from surveys meant to identify arable land. After scattered efforts in the early 20th century, a comprehensive round of surveying was completed 1964–2001 as part of a wider land cover mapping in Norway (Bjørdal, 2007). This survey produced the initial maps from which Norway's most detailed updated land cover datasets are derived – including the AR5 and DMK datasets used in this study and described later. Because of its agricultural and silvicultural focus, the survey covered only productive areas below the tree line, and peat depth was measured only in places judged to be potentially arable or afforestable (Ahlstrøm et al., 2019). Field surveyors carried a 1 m probe and they measured peat depth as a categorical variable: *shallow* (<1 m), or *deep* (>1 m). These classes were assigned to whole polygons, so spatial resolution is on the order of hectares. The number of measurements per unit area was not standardized but probably low.

The push for nature-based climate solutions motivates for extensive mapping of peat depth, to identify peatlands with rich carbon stocks to avoid their conversion and prioritize their preservation or restoration (Strack et al., 2022). Although ground

surveys of peat depth are accurate and feasible across small areas, having landscape-scale maps available before detailed investigation increases the option space for spatial planning and land management. For example, the Norwegian Public Roads Administration routinely measures peat depth during geotechnical work before road construction, but by that time, the route of the road is already set. With a landscape-scale peat depth map, planners could better compare the climate impact of different routes. To make landscape-scale maps possible, more studies need to investigate the relationships between peat depth and topographic or radiometric variables, to determine how strong and consistent these are.

Here, we assess how well remotely sensed topographic and radiometric data can predict peat depth at two contrasting sites, with a view toward revising regional and national maps in Norway. Specifically, we aim to: (1) quantify the accuracy of predictions from topographic and radiometric variables, and (2) identify key predictive variables. Reflecting on our results and the need for continued improvement in peat depth maps, we close by discussing implications for digital soil mapping of peat depth.

To our knowledge, this is the first study to predict continuous peat depth from airborne radiometric data using machine learning. Where airborne radiometric data have previously been used to predict continuous peat depth, it has been through modeling techniques less suited for prediction and spatial extrapolation (e.g., Keaney et al., 2013; Gatis et al., 2019; Siemon et al., 2020). Where machine learning has previously been used with airborne radiometric data, it has been to predict peat extent or depth classes rather than continuous depth (e.g., O'Leary et al., 2022; Pohjankukka et al., 2025).

2 Materials and methods

2.1 Sites

To assess the predictive value of terrain and radiometric data for mapping peat depth, we chose two sites with conspicuously different physical geography: Skrimfjella in eastern Norway (Fig. 1a) and Ørskogfjellet in western Norway (Fig. 1b). These sites were chosen because they were covered by radiometric data from airborne surveys, had relatively little built-up area, and were accessible by road. Ørskogfjellet was additionally chosen for its relatively uniform bedrock geology, which should improve the peat depth signal in radiometric data, all else being equal (Minasny et al., 2019).

At Skrimfjella we delineated a study area of 34 km^2 based on radiometric coverage (limiting to the west) and accessibility (limiting to the south), as part of a pilot project (Fig. 1a). In Norway's AR5 national land cover dataset ("areal resources in scale 1:5000", Ahlstrøm et al., 2019), 1.5 km^2 (4.5 %) of the study area is classified as 'mire' – defined as areas with mire vegetation and at least 30 cm of peat depth. Relatively sparse peatland cover did not disqualify the area for our purposes, since we were also interested in peatland extent mapping in the pilot project. The study area has a diverse bedrock, with 32 % alkali feldspar granite, 26 % marlstone, 10 % granite, 8 % monzonite, 7 % sandstone, 6 % limestone, and five other rock types with 1–5 % coverage (Geological Survey of Norway, 1:250 000 dataset). It is almost without human infrastructure, dominated by conifer forest, and borders on a nature reserve. Its mean elevation is 438 m a.s.l. (range 223–711 m, IQR 351–509 m), and its mean slope at 10 m resolution is $10.8°$ (IQR $4.6 - 15.1°$).

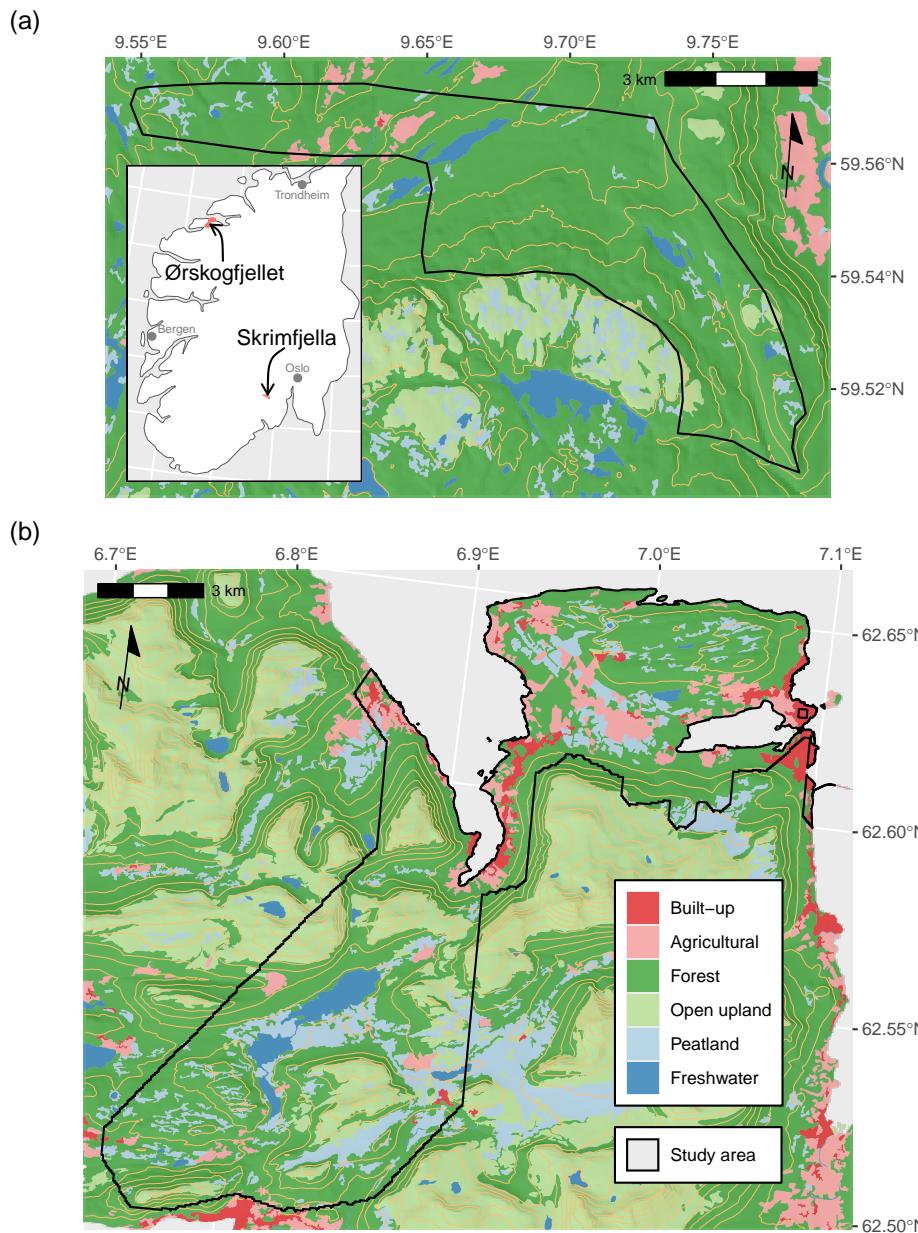



Figure 1. Study areas with land cover at Skrimfjella (a) and Ørskogfjellet (b). To fit the scale of the maps, the land cover shown here is from the AR50 national land resource dataset, which has simplified geometry with respect to the AR5 dataset that is used in the study. Terrain is visualized with light orange contour lines at intervals of 100 m, and a hillshade with slope $45°$ and azimuth $225°$. AR50 data are from Norwegian Institute of Bioeconomy Research under the Norwegian Licence for Open Government Data (https://data.norge.no/nlod/en/1.0) and terrain data are from the Norwegian Mapping Authority under the Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0/).

125   At Ørskogfjellet we defined a study area of 124 km^2 which basically followed the southernmost part of the radiometric survey extent (Fig. 1b). In the AR5 dataset, 15.3 km^2 (12.4 %) of the study area is classified as mire. Bedrock in the area is 84 % granitic gneiss, 11 % granite, and 5 % aluminium silicate gneiss (Geological Survey of Norway, 1:250 000 dataset). This study area is mostly forested, but also contains considerable farmland and open upland, and has several large lakes. Its mean elevation is 211 m above sea level (range 0–807 m, IQR 73–310 m), and its mean slope at 10 m resolution is 13.0° (IQR

130     $4.7 - 18.3°$).

2.2   Peat depth predictors

We created the same suite of peat depth predictors for both sites (25 continuous and 1 categorical; Table 1). All continuous predictors were derived either from an airborne radiometric survey or from a DTM. From the radiometric surveys, we simply used the four variables produced by the Geological Survey of Norway: ground concentration of potassium, thorium, uranium,

135   as well as total count. From the DTMs we calculated several land surface parameters, ranging from simple terrain indices to more complex geomorphometric and hydrological variables (Maxwell and Shobe, 2022). The categorical predictor was peat depth class, from a national map dataset. Predictor preparation is described in more detail below.

    We chose a spatial resolution of 10 m for our predictors, depth sampling, and modeling. We considered this a reasonable compromise between DTM resolution (1 m) and small peatlands on the one hand, and airborne radiometric resolution (50 m)

140   on the other hand.

2.2.1   Radiometric

The Geological Survey of Norway conducted and processed radiometric surveys over our study areas, as reported in Baranwal et al. (2013) and Ofstad (2015). They provided us for each site four variables at 50 m resolution, which we downscaled to 10 m resolution by cubic spline resampling. Sensitivity analysis showed that Pearson correlations between cubic spline and bilinear

145   resampling methods exceeded 0.995 for all radiometric variables, so we are confident that the choice of resampling method did not affect our results.

    The radiometric surveys were conducted with 75–80 m average flight altitude, 88–108 $\mathrm{km\,h^{-1}}$ average flight speed, and 200 m flight line spacing (Table A1). Spectrometer count rates were calibrated to known concentrations of potassium, thorium, and uranium in mobile pads. Data were processed following standard procedures outlined by the International Atomic Energy

150   Association. Processing included: correction for aircraft and cosmic background radiation, correction for radon in the air, window stripping of the gamma ray spectrum, correction for flying height, conversion of count rates to ground concentrations, and finally gridding to 50 m resolution with micro-leveling. At Ørskogfjellet an additional convolution filter was added to smooth the gridded data.

    Although radiometric data must be in units of counts per second to model attenuation directly (O'Leary et al., 2022), we

155   used the radiometric data as provided to us: in units of concentration for potassium, thorium, and uranium (converted from counts per second by scalar calibration factors). The monotonic transformation between counts per second and concentration has no effect on the tree-based machine learning algorithm that we used to model peat depth (Hastie et al., 2009). We also used

Table 1. Twenty-six candidate predictors of peat depth. Note that each continuous predictor contributes one variable to the model while the single categorical predictor (DMK) contributes two binary variables, corresponding to its three levels (one level is used as the reference). This results in a total of 27 variables.

Group	Code	Units	Description
radiometric	radK	%	Potassium ground concentration
	radTh	ppm	Thorium ground concentration
	radU	ppm	Uranium ground concentration
	radTC	$\mathrm{counts\,s^{-1}}$	Total Count of gamma radiation
terrain	elevation	m	mean elevation from DTM with 1 m resolution
	slope1m	degrees	mean of slope at 1 m resolution
	TPI1m	m	mean of Topographic Position Index at 1 m resolution
	TRI1m	m	mean of Terrain Ruggedness Index at 1 m resolution
	roughness1m	m	mean of roughness at 1 m resolution
	slope10m	degrees	slope from DTM with 10 m resolution
	TPI10m	m	Topographic Position Index from DTM with 10 m resolution
	TRI10m	m	Terrain Ruggedness Index from DTM with 10 m resolution
	roughness10m	m	roughness from DTM with 10 m resolution
	MRVBF	dimensionless	Multi-Resolution Valley Bottom Flatness
	TWI5m	dimensionless	mean of Topographic Wetness Index at 5 m resolution
	TWI10m	dimensionless	Topographic Wetness Index at 10 m resolution
	TWI20m	dimensionless	bilinear interpolation of Topographic Wetness Index at 20 m resolution
	TWI50m	dimensionless	bilinear interpolation of Topographic Wetness Index at 50 m resolution
	DTW2500	m	Depth-To-Water index, flow initiation area of 0.25 ha
	DTW5000	m	Depth-To-Water index, flow initiation area of 0.5 ha
	DTW10000	m	Depth-To-Water index, flow initiation area of 1 ha
	DTW20000	m	Depth-To-Water index, flow initiation area of 2 ha
	DTW40000	m	Depth-To-Water index, flow initiation area of 4 ha
	DTW80000	m	Depth-To-Water index, flow initiation area of 8 ha
	DTW160000	m	Depth-To-Water index, flow initiation area of 16 ha
DMK	DMK	categorical	peat depth class in the DMK national map dataset: shallow/deep/unknown

the total count variable as provided, rather than calculating a gamma dose rate based on the potassium, thorium, and uranium (as was done in Gatis et al., 2019), because these were highly correlated at both study sites ($\rho = 0.989, 0.986$), and because conversions to dose rates are approximations (IAEA, 2003).

2.2.2 Terrain

For terrain-derived predictors, we obtained 1 m resolution rasters from the national DTM (CC BY 4.0 Norwegian Mapping Authority). The DTM for Skrimfjella was produced from airborne laser scanning surveys in 2015 and 2022, with laser point density of $5\,\mathrm{pts}\,\mathrm{m}^{-2}$. For Ørskogfjellet, the DTM was produced from a 2015 survey with $2\,\mathrm{pts}\,\mathrm{m}^{-2}$. Where necessary, DTMs were resampled to the coordinate reference system of the radiometric data.

We used the *terra* R package (v1.8) to calculate from the DTMs: slope, Topographic Position Index (difference from mean of eight neighbors), Terrain Ruggedness Index (mean of absolute differences from eight neighbors), and roughness (range in the nine-cell neighborhood). These were derived at two scales to produce eight different predictors; we either calculated the indices at 1 m DTM resolution and then aggregated to 10 m resolution, or aggregated to 10 m DTM resolution and then calculated the indices. This kind of multiscale feature engineering of land surface parameters has been found to improve machine learning predictions of soil properties (Miller et al., 2015; Dornik et al., 2022; Newman et al., 2023). We know that peat depth tends to vary at fine scales in Norway, which is why we chose 1 m and 10 m resolutions (Maxwell and Shobe, 2022). We also calculated the Multi-Resolution Valley Bottom Flatness index, which indicates the degree of valley bottom flatness at a given location via a multiscale algorithm (Gallant and Dowling, 2003). We calculated this index in SAGA GIS (v.9.3.2, Morphometry library, Conrad et al., 2015) with default parameters (initial slope threshold = 16 %, lowness threshold = 0.4, upness threshold = 0.35, slope shape parameter = 4, elevation shape parameter = 3).

The Topographic Wetness Index (Quinn et al., 1991) is notoriously scale-dependent and often matches real hydrological conditions best when calculated from moderate to coarse resolution DTMs (Ågren et al., 2014; Riihimäki et al., 2021), so we calculated it from 5 m, 10 m, 20 m, and 50 m DTM resolution. The calculations were performed with Whitebox software (Lindsay, 2016a), accessed through the *whitebox* R package (v2.4, Wu and Brown, 2022). We filled depressions in the DTM with the algorithm in Wang & Liu (2006), and used the deterministic infinity flow accumulation algorithm (Tarboton, 1997).

The Depth-to-Water index (Murphy et al., 2007) approximates a location's vertical height above the surface water feature (e.g., stream, lake, or sea) that it is likely to drain towards. It is calculated as the minimum cumulative slope (scaled by cell size) to a surface water feature (eq. 5 in Murphy et al., 2009). We calculated unitless slope from the 1 m DTM using Whitebox. We defined surface water features from the DTM by filling depressions and then calculating flow accumulation to define catchment areas for each cell (Schönauer et al., 2021; Schönauer and Maack, 2021). This catchment area layer was then thresholded at seven different levels (flow initiation area 0.25–16 ha) to estimate surface water features under moisture scenarios varying from wet to dry (Murphy et al., 2011; Ågren et al., 2014; Schönauer et al., 2021). In addition, all surface water features mapped in the AR5 dataset were also transferred to the raster layer. For each of the seven surface water layers, we derived Depth-to-Water using the *Distance Accumulation* tool in ArcGIS Pro (v.3.1, ESRI, USA), which has an efficient algorithm to find the cumulative distance over a cost surface to the least-cost source.

2.2.3 Peat depth class

We prepared one categorical predictor – peat depth class – from the national map dataset called *DMK* (Ahlstrøm et al., 2019). The DMK dataset is derived from the same historical surveys as the AR5 dataset, and peat depth classes are: <1 m (*shallow*), >1 m (*deep*), and *unknown*. Surveyors generally assigned peat depth classes to polygons of at least 0.5 ha, although delineating polygons down to 0.2 ha was allowed if peat depth showed a "particularly marked difference" (Bjørdal, 2007). We rasterized the peat depth class attribute to our 10 m grid and this predictor equates to two variables in the model because its three levels become two indicator variables (one level is used as the reference).

2.3 Peat depth sample selection

The places (10 m raster cells) we chose to measure peat depth were sampled from mire areas in the AR5 dataset, and optimized for training a Random Forest (RF) model of peat depth (Brus, 2019). Broadly, we aimed for a sample that was representative of the predictor space defined by the most important predictors of peat depth (Wadoux et al., 2019; Ma et al., 2020). A sample that preserves the properties of the multivariate distribution of predictor and outcome variables is most likely to maintain any complex, non-linear relationships that exist in the population, while avoiding spurious ones (Brus, 2019). Although we implemented the sampling design differently at Skrimfjella (in 2020) than at Ørskogfjellet (in 2023), the overall approach was the same. We chose 105 raster cells at Skrimfjella and 160 raster cells at Ørskogfjellet as our designed samples. A complete description of our sample selection is in Appendix B1.

2.4 Depth measurements

We measured peat depths at Skrimfjella in August 2020 and at Ørskogfjellet in August 2023. We made at least three point measurements of peat depth within each of the raster cells in the designed samples. At both sites, we used manual probing as the primary method of measuring peat depth and ground-penetrating radar (GPR) as a secondary method. That is: peat depth was always measured by probing in our designed samples, and we also used GPR in a set of cells that partly overlapped with the designed samples. We chose this combined approach because probing is a fast and reliable method for point measurements, while GPR can provide higher lateral density of data in the same amount of time (Parry et al., 2014). Probed depths serve to calibrate GPR measurements when calibration by common midpoint survey is not possible – as was the case with our fixed radar antennas – so the methods are complementary. A complete description of our peat depth measurements is in Appendix B2.

At Ørskogfjellet we were also able to use two existing sets of peat depth measurements in addition to our own. We extracted a set of 44 borehole depths (in decimeters) across a 9 ha peatland area, from a paper map made by the Norwegian Soil and Mire Company in 1984. We also used a set of GPR-based depth measurements, commissioned and provided to us by the Norwegian Public Roads Administration. These data were collected in 2020 and 2021 with a dual channel system (70 MHz and 300 MHz; ImpulseRadar AB, Sweden), connected to GNSS with CPOS correction. We used a total of 403440 interpreted and calibrated traces along 7.4 km of transects from this work – discarding data where multiple depths were interpreted for the same location.

All point measurements described above were ultimately aggregated to 10 m resolution by taking the mean of point values
225 within each cell, inversely weighted by their distances to the cell center.

2.5 Predictive models of peat depth

2.5.1 Modeling approach

We used Random Forests (RF) to predict peat depth at both sites. RF is an ensemble machine learning algorithm that builds
many decision trees on bootstrapped samples of the training data, randomly subsets predictors in the trees, and averages the
230 predictions of the trees (Breiman, 2001). We chose RF because it can handle complex interactions between predictors, is
robust to overfitting, and generally shows higher performance in DSM applications than other algorithms (Beguin et al., 2017;
Nussbaum et al., 2018; Lamichhane et al., 2019). It is suited for use on relatively small training datasets and its predictions can
be interrogated to learn about predictor importance (Khaledian and Miller, 2020).

RF by itself is not a spatial model, and it will only predict spatial structure in the peat depth to the degree that spatial structure
235 is captured by predictors. We considered using regression kriging – a hybrid between non-spatial and spatial techniques that
adds to the RF predictions a geostatistically interpolated surface of RF residuals (Hengl et al., 2004). The spatial component
in regression kriging often improves map accuracy (Beguin et al., 2017; Lamichhane et al., 2019; Molla et al., 2023), but it
can do so only if the spatial autocorrelation range in the non-spatial residuals is large compared to distances between samples
and prediction locations (Hengl et al., 2004; Szabó et al., 2019; Takoutsing and Heuvelink, 2022). If the outcome varies at
240 fine scales and the samples are clustered in small parts of the study area, a spatial component will hardly improve overall
map accuracy. We used semivariograms to assess the spatial structure in the residuals of the RF predictions, and found that
(non-spatial) RF was justified at both sites.

We implemented models in the *tidymodels* framework in R (Kuhn and Wickham, 2020), with the *ranger* R package for
RFs (v0.16, Wright and Ziegler, 2017). RFs were fit with 1000 trees, minimum node size of 5, and the number of predictors
245 randomly sampled at each split was the square root of the total number of predictors (*ranger* default). We did not tune these
hyperparameters because RFs are relatively insensitive to tuning (Probst et al., 2019), and because it would require nested
spatial cross-validation to prevent data leakage (Schratz et al., 2019).

2.5.2 Model performance

For both sites, we compared the performance of models with different predictor configurations, where each configuration was
250 one of the seven combinations of the three predictor groups: (1) radiometric (2) terrain and (3) DMK. Comparing the different
configurations allowed us to isolate the added value of each of the predictor groups. The models with only DMK peat depth
class served to provide a fair comparison between the accuracy of the RF models and the existing national map of peat depth,
calibrated on the same data.

We used a spatial cross-validation scheme to evaluate model performance (Wadoux et al., 2021; Meyer and Pebesma,
255 2022). To set the folds we used k-Means Nearest Neighbor Distance Matching (kNNDM), which aims to mimic the spatial

prediction task that is defined as the goal (Linnenbrink et al., 2024). In particular, kNNDM searches for a fold assignment that minimizes the difference between two distributions: nearest neighbor distances between training and test locations in the cross-validation, and nearest neighbor distances between training and prediction locations for the model. That way, the spatial separation between folds is similar to the separation between training and prediction locations – which increases the quality of the map accuracy estimate (Linnenbrink et al., 2024). For spatially clustered training data, this approach strikes a balance between the risk of optimistic metrics from random cross-validation and the risk of pessimistic metrics from other forms of spatial cross-validation (Wadoux et al., 2021). We implemented the kNNDM with the *CAST* R package (v1.0, Meyer et al., 2024), setting prediction locations to all AR5 mire cells in the study area, and choosing a number of folds (k = 5–20) that produced the best match between the two NND distributions. From the cross-validation we quantified mean absolute error (error magnitude, original scale), R^2 (explained variation, standardized scale), and Lin's concordance correlation coefficient (error magnitude and explained variation, standardized scale). We formally assessed the effect of predictor configuration on performance metrics using mixed-effects models to account for the cross-validation fold structure (folds as random effects), and testing pairwise differences between configurations.

DSM products have much more value when their predictions are accompanied by uncertainty estimates, and all DSM should strive to assess uncertainty (Arrouays et al., 2020; Wadoux et al., 2020) and evaluate uncertainty estimates (Heuvelink and Webster, 2022). Therefore, we produced prediction intervals with quantile regression forests (Meinshausen, 2006), and used the same spatial cross-validation to evaluate the prediction interval coverage probability (Shrestha and Solomatine, 2006). The quantile regression forests were trained with the predictor configuration that showed the highest performance at each site, and we extracted 90 % prediction intervals.

2.5.3 Model interpretation

We quantified global variable importance (predictor influence across all locations) and examined partial dependence plots (curves of fitted relationships) for the best-performing predictor configuration at each site. Both are useful for understanding the mechanisms behind the model's predictions and the roles of the predictors in the model.

For both sites, we interpreted a model trained on a non-collinear subset of variables from the best performing predictor configuration – because correlation between predictors degrades variable importance metrics (Strobl et al., 2008; Biau and Scornet, 2016) and can produce misleading visualizations of predictor–outcome relationships (Biecek and Burzykowski, 2021; Dwivedi et al., 2023). Specifically, we eliminated variables from the best performing predictor configuration to obtain a set with no pairwise Pearson correlation coefficient above 0.7 (an arbitrary but conventional threshold for this purpose). Thus, highly correlated sets of variables are represented by a single variable for the purposes of model interpretation.

We calculated variable importance with the *vip* R package (v0.4), by three different methods: *FIRM*, *permutation*, and *Shapley* (Greenwell and Boehmke, 2020). *FIRM* values measure the flatness of the partial dependence plot, *permutation* values measure the decrease in model performance when the predictor is permuted, and *Shapley* values are aggregated from local, game-theoretical measures of variable importance (Greenwell and Boehmke, 2020). Since *FIRM* reflects the flatness of

the partial dependence plot, it captures functional complexity rather than overall predictive impact. *Permutation* values were
290   obtained from ten iterations, with root mean square error as the performance measure.

We calculated partial dependence with the *pdp* R package (v0.8, Greenwell, 2017). For the six most important variables, we plotted partial dependence to show the average effect of the predictor on the outcome and individual conditional expectations to show variation in the effect across observations (Goldstein et al., 2015). Non-parallel individual conditional expectation lines indicate the presence of interactions between predictors.

295   ## 3   Results

Our point measurements of peat depth (all sources) produced aggregated depths for 372 cells or 2.4 % of AR5 mire area at Skrimfjella, and 1878 cells or 1.2 % of AR5 mire area at Ørskogfjellet. Approximately 80 % of these 10 m cells were within AR5 mires, and the remainder in forest, open upland, or farmland (Table 2). DMK peat depth had higher coverage at Ørskogfjellet than at Skrimfjella (80 % versus 27 % of cells not *unknown*), and at Ørskogfjellet the *deep* and *shallow* classes
300   showed a larger difference in measured depth. Overall, mean peat depths were similar at Skrimfjella and Ørskogfjellet (119 cm versus 126 cm) but the distribution was more right-skewed at Ørskogfjellet (Fig. 2).

Table 2. Selected attributes of the peat depth datasets at Skrimfjella and Ørskogfjellet. The first row summarizes all 10 m cells, while subsequent rows stratify the dataset by AR5 land class and DMK peat depth class.

		Skrimfjella			Ørskogfjellet		
		n	%	mean (SD) depth	n	%	mean (SD) depth
all		372	100	119 (83) cm	1878	100	126 (119) cm
AR5	Agricultural				21	1	30 (8) cm
	Forest	52	14	65 (63) cm	272	14	71 (68) cm
	Open upland	19	5	98 (83) cm	134	7	39 (21) cm
	Peatland	301	81	130 (82) cm	1451	77	145 (125) cm
DMK	deep (>100 cm)	94	25	100 (53) cm	659	35	219 (143) cm
	shallow (<100 cm)	6	2	50 (31) cm	838	45	82 (54) cm
	unknown	272	73	127 (90) cm	381	20	60 (62) cm

3.1   Model performance

None of the models were able to predict peat depth across the study areas with high accuracy (Fig. 3). For Skrimfjella, the best model achieved a concordance correlation coefficient of 0.30, an R^2 of 0.34, and a mean absolute error of 60 cm. For
305   Ørskogfjellet, the concordance correlation coefficient was 0.39, R^2 was 0.33, and mean absolute error was 56 cm, so the best model at Ørskogfjellet was slightly more accurate than the best model at Skrimfjella. These values were derived from kNNDM spatial cross-validation with 20 folds at Skrimfjella and 10 folds at Ørskogfjellet.

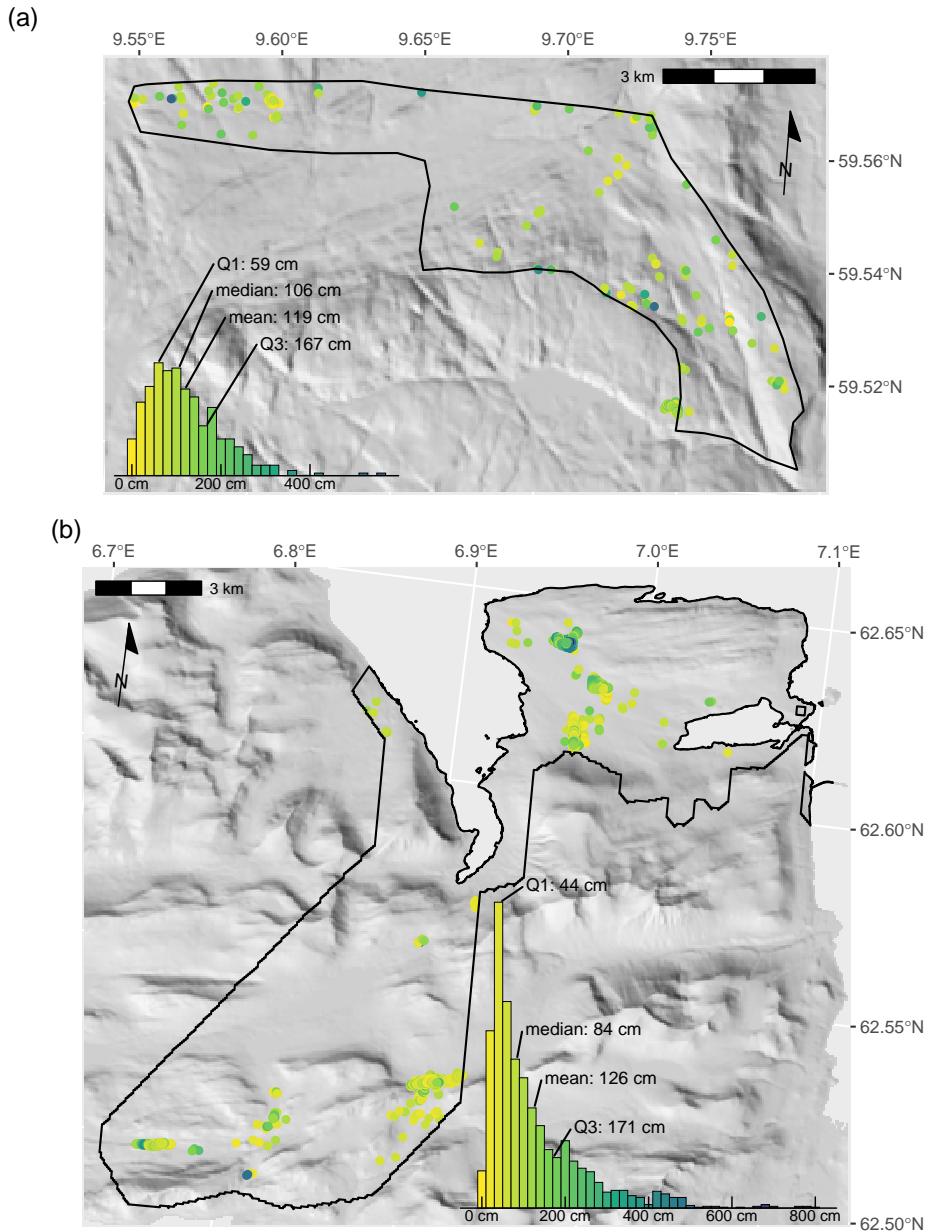



Figure 2. Spatial and statistical distributions of mean peat depth in 10 m cells at Skrimfjella, N = 372 (a) and Ørskogfjellet, N = 1878 (b). Terrain is visualized with a hillshade with slope $45°$ and azimuth $225°$. Terrain data are from the Norwegian Mapping Authority under the Creative Commons Attribution 4.0 International license (https://creativecommons.org/licenses/by/4.0/).

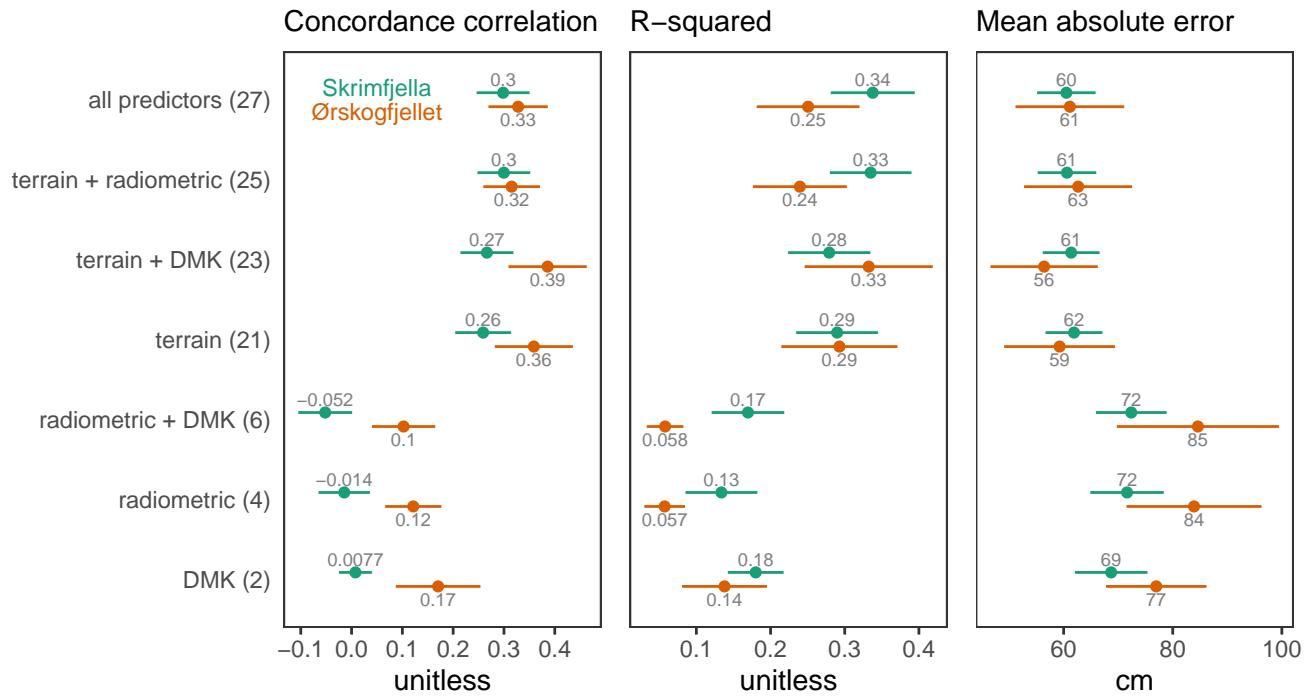



Figure 3. Performance of peatland depth models with different predictor configurations, evaluated via spatial cross-validation. Parentheses denote the number of variables in each predictor configuration, and point estimates are shown +/- their standard error.

For Skrimfjella, the best predictor configuration was *all predictors*, followed closely by *terrain + radiometric* (Fig. 3). The *terrain* configuration outperformed the *radiometric* configuration by 0.27 in concordance correlation coefficient ($p < 0.001$; Appendix Table C1), by 0.16 in R^2 ($p = 0.190$; Appendix Table C2), and by 10 cm in mean absolute error ($p = 0.038$; Appendix Table C3). The *terrain + DMK* configuration outperformed the *radiometric + DMK* configuration by 0.32 in concordance correlation coefficient ($p < 0.001$; Appendix Table C1), by 0.11 in R^2 ($p = 0.605$; Appendix Table C2), and by 11 cm in mean absolute error ($p = 0.011$; Appendix Table C3).

For Ørskogfjellet, the best predictor configuration was *terrain + DMK*, followed by *terrain* (Fig. 3). Adding radiometric predictors to these configurations slightly worsened model performance. The *terrain* configuration outperformed the *radiometric* configuration by 0.24 in concordance correlation coefficient ($p = 0.068$; Appendix Table C4), by 0.24 in R^2 ($p = 0.004$; Appendix Table C5), and by 25 cm in mean absolute error ($p = 0.086$; Appendix Table C6). The *terrain + DMK* configuration outperformed the *radiometric + DMK* configuration by 0.28 in concordance correlation coefficient ($p = 0.015$; Appendix Table C4), by 0.27 in R^2 ($p < 0.001$; Appendix Table C5), and by 28 cm in mean absolute error ($p = 0.031$; Appendix Table C6).

The best models at both sites overpredicted shallow peats and strongly underpredicted very deep peats (Fig. 4). The mean error (bias) of these models was 10 cm at Skrimfjella and -4 cm at Ørskogfjellet. Although the prediction intervals from the

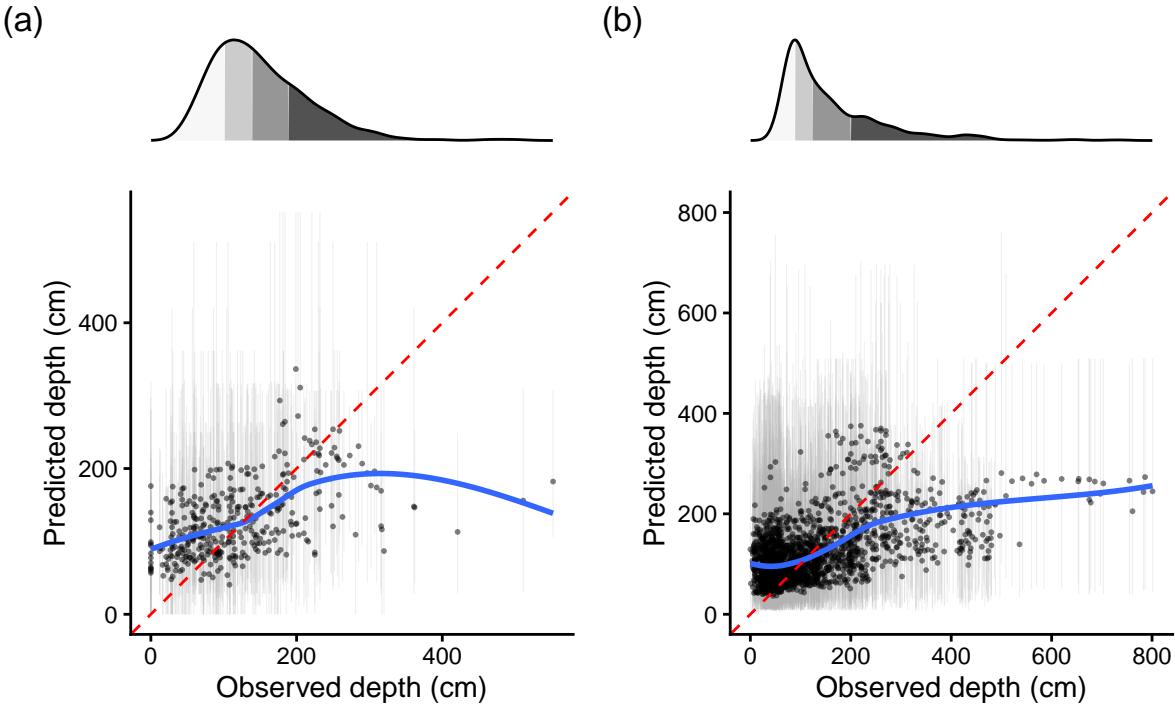
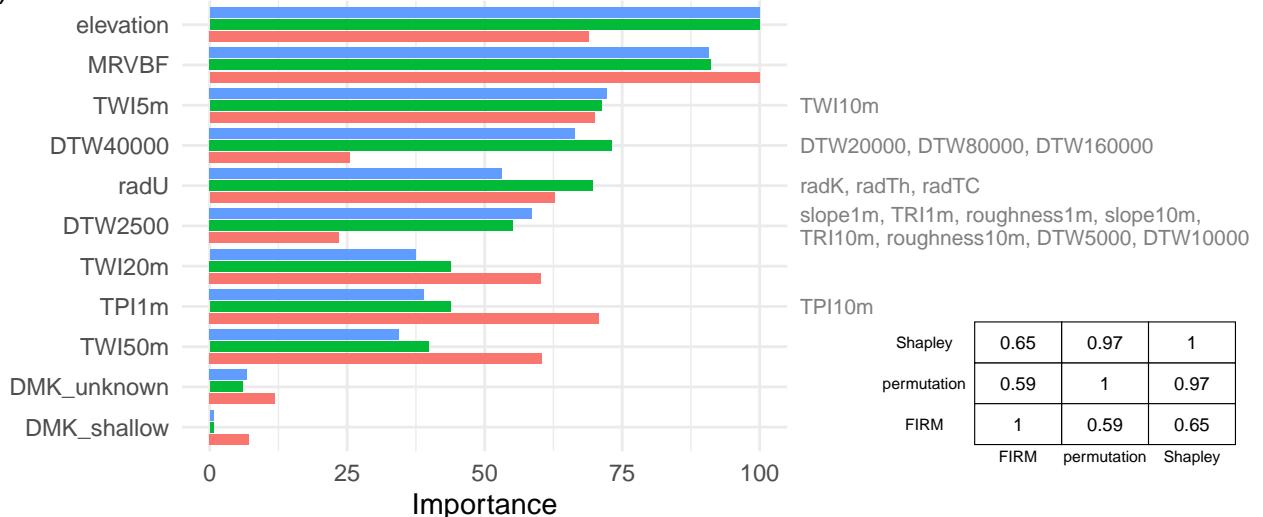



Figure 4. Calibration plots for the best-performing models at Skrimfjella (a) and Ørskogfjellet (b), with predictions and 90 % prediction intervals (grey vertical lines) from spatial cross-validation. Points have transparency to show overlap. Blue lines are local polynomial regressions and the red dashed line in each panel shows the 1:1 line. Marginal distributions (top) are shaded by quartile.

quantile regression forests were wide, they were well calibrated. At Skrimfjella, the prediction interval coverage probability was 92 %, and at Ørskogfjellet it was 84 % (both compared to the target value of 90 %). Observations outside of the prediction intervals showed no obvious spatial pattern.

3.2   Model interpretation

For the purpose of model interpretation, the *all predictors* configuration for Skrimfjella was reduced from 27 variables to 11 non-collinear variables, by removing one of the variables in each highly-correlated pair. Similarly, the *terrain + DMK* configuration for Ørskogfjellet was reduced from 23 variables to 11 non-collinear variables. The *permutation* and *Shapley* methods of variable importance showed high Spearman rank correlation at both sites, while the *FIRM* method ranked variable importance differently (Fig. 5).

3.2.1 Variable importance

At both sites, elevation and Multi-Resolution Valley Bottom Flatness were important predictors (Fig. 5). At Skrimfjella these two predictors were of similar importance, while at Ørskogfjellet elevation had higher predictive impact (*permutation* and *Shapley*) but lower functional complexity (*FIRM*) than Multi-Resolution Valley Bottom Flatness. DMK was also important
335 – the *shallow* class in particular – but only at Ørskogfjellet. Some realizations of the hydrological predictors Topographic Wetness Index and Depth-to-Water showed considerable importance, while others showed little – with no clear consistency between sites. At both sites, the most important hydrological predictors were more important than the simple terrain indices slope, Terrain Ruggedness Index, Topographic Position Index, and roughness. At Skrimfjella, the radiometric predictor uranium ground concentration – which was highly correlated with all other radiometric variables – showed moderate importance.

340 ### 3.2.2 Partial dependence

Many of the most important predictors in the best performing models showed non-monotonic effects on peat depth (Fig. 6). At Ørskogfjellet, for example, increasing elevation was predictive of deeper peat up to about 75 m above sea level, after which a further increase was predictive of shallower peat. At Skrimfjella the partial dependence on elevation had the opposite shape but covers a higher elevation range, with the shallowest peats predicted around 350 m above sea level. TWI50m at Ørskogfjellet
345 and DTW4000 and uranium ground concentration at Skrimfjella were other predictors that showed non-monotonic associations with peat depth. The radiometric predictor in particular displayed an idiosyncratic effect, with a marked dip in predicted depth at intermediate values of uranium ground concentration. On the other hand, the partial effects of some important predictors were more straightforward. The partial dependence on Multi-Resolution Valley Bottom Flatness was quite similar across sites, with the deepest peats predicted in the very flattest valley bottoms. Also, TWI5m and DTW2500 at Skrimfjella showed
350 monotonically positive and negative predictive effects, respectively.

Individual conditional expectation lines indicated some interactions between predictors (Fig. 6). For example, the positive association of depth with elevation at Ørskogfjellet was larger for some observations than others; depth increases ranged 40–100 cm over the same elevation gain. Similarly, individual conditional expectations along uranium ground concentration at Skrimfjella were non-parallel, with some locations showing decreasing peat depth predictions (unlike the average effect).
355 Nevertheless, most individual conditional expectations were approximately parallel, so the average effects of the predictors are mostly representative of their overall effects.

Figure 5. Global variable importance in the best-performing models at Skrimfjella (a) and Ørskogfjellet (b), as measured by three different metrics: Shapley values, permutation importance, and the Feature Importance Ranking Measure (Greenwell and Boehmke, 2020). Variables removed due to collinearity are shown to the right of the variable with which they are most correlated. Spearman rank correlations between the three metrics at each site are displayed in matrices.

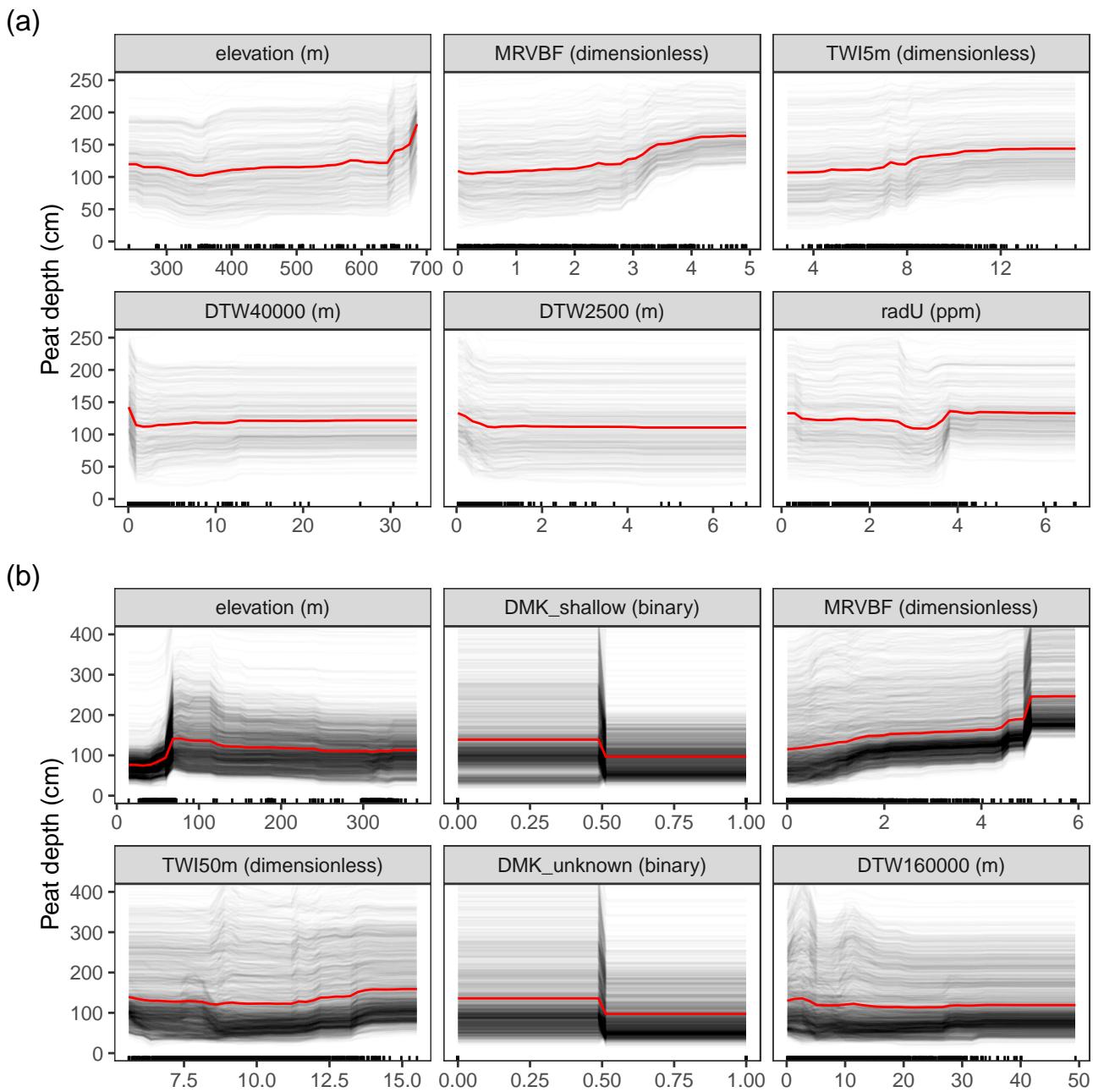



Figure 6. Partial dependence plots of the six most important variables in the best-performing models at Skrimfjella (a) and Ørskogfjellet (b). The average effect of the predictor on the outcome (red line) overlays the individual conditional expectations, which show the variation in the effect across observations (grey lines). The variables' training data distributions are indicated with rug plots along the x-axes.

4 Discussion

4.1 Can we improve Norway's peat depth maps?

4.1.1 Remotely sensed variables are weak predictors of peat depth

360   Our ability to predict peat depth in the study sites based on terrain and radiometric data was limited. Mean absolute errors of 60 and 56 cm at the two sites — relative to mean depths of 119 and 126 cm — illustrate the practical limitations of these maps. Since any given 10 m cell will miss by about 60 cm, applications requiring detailed peat depth in a small area (e.g., < 1 ha) would benefit from measuring depth on the ground rather than relying on the DSM alone.

Although model performance was limited, we improved on the best available map of peat depth (DMK depth class), which
365   is based on field measurements only. This highlights the general value of remotely sensed data, whose complete coverage can improve maps even when their association with the variable of interest is weak (Mulder et al., 2011). Since remotely sensed data are widely available, improvements to soil maps as shown here are readily achievable at low cost (Minasny et al., 2019).

DMK peat depth classes were a worse predictor of peat depth than our models even though we calibrated them with the same data (Fig. 3). If we had taken the DMK peat depth classes at face value and assumed depths according to their class definitions
370   (<1 m, >1 m), they would have performed worse and the advantage of the DSM would be greater. The advantage of the DSM was not large in absolute terms (9 cm and 21 cm improvements in mean absolute error), but it explained much more of the variation in depth (improvements in R^2 of 0.16 and 0.19). We attribute this result to the poor spatial and thematic resolution of DMK peat depth, which precludes a robust correlation with peat depths.

The accuracy of our mapping could have been improved with a spatially explicit approach like regression kriging. However,
375   residuals of the RF predictions showed weak spatial structure at Skrimfjella, and only up to a range of 150 m at Ørskogfjellet. Therefore, the improvement from regression kriging would be small overall and limited to parts of the maps close to measurements.

4.1.2 Similar error but less explanatory power in Norwegian peatlands

Compared to other studies using terrain and radiometric data to predict peat depth, our models explained less variability in peat depth (R^2) but generally produced better or comparable error magnitude (mean absolute error or root mean square error,
380   Wadoux et al., 2022). It is important to keep in mind that differences in peat depth distributions, spatial scales, and evaluation methods make direct performance comparisons precarious. For example, R^2 is sensitive to high leverage, extreme values, so it will evaluate a right-skewed distribution differently than a symmetrical distribution. We evaluated our models with respect to the explicit purpose of creating peat depth maps across the study areas, but not all studies tailored evaluation to match an explicitly formulated problem (Milà et al., 2022).

385   Gatis et al. (2019) used similar predictors and the same spatial grain, finding a much stronger relationship between predicted and observed peat depth ($R^2 = 0.68$). Although their random evaluation data partition could make performance estimates too optimistic (Roberts et al., 2017; Wadoux et al., 2021), the confounding effect of spatial structure is probably small because they used linear regression and few predictors, so their model had limited opportunity to overfit to the spatial structure in peat

depth. The most salient difference in Gatis et al. (2019) compared to our study is the character of the study area. They study a flatter area with a higher proportion of peatland cover, and their peatland is primarily blanket bog. The peat surface of a blanket bog is tied more closely to its underlying topography than the peat surface of raised bogs or fens (Lindsay, 2016b), which may make blanket bog depth easier to predict. The predominance of fens and smaller peatland extent may have contributed to worse performance in our study.

Marchant (2021) examined a subset of the area studied in Gatis et al. (2019) at 100 m resolution, using splines to relax linearity between radiometry/terrain and peat depth. He found that potassium ground concentration alone predicted peat depth with much higher concordance than our models (concordance correlation coefficient = 0.76) and that elevation alone produced comparable performance (concordance correlation coefficient = 0.27). The root mean square error from these univariate models was 46–68 cm (cf. 78 and 75 cm for Skrimfjella and Ørskogfjellet).

Koganti et al. (2023) had a peat depth distribution and predictors similar to ours, but at much smaller spatial grain and extent. Their training and validation points are closer than the range of spatial autocorrelation in peat depth, so our results are best compared to their non-spatial models. They accounted for more variability in peat depth (adjusted $R^2 = 0.71$) but had larger errors (RSME = 110 cm). Koganti et al.'s (2023) linear regression models produced negative predictions, and it is unclear whether the values quoted above include these. If we disregard the negative predictions, their model showed the same pattern as ours in overpredicting shallow peats and underpredicting deep peats, although their underprediction was less severe. An important difference between Koganti et al. (2023) and our study (besides spatial scale) is that they measured radiometrics on the ground, rather than using airborne survey data. Thus, the footprint of their detector was smaller and they captured variation in radiation at finer scales.

4.1.3   Asymmetries in depth predictions for land use planning and carbon accounting

Our models erred most for the deepest peats (Fig. 4). Where overprediction occurred, the error was smaller. This is not unexpected for the right-skewed distributions of peat depth, but it has implications for potential users of the maps. For example, it makes the maps more suited for "red-lighting" than "green-lighting" peatland conversion (assuming depth is the determinative factor). Identifying the deepest peats will require additional field work in candidate areas – where candidate areas could be defined by some upper quantile of predicted depth.

4.2   Which variables predict peat depth?

4.2.1   Airborne radiometrics do not predict Norwegian peat depth

Radiometric data had little to no predictive value at either site (poor performance of *radiometric* configuration), although they did contribute to the best model at Skrimfjella (marginal improvement in *all predictors* configuration compared to *terrain + DMK*). These results contrast with earlier studies that found that radiometrics were useful predictors of peat depth (Keaney et al., 2013; Gatis et al., 2019; Koganti et al., 2023; Pohjankukka et al., 2025). The bedrock is more homogeneous at Ørskogfjellet than at Skrimfjella, so uneven radiogenesis is not a good explanation for the higher predictive value of radiometric

variables at Skrimfjella (Beamish, 2014; Reinhardt and Herrmann, 2019). All four radiometric variables were highly correlated within the peatland parts of our study sites, so there could be no large differences in their predictive value. This contrasts with Koganti et al. (2023), who found that radiometric total count and potassium ground concentration were better predictors than thorium or uranium ground concentration.

425   We suspect that the primary reason for the poor predictive value of the radiometric data was the large footprint of the detector in the airborne survey. With an average flight altitude of 75 m, less than half of the radiation reaching the detector comes from inside the 100 m diameter circle directly below it (Beamish, 2016; Beamish and White, 2024). The rest of the measured activity integrates a much wider area. For comparison, empirical variograms of peat depth at Skrimfjella and Ørskogfjellet showed no spatial autocorrelation beyond 50 and 75 m (among GPR data) or 110 and 230 m (among 10 m cells). Basically, the
430   airborne radiometric data will not capture large variation over short (< 100 m) distances; the instrument's field of view has a large smoothing effect on the data (Beamish, 2016; Reinhardt and Herrmann, 2019). Different landforms and the changes they cause in the geometry between the radioactive source and the detector can also distort airborne measurements (Reinhardt and Herrmann, 2019).

    Studies comparing airborne and ground radiometric surveys confirm that they are poorly correlated in low-activity areas
435   like peatlands (Kock and Samuelsson, 2011; Karjalainen et al., 2025). Karjalainen et al. (2025) found that ground-based measurements predicted peat depth better than airborne measurements. Nevertheless, a large radiometric footprint did not prevent strong associations with peat depth in Gatis et al. (2019) and Marchant (2021), perhaps because the extensive blanket bog landscape in these studies has more gradual changes in depth (Lindsay, 1995). We are unsure whether short-range depth changes explain the weak associations that Siemon et al. (2020) found in a large raised bog.

440   Another possible reason for the poor predictive value of the radiometric data could be that other physical parameters influencing the amount of intercepted radiation varied too much within sites. Initial source strength, soil moisture, bulk density, and porosity all affect the amount of radiation that reaches the detector (Beamish, 2013; Reinhardt and Herrmann, 2019). Therefore, variation in these parameters could have masked the relationship between peat depth and radiometric data. This makes physical modeling of peat depth from radiometric data an undetermined problem. We chose the Ørskogfjellet site in
445   part because it has a relatively homogeneous bedrock, which should reduce the variation in initial source strength. Soil moisture, bulk density, and porosity, however, are not easily measured across landscape scales and were assumed to be homogeneous. Uneven snow cover and air moisture during the Ørskogfjellet radiometric survey may also have masked the soil signal in these data, as Ofstad (2015) reports large variation in weather conditions. If maps of these other physical parameters at the time of the radiometric survey were available and included in the model, the predictive value of radiometric data might improve, but
450   this is not a practical solution for digital soil mapping of peat depth.

    We do not believe that the poor predictive value of the radiometric data in this study was caused primarily by full attenuation of radioactivity. The RF algorithm's flexibility means that radiometrics could be used for shallower peats if they provided predictive value for that part of the depth distribution, but there is no evidence of that in our results. In the partial dependence plot of uranium ground concentration at Skrimfjella, the expected negative relationship between depth and uranium ground

concentration cannot be found by ignoring the left (highly attenuated) side of the distribution (Fig. 6). More than a quarter of the peats in each of our study areas were less than 60 cm deep, and full attenuation is unlikely for these (Beamish, 2013).

Although we do not believe full attenuation is the primary reason for poor performance in our study, it may limit peat depth mapping under other circumstances. First-principle calculations suggest that radiation should be 90 % attenuated after about 50–60 cm of typical, wet peat or 85 cm of unnaturally dry peat (Beamish, 2013; Beamish and White, 2024), and some field tests support these values (Billen et al., 2015). It is remarkable that particular studies detected radiation differences up to several meters deep (Gatis et al., 2019; Koganti et al., 2023), but these may be the exceptions rather than the rule. Perhaps relatively deeper water tables in these study sites (blanket bog, drained fen, Price et al., 2016) contributed to better penetration.

4.2.2 Terrain-based variables can predict peat depth

At both our sites, LiDAR-derived terrain variables predicted peat depth much better than radiometric variables (Fig. 3). This is consistent with the findings of Pohjankukka et al. (2025), who mapped peat depth classes at 50 m resolution across Finland.

At both our sites, elevation was the most important predictor, and peat depth showed non-monotonic responses to changes in elevation (Figs. 5, 6). We believe that the idiosyncratic elevational relationships we detected are mostly not generalizable beyond the study areas, because we see no simple mechanism (e.g., via climate) to explain the observed patterns. For example, the increase in peat depth from 350 to 700 m.a.s.l. at Skrimfjella is opposite to the general pattern of deeper peats in lowland than upland Norway (Lyngstad et al., 2023). Moreover, elevation at Ørskogfjellet seems to interact with other variables (non-parallel individual conditional expectation lines), complicating its interpretation. Relationships between elevation and peat depth have previously shown opposite shapes in different areas (Finlayson et al., 2021). Nonetheless, a relationship that is not generalizable beyond the mapping area is still useful for DSM, as long as it is evaluated to demonstrate its robustness for the predictive task (e.g., through kNNDM spatial cross-validation).

One interesting feature of the elevational relationships we found may be generalizable: a steep increase in peat depth near the marine limit after the last ice age. At Ørskogfjellet, the marine limit is about 75 meters above today's sea level (Geological Survey of Norway, Høgaas et al., 2012), where the partial dependence plot of elevation shows a sharp increase in peat depth. In areas under the marine limit there has been less time for peat accumulation since the ice sheets retreated, and it is plausible that this makes peats there shallower, all else being equal. This hypothesis is supported by similar findings at a coarser scale in the Hudson Bay Lowlands, where a strong positive relationship between peat depth and distance from the coast can be explained by isostatic uplift and time since peat initiation (Li et al., 2025). The same has been reported for Finland (Pohjankukka et al., 2025). We cannot evaluate this effect at Skrimfjella, where the marine limit is below our study area (at 175 m.a.s.l.).

Another influential terrain-based predictor was Multi-Resolution Valley Bottom Flatness (Fig. 5). Unlike elevation, it showed a monotonic effect on peat depth: greater valley bottom flatness was always associated with increases in peat depth (Fig. 6). Delineating a valley bottom involves ambiguity, but the Multi-Resolution Valley Bottom Flatness index is a pragmatic approach that considers a location a valley bottom if it is sufficiently low and flat at a particular scale (Gallant and Dowling, 2003). The multiscale nature of the index allows small elevated but flat areas (including saddles) to be characterized as having high valley bottom flatness (Gallant and Dowling, 2003). Our results suggest that Multi-Resolution Valley Bottom Flatness is a robust

indicator of high water tables (and peat accumulation) over millennial time scales, corroborating other studies (Rudiyanto et al., 2018; Deragon et al., 2023).

Other terrain-derived predictors with predictive value in our study are hydrological (Topographic Wetness Index and Depth-to-Water). Notably, slope was inferior to (at Ørskogfjellet) or highly correlated with (at Skrimfjella) these hydrological indices. Mappers of peat depth should not assume that slope is the best terrain-derived predictor, despite its prevalence in the literature (Pohjankukka et al., 2025). Wetter locations (high Topographic Wetness Index and low Depth-to-Water) were generally associated with deeper peat, but these relationships were not as strong or consistent as with Multi-Resolution Valley Bottom Flatness (Figs. 5, 6). The optimal scale for Topographic Wetness Index and Depth-to-Water varied, and likely depends on both the dominant peat formation processes and the typical size of peatland features in a landscape. Including multiple scales of these variables allows the model to capture different hydrological mechanisms or patterns operating at different spatial scales.

4.2.3 Legacy depth maps have inconsistent predictive value

DMK peat depth class proved an inconsistent predictor of peat depth. At Skrimfjella, it barely improved model performance (Fig. 3). At Ørskogfjellet, it increased performance more, and both indicator variables (*shallow* and *unknown*) were among the most important in the model (Fig. 5). We suspect the discrepancy between sites is due to different levels of effort and coverage during the historical surveys; more lowland peatland near agriculture at Ørskogfjellet may have caused more purposeful surveying. This is evidenced by the fact that 73 % of the cells measured at Skrimfjella had unknown depth in DMK, compared to 20 % at Ørskogfjellet. Interactions between DMK and other variables also underline the inconsistency of DMK depth maps, even within a site. For example, *shallow* peat classification in DMK sometimes increased rather than decreased predicted depth at Ørskogfjellet.

4.3 Implications for digital soil mapping of peat depth

The performance gap between the best models and the DMK models shows that DSM of peat depth has value in Norway. Where calibrating measurements are available, better maps than DMK peat depth class can be produced at low cost. Moreover, DSM can make the production of maps transparent, reproducible, and updatable. We can apply the same approach across different areas and make maps with full spatial coverage, continuous values, and validated uncertainty. The large difference in coverage and quality of DMK peat depth at Skrimfjella versus Ørskogfjellet underlines these advantages.

4.3.1 Peat depth measurements should be organized

High-quality DTMs are available for mainland Norway, which makes peat depth measurements the critical training data needed for DSM. Better infrastructure to make depth measurements findable, accessible, interoperable, and reusable would help DSM and other applications. Geoportal access and data exchange standards (like Natural England's for peat surveys, 2023) are important. Peat depth is quick and easy to measure, so integrating its measurement into existing national field programs, like Norway's spatially representative nature monitoring or national forest inventory, would be helpful (although not sufficient for

520 regional or local mapping). Low-altitude, drone-mounted GPR may prove an efficient approach for collecting many, accurate depth data in a landscape, by combining the advantages of airborne deployment and active sensing (Pelletier et al., 1991; Ruols et al., 2023).

4.3.2 Spatial scale affects model performance and utility

Peat depths typically vary over short distances (e.g., < 1 m), so mapping at 10 m resolution implies that the map will compress
525 much of the fine-scale variation. Terrain–depth relationships might be stronger at finer resolution than 10 m, especially considering the hummock–hollow microtopography of many peatlands (Rydin et al., 1999; Lindsay, 2010). Therefore, it may be advantageous to model peat depth at 1 m resolution – even if land use planning and carbon accounting do not operate at such fine scale.

To define a spatial extent for a particular DSM, kNNDM can be used to investigate how far the map can extend beyond a
530 set of point measurements; if the sample-to-prediction nearest neighbor distribution cannot be simulated by any set of cross-validation folds, then the extent is too expansive (Meyer and Pebesma, 2022; Linnenbrink et al., 2024). However, the tradeoff between multiple small-extent DSM and fewer large-extent DSM needs research. Bohn and Miller (2024) advocate for bottom-up stitching of local DSM, and for peat depth we assert that these should at least stay within peatland regions (or 'supertopes'), where the composition of mesotopes is similar – e.g., regions dominated by raised bogs versus regions dominated by sloping
535 fens (Moen, 1999; Joosten and Clarke, 2002). Depth varies systematically between bogs and fens (Lindsay, 2016b), between peats formed by terrestrialization versus paludification (Buffam et al., 2010), and probably along other axes of peatland typology. Therefore, DSM is more likely to uncover consistent predictor–depth relationships within peatland regions than across them.

4.3.3 Machine learning approaches can build on success

540 The DSM literature and our results support using flexible machine learning algorithms like RF to predict peat depth. RF avoids negative predictions (c.f. Koganti et al., 2023) and produces good uncertainty estimates (our study, Vaysse and Lagacherie, 2017; Takoutsing and Heuvelink, 2022). As depth data become more abundant, pixel-based learners may be surpassed by deep learning approaches like convolutional neural networks. The success of Multi-Resolution Valley Bottom Flatness as a predictor of peat depth demonstrates that multiscale spatial patterns matter for peat depth, and convolutional neural networks
545 are designed to learn such patterns (Borowiec et al., 2022). The kind of relationship described in Buffam et al. (2010), where peat depth in basins related to terrain slope at the basin edge, is also something a convolutional neural network could learn. However, since this approach is data-hungry, we should build soil knowledge into the DSM where we can (Minasny et al., 2024b).

4.3.4 Peat extent and depth should be mapped together

Finally, we would like to highlight briefly the need for peatland extent mapping and peat depth mapping to be better integrated. Since peatland extent is defined by non-zero peat depth (the specific threshold varies by definition, Minasny et al., 2024a), the lateral and vertical dimensions are fundamentally linked. The goal, therefore, should be a unified prediction framework for extent and depth. Research is needed to determine whether it is better to parameterize a single model of peat depth across a full landscape, or to break down the problem into a hurdle model by classifying zero depth and then regressing non-zero depths. Though peatland definitions may encourage reducing continuous depth predictions to arbitrary classes, we caution against this practice (as in Ivanovs et al., 2024; Karjalainen et al., 2025).

5 Conclusions

This study demonstrates that digital soil mapping at 10 m resolution can improve upon existing peat depth maps in Norway, though the strength of the relationship between available predictors and peat depth remains limited. Our findings show that terrain-derived variables, particularly elevation and Multi-Resolution Valley Bottom Flatness, provide predictive value for peat depth mapping within peatland extents. In contrast, airborne radiometric data showed little to no predictive value at either of two study sites, possibly because of the large footprint of airborne spectrometers relative to the fine-scale variation in peat depths.

The best models achieved mean absolute errors of 56–60 cm against mean depths of 119–126 cm, and explained approximately one-third of the variation in peat depth across landscapes with aggregate peatland areas of 1.5–15.3 km^2. Though field measurements remain necessary for local, detailed assessments of peat depth, digital soil maps at 10 m resolution can provide valuable information for landscape-scale planning and regional carbon assessments. The models' tendency to underpredict the deepest peats has important practical implications, making them more suitable for precautionary screening than comprehensive coverage. As Norway and other nations pursue nature-based climate solutions, these findings highlight both the potential and limitations of remote sensing for peatland carbon mapping.

Appendix A

Table A1. Attributes of the radiometric surveys, as reported in Baranwal et al., (2013) and Ofstad (2015).

	Skrimfjella	Ørskogfjellet
Survey period	2008–2011	2014.12-2015.01
Average flight altitude (m)	75	80
Average flight speed ($\mathrm{km\,h^{-1}}$)	108	88
Flight line spacing (m)	200	200

Appendix B

B1 Peat depth sample selection

B1.1 Skrimfjella

575 At Skrimfjella, we used the *eSample* function in the *iSDM* R package (v1.0) to stratify our sample across elevation, slope, and potassium concentration. This function defines the environmental space as a two-dimensional convex hull around the PCA-ordinated data, then creates a regular grid across that space, and lastly finds for each grid cell the datum that is nearest (Hattab et al., 2017). We set a target sample size of 100, excluded the top and bottom percentile from the convex hull, and *eSample* returned 105 raster cells.

580 #### B1.2 Ørskogfjellet

At Ørskogfjellet, we first determined a minimal sample size that would adequately capture the slope and radiometric properties (potassium, thorium, uranium, and total count) of the entire AR5 mire area (Saurette et al., 2023). Specifically, we identified an elbow point in a curve of similarity between sample and population (Malone et al., 2019). For a sequence of sample sizes (50–500) (ten replicates each, drawn by conditioned latin hypercube sampling, Minasny and McBratney, 2006; Roudier, 585 2011), we calculated the mean Kullback–Leibler divergence between sample and population distributions (Malone et al., 2019; Saurette et al., 2023). Then we fitted an asymptotic regression of mean divergence on sample size, and found that the curve reached 95 % of the fitted asymptote at a sample size of 160.

To choose 160 locations, we performed feature space coverage sampling, implemented using the *kmeans* function in base R and the *rdist* function in the *fields* package (v14.2). Feature space coverage sampling chooses locations that are closest to 590 cluster centers in standardized predictor space (Brus, 2019). This approach has been found to produce higher accuracy in RFs than conditioned latin hypercube sampling (Wadoux et al., 2019; Ma et al., 2020). Feature space coverage sampling works best when all dimensions are important predictors of the outcome (Wadoux et al., 2019), and we used the same five predictors that we used to choose sample size: slope and four radiometrics.

We adjusted the feature space coverage sampling to ensure that locations were accessible within time constraints, and 595 assessed how this changed our sample from an ideal feature space coverage sample. Adjusting for accessibility is justified because the smaller sample size that would result if accessibility were ignored can degrade model accuracy as much or more as deviations from ideal sampling designs (Wadoux et al., 2019; Ma et al., 2020). To adjust, we first restricted the sampling population to AR5 mire areas that were within an arbitrary cost distance of publicly accessible roads. Cost distance was calculated using GRASS's *r.walk* function, with friction costs defined by AR5 land classes (GRASS Development Team, 2022). 600 After creating a feature space coverage sample with this restriction, we also inspected a map of the sample and substituted 16 inaccessible locations with accessible locations from the same or a nearby cluster. Our two accessibility adjustments increased the distance in standardized predictor space between sample locations and cluster centers by 78 % (compared to the ideal

sample), but distance in our sample was still only 46 % of the mean distance to cluster centers – i.e., accessibility did not force locations far from cluster centers relative to the size of the clusters.

B2   Depth measurements

B2.1   Peat probing

We navigated to the centers of the raster cells in our samples using handheld (Skrimfjella) or real-time kinematic (Ørskogfjellet) global navigation satellite system (GNSS) receivers. We dampened the effect of outlying measurements by probing three times at each location (Parry et al., 2014), at the vertices of a centered triangle with 2 m (Skrimfjella) or 4.5 m (Ørskogfjellet) sides. We used changes in resistance to indicate the base of the peat column. Probe locations were adjusted up to 20 cm if the base of the peat column seemed to be blocked by an obvious artifact, like a buried rock. Where the peat column was deeper than the extendable probe could be manually inserted and extracted by a pair of operators, we recorded a right-censored result (one at Skrimfjella, five at Ørskogfjellet).

B2.2   Ground-penetrating radar

We performed GPR surveys in three subjectively chosen peatlands at Skrimfjella and in areas with a high density of sampled raster cells at Ørskogfjellet. We used the Malå ProEx GPR system (Guideline Geo AB, Sweden) with a GNSS-enabled control unit connected to a 500 MHz shielded antenna mounted in a plastic sledge (transmitter–receiver separation 0.18 m, trace frequency 10 Hz). For some transects at Ørskogfjellet we substituted a 100 MHz Malå rough terrain antenna (transmitter–receiver separation 2.2 m, trace frequency 5 Hz), because the lower frequency antenna gives greater penetration depth. In all cases the system was towed by a walking GPR operator.

GPR traces were recorded along zigzag (Skrimfjella) or snaking (Ørskogfjellet) transects. At Ørskogfjellet, transects were predetermined to pass through the centers of sampled raster cells, and we marked these precisely with flags to guide the GPR operator. A GPR records the time taken for a radio wave to travel from the transmitter to a reflector and back to the receiver, and the velocity of the wave varies with properties of the peat column. Therefore, wave velocity has to be calibrated to convert travel time to peat depth, and we probed peat depth along the transects.

We processed the GPR data with Reflex2DQuick (v.3.0; Skrimfjella) or REFLEXW (v.8.5; Ørskogfjellet) software (Sandmeier Scientific Software, Germany). We applied a time-zero correction, a dewow filter, and a gain filter based on observed energy decay. With Ørskogfjellet data, we also applied a bandpass filter and a dynamic correction that accounts for the non-vertical wave path between offset transmitter and receiver antennas. The latter is important for the rough terrain antenna, where the antenna separation is comparable to typical peat depths. From the processed radargrams, we picked travel times from strong reflectors that we interpreted as the base of the peat column.

We used picks near probed depths to calibrate wave speed velocity – separately for each site. Calibration data were created by matching marked trace locations to a corresponding depth probe (Skrimfjella), or by a spatial join that identified interpreted traces and depth probes within 2 m of each other (Ørskogfjellet). We had sufficient calibration point density to avoid bias

in wave velocity as a major source of error (Rosa et al., 2009): 46 calibration points along 3.5 km of interpretable traces at Skrimfjella, and 78 along 7.8 km at Ørskogfjellet. We fitted site-specific linear regressions of probed depth on picked travel time, with the intercept fixed at zero, to estimate wave velocities. Notwithstanding a few outlying points, our regressions showed good fits and the resulting velocities are within the range reported for peat (Parsekian et al., 2012): $0.0387\,\mathrm{m\,ns^{-1}}$, $R^2 = 0.874$ at Skrimfjella and $0.0427\,\mathrm{m\,ns^{-1}}$, $R^2 = 0.946$ at Ørskogfjellet (Fig. B1). Finally, we used these two wave velocities to convert the travel times of all picks to calibrated peat depths. In total, the GPR surveys produced 48579 point measurements of peat depth at Skrimfjella and 32653 at Ørskogfjellet.

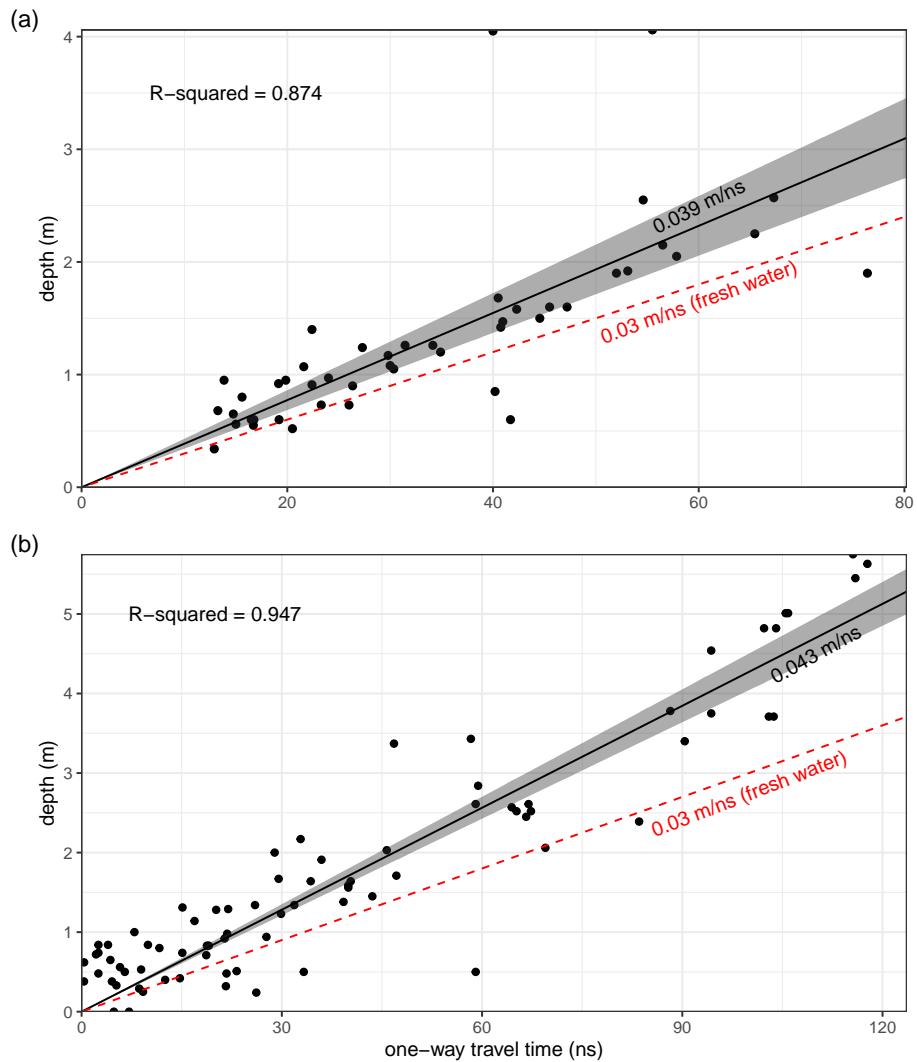



Figure B1. Calibration of GPR wave velocity at Skrimfjella (a) and Ørskogfjellet (b) by regression of probed depth against wave travel time.

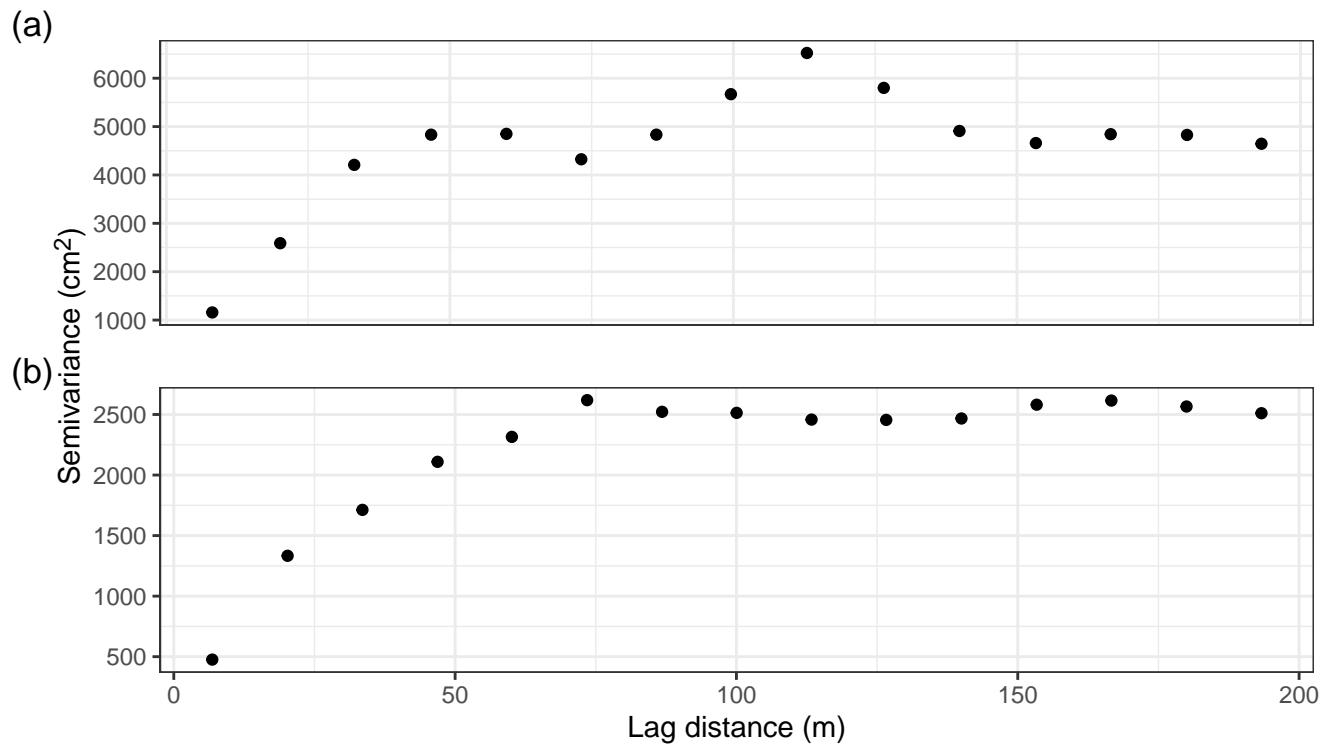



Figure B2. Empirical semivariograms of peat depth point measurements at Skrimfjella (a) and Ørskogfjellet (b).

Appendix C

Table C1. Pairwise comparisons of the concordance correlation coefficient among models for Skrimfjella, based on 10-fold cross-validation. The comparisons use estimated marginal means with Tukey HSD correction and Kenward-Roger degrees of freedom approximation (`emmeans` package) on mixed-effects models (`lme4` package) to account for the cross-validation fold structure.

Comparison	Estimate	Std.error	Df	Statistic	Adj.p.value	Significance
DMK - Radiometric	0.02	0.05	114	0.44	0.999	
DMK - RadiometricDMK	0.06	0.05	114	1.19	0.896	
DMK - RadiometricTerrain	-0.29	0.05	114	-5.84	< 0.001	***
DMK - RadiometricTerrainDMK	-0.29	0.05	114	-5.81	< 0.001	***
DMK - Terrain	-0.25	0.05	114	-5.03	< 0.001	***
DMK - TerrainDMK	-0.26	0.05	114	-5.17	< 0.001	***
Radiometric - RadiometricDMK	0.04	0.05	114	0.75	0.989	
Radiometric - RadiometricTerrain	-0.31	0.05	114	-6.28	< 0.001	***
Radiometric - RadiometricTerrainDMK	-0.31	0.05	114	-6.25	< 0.001	***
Radiometric - Terrain	-0.27	0.05	114	-5.47	< 0.001	***
Radiometric - TerrainDMK	-0.28	0.05	114	-5.62	< 0.001	***
RadiometricDMK - RadiometricTerrain	-0.35	0.05	114	-7.03	< 0.001	***
RadiometricDMK - RadiometricTerrainDMK	-0.35	0.05	114	-7.00	< 0.001	***
RadiometricDMK - Terrain	-0.31	0.05	114	-6.22	< 0.001	***
RadiometricDMK - TerrainDMK	-0.32	0.05	114	-6.37	< 0.001	***
RadiometricTerrain - RadiometricTerrainDMK	0.00	0.05	114	0.03	1	
RadiometricTerrain - Terrain	0.04	0.05	114	0.81	0.983	
RadiometricTerrain - TerrainDMK	0.03	0.05	114	0.67	0.994	
RadiometricTerrainDMK - Terrain	0.04	0.05	114	0.79	0.986	
RadiometricTerrainDMK - TerrainDMK	0.03	0.05	114	0.64	0.995	
Terrain - TerrainDMK	-0.01	0.05	114	-0.15	1	

Table C2. Pairwise comparisons of R^2 among models for Skrimfjella, based on 10-fold cross-validation. The comparisons use estimated marginal means with Tukey HSD correction and Kenward-Roger degrees of freedom approximation (`emmeans` package) on mixed-effects models (`lme4` package) to account for the cross-validation fold structure.

Comparison	Estimate	Std.error	Df	Statistic	Adj.p.value	Significance
DMK - Radiometric	0.04	0.08	108.23	0.55	0.998	
DMK - RadiometricDMK	0.01	0.08	108.23	0.08	1.000	
DMK - RadiometricTerrain	-0.16	0.08	108.23	-2.07	0.378	
DMK - RadiometricTerrainDMK	-0.16	0.08	108.23	-2.11	0.357	
DMK - Terrain	-0.11	0.08	108.23	-1.48	0.755	
DMK - TerrainDMK	-0.10	0.08	108.23	-1.34	0.830	
Radiometric - RadiometricDMK	-0.04	0.06	105.01	-0.56	0.998	
Radiometric - RadiometricTerrain	-0.20	0.06	105.01	-3.16	0.033	*
Radiometric - RadiometricTerrainDMK	-0.20	0.06	105.01	-3.20	0.029	*
Radiometric - Terrain	-0.16	0.06	105.01	-2.45	0.190	
Radiometric - TerrainDMK	-0.15	0.06	105.01	-2.28	0.263	
RadiometricDMK - RadiometricTerrain	-0.17	0.06	105.01	-2.60	0.137	
RadiometricDMK - RadiometricTerrainDMK	-0.17	0.06	105.01	-2.64	0.124	
RadiometricDMK - Terrain	-0.12	0.06	105.01	-1.89	0.494	
RadiometricDMK - TerrainDMK	-0.11	0.06	105.01	-1.72	0.605	
RadiometricTerrain - RadiometricTerrainDMK	0.00	0.06	105.01	-0.04	1.000	
RadiometricTerrain - Terrain	0.05	0.06	105.01	0.71	0.992	
RadiometricTerrain - TerrainDMK	0.06	0.06	105.01	0.88	0.975	
RadiometricTerrainDMK - Terrain	0.05	0.06	105.01	0.75	0.989	
RadiometricTerrainDMK - TerrainDMK	0.06	0.06	105.01	0.92	0.968	
Terrain - TerrainDMK	0.01	0.06	105.01	0.17	1.000	

Table C3. Pairwise comparisons of mean absolute error among models for Skrimfjella, based on 10-fold cross-validation. The comparisons use estimated marginal means with Tukey HSD correction and Kenward-Roger degrees of freedom approximation (`emmeans` package) on mixed-effects models (`lme4` package) to account for the cross-validation fold structure.

Comparison	Estimate	Std.error	Df	Statistic	Adj.p.value	Significance
DMK - Radiometric	-2.91	3.13	114	-0.93	0.967	
DMK - RadiometricDMK	-3.69	3.13	114	-1.18	0.902	
DMK - RadiometricTerrain	8.10	3.13	114	2.58	0.140	
DMK - RadiometricTerrainDMK	8.21	3.13	114	2.62	0.130	
DMK - Terrain	6.81	3.13	114	2.17	0.319	
DMK - TerrainDMK	7.33	3.13	114	2.34	0.235	
Radiometric - RadiometricDMK	-0.77	3.13	114	-0.25	1.000	
Radiometric - RadiometricTerrain	11.02	3.13	114	3.51	0.011	*
Radiometric - RadiometricTerrainDMK	11.13	3.13	114	3.55	0.010	*
Radiometric - Terrain	9.72	3.13	114	3.10	0.038	*
Radiometric - TerrainDMK	10.24	3.13	114	3.27	0.024	*
RadiometricDMK - RadiometricTerrain	11.79	3.13	114	3.76	0.005	**
RadiometricDMK - RadiometricTerrainDMK	11.90	3.13	114	3.80	0.004	**
RadiometricDMK - Terrain	10.50	3.13	114	3.35	0.018	*
RadiometricDMK - TerrainDMK	11.01	3.13	114	3.51	0.011	*
RadiometricTerrain - RadiometricTerrainDMK	0.11	3.13	114	0.04	1.000	
RadiometricTerrain - Terrain	-1.29	3.13	114	-0.41	1.000	
RadiometricTerrain - TerrainDMK	-0.77	3.13	114	-0.25	1.000	
RadiometricTerrainDMK - Terrain	-1.40	3.13	114	-0.45	0.999	
RadiometricTerrainDMK - TerrainDMK	-0.88	3.13	114	-0.28	1.000	
Terrain - TerrainDMK	0.52	3.13	114	0.17	1.000	

Table C4. Pairwise comparisons of the concordance correlation coefficient among models for Ørskogfjellet, based on 10-fold cross-validation. The comparisons use estimated marginal means with Tukey HSD correction and Kenward-Roger degrees of freedom approximation (`emmeans` package) on mixed-effects models (`lme4` package) to account for the cross-validation fold structure.

Comparison	Estimate	Std.error	Df	Statistic	Adj.p.value	Significance
DMK - Radiometric	0.05	0.08	54	0.61	0.996	
DMK - RadiometricDMK	0.07	0.08	54	0.84	0.979	
DMK - RadiometricTerrain	-0.14	0.08	54	-1.79	0.559	
DMK - RadiometricTerrainDMK	-0.16	0.08	54	-1.95	0.458	
DMK - Terrain	-0.19	0.08	54	-2.33	0.248	
DMK - TerrainDMK	-0.22	0.08	54	-2.67	0.127	
Radiometric - RadiometricDMK	0.02	0.08	54	0.24	1.000	
Radiometric - RadiometricTerrain	-0.19	0.08	54	-2.40	0.220	
Radiometric - RadiometricTerrainDMK	-0.21	0.08	54	-2.55	0.161	
Radiometric - Terrain	-0.24	0.08	54	-2.94	0.068	
Radiometric - TerrainDMK	-0.26	0.08	54	-3.27	0.029	*
RadiometricDMK - RadiometricTerrain	-0.21	0.08	54	-2.63	0.136	
RadiometricDMK - RadiometricTerrainDMK	-0.23	0.08	54	-2.79	0.096	
RadiometricDMK - Terrain	-0.26	0.08	54	-3.17	0.037	*
RadiometricDMK - TerrainDMK	-0.28	0.08	54	-3.51	0.015	*
RadiometricTerrain - RadiometricTerrainDMK	-0.01	0.08	54	-0.16	1.000	
RadiometricTerrain - Terrain	-0.04	0.08	54	-0.54	0.998	
RadiometricTerrain - TerrainDMK	-0.07	0.08	54	-0.88	0.975	
RadiometricTerrainDMK - Terrain	-0.03	0.08	54	-0.38	1.000	
RadiometricTerrainDMK - TerrainDMK	-0.06	0.08	54	-0.72	0.991	
Terrain - TerrainDMK	-0.03	0.08	54	-0.33	1.000	

Table C5. Pairwise comparisons of R^2 among models for Ørskogfjellet, based on 10-fold cross-validation. The comparisons use estimated marginal means with Tukey HSD correction and Kenward-Roger degrees of freedom approximation (`emmeans` package) on mixed-effects models (`lme4` package) to account for the cross-validation fold structure.

Comparison	Estimate	Std.error	Df	Statistic	Adj.p.value	Significance
DMK - Radiometric	0.08	0.06	54	1.36	0.818	
DMK - RadiometricDMK	0.08	0.06	54	1.35	0.823	
DMK - RadiometricTerrain	-0.10	0.06	54	-1.72	0.607	
DMK - RadiometricTerrainDMK	-0.11	0.06	54	-1.91	0.485	
DMK - Terrain	-0.15	0.06	54	-2.62	0.14	
DMK - TerrainDMK	-0.19	0.06	54	-3.29	0.028	*
Radiometric - RadiometricDMK	0.00	0.06	54	-0.01	1	
Radiometric - RadiometricTerrain	-0.18	0.06	54	-3.08	0.048	*
Radiometric - RadiometricTerrainDMK	-0.19	0.06	54	-3.27	0.029	*
Radiometric - Terrain	-0.24	0.06	54	-3.98	0.004	**
Radiometric - TerrainDMK	-0.28	0.06	54	-4.65	< 0.001	***
RadiometricDMK - RadiometricTerrain	-0.18	0.06	54	-3.07	0.049	*
RadiometricDMK - RadiometricTerrainDMK	-0.19	0.06	54	-3.26	0.03	*
RadiometricDMK - Terrain	-0.24	0.06	54	-3.97	0.004	**
RadiometricDMK - TerrainDMK	-0.27	0.06	54	-4.64	< 0.001	***
RadiometricTerrain - RadiometricTerrainDMK	-0.01	0.06	54	-0.19	1	
RadiometricTerrain - Terrain	-0.05	0.06	54	-0.90	0.971	
RadiometricTerrain - TerrainDMK	-0.09	0.06	54	-1.57	0.702	
RadiometricTerrainDMK - Terrain	-0.04	0.06	54	-0.71	0.991	
RadiometricTerrainDMK - TerrainDMK	-0.08	0.06	54	-1.38	0.809	
Terrain - TerrainDMK	-0.04	0.06	54	-0.67	0.994	

Table C6. Pairwise comparisons of mean absolute error among models for Ørskogfjellet, based on 10-fold cross-validation. The comparisons use estimated marginal means with Tukey HSD correction and Kenward-Roger degrees of freedom approximation (`emmeans` package) on mixed-effects models (`lme4` package) to account for the cross-validation fold structure.

Comparison	Estimate	Std.error	Df	Statistic	Adj.p.value	Significance
DMK - Radiometric	-6.93	8.69	54	-0.80	0.984	
DMK - RadiometricDMK	-7.63	8.69	54	-0.88	0.974	
DMK - RadiometricTerrain	14.31	8.69	54	1.65	0.653	
DMK - RadiometricTerrainDMK	15.83	8.69	54	1.82	0.539	
DMK - Terrain	17.71	8.69	54	2.04	0.403	
DMK - TerrainDMK	20.55	8.69	54	2.37	0.233	
Radiometric - RadiometricDMK	-0.70	8.69	54	-0.08	1.000	
Radiometric - RadiometricTerrain	21.24	8.69	54	2.45	0.200	
Radiometric - RadiometricTerrainDMK	22.76	8.69	54	2.62	0.140	
Radiometric - Terrain	24.64	8.69	54	2.84	0.086	
Radiometric - TerrainDMK	27.48	8.69	54	3.16	0.039	*
RadiometricDMK - RadiometricTerrain	21.95	8.69	54	2.53	0.170	
RadiometricDMK - RadiometricTerrainDMK	23.47	8.69	54	2.70	0.118	
RadiometricDMK - Terrain	25.35	8.69	54	2.92	0.071	
RadiometricDMK - TerrainDMK	28.19	8.69	54	3.24	0.031	*
RadiometricTerrain - RadiometricTerrainDMK	1.52	8.69	54	0.17	1.000	
RadiometricTerrain - Terrain	3.40	8.69	54	0.39	1.000	
RadiometricTerrain - TerrainDMK	6.24	8.69	54	0.72	0.991	
RadiometricTerrainDMK - Terrain	1.88	8.69	54	0.22	1.000	
RadiometricTerrainDMK - TerrainDMK	4.72	8.69	54	0.54	0.998	
Terrain - TerrainDMK	2.84	8.69	54	0.33	1.000	

. *Code and data availability.* R code used in this study is available at https://github.com/julienvollering/DSMdepth. The depth measurements we used are archived at https://doi.org/10.6073/pasta/6ce440152f693f2156bf5b692a2e7917 and follow data and metadata standards.

645   . *Author contributions.* JV: Conceptualization, Investigation, Data curation, Formal analysis, Writing – original draft. NG: Conceptualization, Methodology, Writing - review & editing. MKG: Investigation, Writing - review & editing. KKM: Investigation, Data curation, Writing - review & editing. SDN: Investigation, Writing - review & editing. KR: Conceptualization, Investigation, Writing - review & editing. MS: Investigation, Data curation, Writing - review & editing.

. *Competing interests.* The authors declare that they have no conflict of interest.

650   . *Disclaimer.* The authors declare that the results, discussions, and interpretations presented in this study are solely their own. The views expressed herein do not necessarily reflect those of their respective institutions or funding agencies.

. *Acknowledgements.* We thank the Norwegian Public Roads Administration for sharing data from ground-penetrating radar surveys. We also thank Vikas Baranwal from the Geological Survey of Norway for helping us access the radiometric data from Skrim. This work contains data under the following licenses: (1) Creative Commons Attribution 4.0 International, ©Kartverket, (2) *Norge digitalt* license, Norwegian

655   Institute of Bioeconomy Research (NIBIO), ©Geovekst, and (3) the Norwegian License for Public Data (NLOD), made available by the Geological Survey of Norway. Large language models have been used to expedite coding of the analysis, with author oversight. Large language models have also been used during the drafting and editing of this manuscript, with author oversight. We maintain full responsibility for the scientific output, as per the European Commission's *Living guidelines on the responsible use of generative AI in research*.

References

660 Ågren, A. M., Lidberg, W., Strömgren, M., Ogilvie, J., and Arp, P. A.: Evaluating Digital Terrain Indices for Soil Wetness Mapping – a Swedish Case Study, Hydrology and Earth System Sciences, 18, 3623–3634, https://doi.org/10.5194/hess-18-3623-2014, 2014.

Ahlstrøm, A., Bjørkelo, K., and Fadnes, K. D.: AR5 Klassifikasjonssystem, Tech. rep., NIBIO, 2019.

Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A. C., Morgan, C. L. S., Roudier, P., Poggio, L., and Mulder, V. L.: Impressions of Digital Soil Maps: The Good, the Not so Good, and Making Them Ever Better, Geoderma Regional, 20, e00 255,
665 https://doi.org/10.1016/j.geodrs.2020.e00255, 2020.

Bakkestuen, V., Venter, Z., Ganerød, A. J., and Framstad, E.: Delineation of Wetland Areas in South Norway from Sentinel-2 Imagery and LiDAR Using TensorFlow, U-Net, and Google Earth Engine, Remote Sensing, 15, 1203, https://doi.org/10.3390/rs15051203, 2023.

Baranwal, V., Rodionov, A., Ofstad, F., Koziel, J., and Lynum, R.: Helicopter-Borne Magnetic, Electromagnetic and Radiometric Geophysical Surveys in the Kongsberg Region: Krøderen, Sokna, Høneføss, Kongsberg and Numedalen., Tech. Rep. 2013.029, Geological Survey of
670 Norway, 2013.

Baranwal, V. C. and Rønning, J. S.: Airborne Geophysical Surveys and Their Integrated Interpretation, in: Advances in Modeling and Interpretation in Near Surface Geophysics, edited by Biswas, A. and Sharma, S. P., pp. 377–400, Springer International Publishing, Cham, ISBN 978-3-030-28908-9 978-3-030-28909-6, https://doi.org/10.1007/978-3-030-28909-6_14, 2020.

Beamish, D.: Gamma Ray Attenuation in the Soils of Northern Ireland, with Special Reference to Peat, Journal of Environmental
675 Radioactivity, 115, 13–27, https://doi.org/10.1016/j.jenvrad.2012.05.031, 2013.

Beamish, D.: Environmental Radioactivity in the UK: The Airborne Geophysical View of Dose Rate Estimates, Journal of Environmental Radioactivity, 138, 249–263, https://doi.org/10.1016/j.jenvrad.2014.08.025, 2014.

Beamish, D.: Enhancing the Resolution of Airborne Gamma-Ray Data Using Horizontal Gradients, Journal of Applied Geophysics, 132, 75–86, https://doi.org/10.1016/j.jappgeo.2016.07.006, 2016.

680 Beamish, D. and White, J. C.: On the Detailed Mapping of Peat (Raised Bogs) Using Airborne Radiometric Data, Journal of Environmental Radioactivity, 277, 107 462, https://doi.org/10.1016/j.jenvrad.2024.107462, 2024.

Beguin, J., Fuglstad, G.-A., Mansuy, N., and Paré, D.: Predicting Soil Properties in the Canadian Boreal Forest with Limited Data: Comparison of Spatial and Non-Spatial Statistical Approaches, Geoderma, 306, 195–205, https://doi.org/10.1016/j.geoderma.2017.06.016, 2017.

685 Beilman, D. W., Vitt, D. H., Bhatti, J. S., and Forest, S.: Peat Carbon Stocks in the Southern Mackenzie River Basin: Uncertainties Revealed in a High-Resolution Case Study, Global Change Biology, 14, 1221–1232, https://doi.org/10.1111/j.1365-2486.2008.01565.x, 2008.

Biau, G. and Scornet, E.: A Random Forest Guided Tour, TEST, 25, 197–227, https://doi.org/10.1007/s11749-016-0481-7, 2016.

Biecek, P. and Burzykowski, T.: Explanatory Model Analysis, Chapman and Hall/CRC, New York, ISBN 978-0-367-13559-1, 2021.

Billen, N., Kuhfeld, H., Kern, S., and Herrmann, L.: Eignung der Gammaspektrometrie zum Kartieren der Mächtigkeit und der C-Vorräte
690 von Moorböden, in: Jahrestagung der DBG 2015: Unsere Böden - unser Leben, München, 2015.

Bjørdal, I.: Markslagsklassifikasjon i Økonomisk Kartverk. 2007-utgåva, Report, Norsk institutt for skog og landskap, ISBN 9788231100096, 2007.

Bohn, M. P. and Miller, B. A.: Locally Enhanced Digital Soil Mapping in Support of a Bottom-up Approach Is More Accurate than Conventional Soil Mapping and Top-down Digital Soil Mapping, Geoderma, 442, 116 781,
695 https://doi.org/10.1016/j.geoderma.2024.116781, 2024.

Borowiec, M. L., Dikow, R. B., Frandsen, P. B., McKeeken, A., Valentini, G., and White, A. E.: Deep Learning as a Tool for Ecology and Evolution, Methods in Ecology and Evolution, 13, 1640–1660, https://doi.org/10.1111/2041-210X.13901, 2022.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

700   Brus, D. J.: Sampling for Digital Soil Mapping: A Tutorial Supported by R Scripts, Geoderma, 338, 464–480, https://doi.org/10.1016/j.geoderma.2018.07.036, 2019.

Bryn, A., Strand, G.-H., Angeloff, M., and Rekdal, Y.: Land Cover in Norway Based on an Area Frame Survey of Vegetation Types, Norwegian Journal of Geography, 72, 131–145, https://doi.org/10.1080/00291951.2018.1468356, 2018.

Buffam, I., Carpenter, S. R., Yeck, W., Hanson, P. C., and Turner, M. G.: Filling Holes in Regional Carbon Budgets: Predicting Peat Depth in a North Temperate Lake District, Journal of Geophysical Research: Biogeosciences, 115, G01 005, https://doi.org/10.1029/2009JG001034, 705   2010.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J.: System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, Geoscientific Model Development, 8, 1991–2007, https://doi.org/10.5194/gmd-8-1991-2015, 2015.

Deragon, R., Saurette, D. D., Heung, B., and Caron, J.: Mapping the Maximum Peat Thickness of Cultivated Organic Soils in the Southwest 710   Plain of Montreal, Canadian Journal of Soil Science, 103, 103–120, https://doi.org/10.1139/cjss-2022-0031, 2023.

Dornik, A., Cheţan, M. A., Drăguţ, L., Dicu, D. D., and Iliuţă, A.: Optimal Scaling of Predictors for Digital Mapping of Soil Properties, Geoderma, 405, 115 453, https://doi.org/10.1016/j.geoderma.2021.115453, 2022.

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., and Ranjan, R.: Explainable AI (XAI): Core Ideas, Techniques, and Solutions, ACM Comput. Surv., 55, 194:1–194:33, https://doi.org/10.1145/3561048, 2023.

715   Finlayson, A., Marchant, B. P., Whitbread, K., Hughes, L., and Barron, H. F.: Estimating Organic Surface Horizon Depth for Peat and Peaty Soils across a Scottish Upland Catchment Using Linear Mixed Models with Topographic and Geological Covariates, Soil Use and Management, 37, 628–639, https://doi.org/10.1111/sum.12596, 2021.

Flaget, V. U., Vasstrøm, M., and Overvåg, K.: Myr Mellom Lokal Og Nasjonal Planlegging Og Styring – Planlegging Av Fjelldestinasjonen Turufjell, Kart og Plan, 117, 276–295, https://doi.org/10.18261/kp.117.2.14, 2024.

720   Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., 725   Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, Earth System Science Data, 12, 3269–3340, https://doi.org/10.5194/essd-12-3269-2020, 2020.

730   Gallant, J. C. and Dowling, T. I.: A Multiresolution Index of Valley Bottom Flatness for Mapping Depositional Areas, Water Resources Research, 39, 1347, https://doi.org/10.1029/2002WR001426, 2003.

Gatis, N., Luscombe, D., Carless, D., Parry, L., Fyfe, R., Harrod, T., Brazier, R., and Anderson, K.: Mapping Upland Peat Depth Using Airborne Radiometric and Lidar Survey Data, Geoderma, 335, 78–87, https://doi.org/10.1016/j.geoderma.2018.07.041, 2019.

Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E.: Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation, Journal of Computational and Graphical Statistics, 24, 44–65, https://doi.org/10.1080/10618600.2014.907095, 2015.

GRASS Development Team: Geographic Resources Analysis Support System (GRASS GIS) Software, Version 8.2, Open Source Geospatial Foundation, 2022.

Greenwell, B. M.: Pdp: An R Package for Constructing Partial Dependence Plots, The R Journal, 9, 421–436, 2017.

Greenwell, B. M. and Boehmke, B. C.: Variable Importance Plots—An Introduction to the Vip Package, The R Journal, 12, 343–366, 2020.

Hastie, A., Honorio Coronado, E. N., Reyna, J., Mitchard, E. T. A., Åkesson, C. M., Baker, T. R., Cole, L. E. S., Oroche, C. J. C., Dargie, G., Dávila, N., De Grandi, E. C., Del Águila, J., Del Castillo Torres, D., De La Cruz Paiva, R., Draper, F. C., Flores, G., Grández, J., Hergoualc'h, K., Householder, J. E., Janovec, J. P., Lähteenoja, O., Reyna, D., Rodríguez-Veiga, P., Roucoux, K. H., Tobler, M., Wheeler, C. E., Williams, M., and Lawson, I. T.: Risks to Carbon Storage from Land-Use Change Revealed by Peat Thickness Maps of Peru, Nature Geoscience, 15, 369–374, https://doi.org/10.1038/s41561-022-00923-4, 2022.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer Series in Statistics, Springer, New York, USA, 2 edn., ISBN 978-0-387-21606-5, 2009.

Hattab, T., Garzón-López, C. X., Ewald, M., Skowronek, S., Aerts, R., Horen, H., Brasseur, B., Gallet-Moron, E., Spicher, F., Decocq, G., Feilhauer, H., Honnay, O., Kempeneers, P., Schmidtlein, S., Somers, B., Kerchove, R. V. D., Rocchini, D., and Lenoir, J.: A Unified Framework to Model the Potential and Realized Distributions of Invasive Species within the Invaded Range, Diversity and Distributions, 23, 806–819, https://doi.org/10.1111/ddi.12566, 2017.

Hengl, T., Heuvelink, G. B. M., and Stein, A.: A Generic Framework for Spatial Prediction of Soil Variables Based on Regression-Kriging, Geoderma, 120, 75–93, https://doi.org/10.1016/j.geoderma.2003.08.018, 2004.

Henrion, M., Li, Y., Koganti, T., Bechtold, M., Jonard, F., Opfergelt, S., Vanacker, V., Van Oost, K., and Lambot, S.: Mapping and Monitoring Peatlands in the Belgian Hautes Fagnes: Insights from Ground-penetrating Radar and Electromagnetic Induction Characterization, Geoderma Regional, 37, e00 795, https://doi.org/10.1016/j.geodrs.2024.e00795, 2024.

Heuvelink, G. B. M. and Webster, R.: Spatial Statistics and Soil Mapping: A Blossoming Partnership under Pressure, Spatial Statistics, 50, 100 639, https://doi.org/10.1016/j.spasta.2022.100639, 2022.

Høgaas, F., Olsen, L., Sveian, H., Rindstad, B. I., and Hansen, L.: Database for registrering av marin grense (MG) i Norge, Report, Geological Survey of Norway, ISSN 0800-3416, 2012.

Holden, N. M. and Connolly, J.: Estimating the Carbon Stock of a Blanket Peat Region Using a Peat Depth Inference Model, CATENA, 86, 75–85, https://doi.org/10.1016/j.catena.2011.02.002, 2011.

IAEA: Guidelines for Radioelement Mapping Using Gamma Ray Spectrometry Data, International Atomic Energy Agency, Vienna, ISBN 978-92-0-108303-6, 2003.

Ivanovs, J., Haberl, A., and Melniks, R.: Modeling Geospatial Distribution of Peat Layer Thickness Using Machine Learning and Aerial Laser Scanning Data, Land, 13, 466, https://doi.org/10.3390/land13040466, 2024.

Joosten, H. and Clarke, D.: Wise Use of Mires and Peatlands: Background and Principles Including a Framework for Decision-Making, International Mire Conservation Group and International Peat Society, Saarijärvi, Finland, ISBN 951-97744-8-3, 2002.

Joosten, H., Sirin, A., Couwenberg, J., Laine, J., and Smith, P.: The Role of Peatlands in Climate Regulation, in: Peatland Restoration and Ecosystem Services, edited by Bonn, A., Allott, T., Evans, M., Joosten, H., and Stoneman, R., pp. 63–76, Cambridge University Press, Cambridge, ISBN 978-1-139-17778-8, https://doi.org/10.1017/CBO9781139177788.005, 2016.

Karjalainen, V., Malinen, J., and Tokola, T.: Comparison of Two Gamma-Ray Datasets Measured with Different Methods and Assessment of Their Performance to Predict Soil Properties, International Journal of Forest Engineering, 36, 29–42, https://doi.org/10.1080/14942119.2024.2398947, 2025.

775 Keaney, A., McKinley, J., Graham, C., Robinson, M., and Ruffell, A.: Spatial Statistics to Estimate Peat Thickness Using Airborne Radiometric Data, Spatial Statistics, 5, 3–24, https://doi.org/10.1016/j.spasta.2013.05.003, 2013.

Khaledian, Y. and Miller, B. A.: Selecting Appropriate Machine Learning Methods for Digital Soil Mapping, Applied Mathematical Modelling, 81, 401–418, https://doi.org/10.1016/j.apm.2019.12.016, 2020.

Kock, P. and Samuelsson, C.: Comparison of Airborne and Terrestrial Gamma Spectrometry Measurements - Evaluation of Three Areas in

780 Southern Sweden, Journal of Environmental Radioactivity, 102, 605–613, https://doi.org/10.1016/j.jenvrad.2011.03.010, 2011.

Koganti, T., Vigah Adetsu, D., Triantafilis, J., Greve, M. H., and Beucher, A. M.: Mapping Peat Depth Using a Portable Gamma-Ray Sensor and Terrain Attributes, Geoderma, 439, 116 672, https://doi.org/10.1016/j.geoderma.2023.116672, 2023.

Kuhn, M. and Wickham, H.: Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles., 2020.

Lamichhane, S., Kumar, L., and Wilson, B.: Digital Soil Mapping Algorithms and Covariates for Soil Organic Carbon Mapping and Their

785 Implications: A Review, Geoderma, 352, 395–413, https://doi.org/10.1016/j.geoderma.2019.05.031, 2019.

Li, Y., Henrion, M., Moore, A., Lambot, S., Opfergelt, S., Vanacker, V., Jonard, F., and Van Oost, K.: Factors Controlling Peat Soil Thickness and Carbon Storage in Temperate Peatlands Based on UAV High-Resolution Remote Sensing, Geoderma, 449, 117 009, https://doi.org/10.1016/j.geoderma.2024.117009, 2024.

Li, Y., Han, D., Rogers, C. A., Finkelstein, S. A., Hararuk, O., Waddington, J. M., Barreto, C., McLaughlin, J. W., Snider, J., and

790 Gonsamo, A.: Peat Depth and Carbon Storage of the Hudson Bay Lowlands, Canada, Geophysical Research Letters, 52, e2024GL110 679, https://doi.org/10.1029/2024GL110679, 2025.

Lindsay, J. B.: Whitebox GAT: A Case Study in Geomorphometric Analysis., Computers & Geosciences, 95, 75–84, 2016a.

Lindsay, R.: Bogs: The Ecology, Classification and Conservation of Ombrotrophic Mires, Technical report, Scottish Natural Heritage, Edinburgh, ISBN 9781853971006, 1995.

795 Lindsay, R.: Peatbogs and Carbon: A Critical Synthesis to Inform Policy Development in Oceanic Peat Bog Conservation and Restoration in the Context of Climate Change, Technical report, RSPB Scotland, 2010.

Lindsay, R.: Peatland (Mire Types): Based on Origin and Behavior of Water, Peat Genesis, Landscape Position, and Climate, in: The Wetland Book: II: Distribution, Description and Conservation, edited by Finlayson, C. M., Milton, G. R., Prentice, R. C., and Davidson, N. C., pp. 1–23, Springer Netherlands, ISBN 978-94-007-6173-5, 2016b.

800 Linnenbrink, J., Milà, C., Ludwig, M., and Meyer, H.: kNNDM CV: K-Fold Nearest-Neighbour Distance Matching Cross-Validation for Map Accuracy Estimation, Geoscientific Model Development, 17, 5897–5912, https://doi.org/10.5194/gmd-17-5897-2024, 2024.

Loisel, J., Yu, Z., Beilman, D. W., Camill, P., Alm, J., Amesbury, M. J., Anderson, D., Andersson, S., Bochicchio, C., Barber, K., Belyea, L. R., Bunbury, J., Chambers, F. M., Charman, D. J., De Vleeschouwer, F., Fiałkiewicz-Kozieł, B., Finkelstein, S. A., Gałka, M., Garneau, M., Hammarlund, D., Hinchcliffe, W., Holmquist, J., Hughes, P., Jones, M. C., Klein, E. S., Kokfelt, U., Korhola, A., Kuhry, P., Lamarre, A.,

805 Lamentowicz, M., Large, D., Lavoie, M., MacDonald, G., Magnan, G., Mäkilä, M., Mallon, G., Mathijssen, P., Mauquoy, D., McCarroll, J., Moore, T. R., Nichols, J., O'Reilly, B., Oksanen, P., Packalen, M., Peteet, D., Richard, P. J., Robinson, S., Ronkainen, T., Rundgren, M., Sannel, A. B. K., Tarnocai, C., Thom, T., Tuittila, E.-S., Turetsky, M., Väliranta, M., van der Linden, M., van Geel, B., van Bellen, S., Vitt, D., Zhao, Y., and Zhou, W.: A Database and Synthesis of Northern Peatland Soil Properties and Holocene Carbon and Nitrogen Accumulation, The Holocene, 24, 1028–1042, https://doi.org/10.1177/0959683614538073, 2014.

810 Lyngstad, A., Moen, A., Halvorsen, R., and Øien, D.-I.: Beskrivelser av torvmassivenheter. Kunnskapsgrunnlag for NiN 3.0, Tech. rep., NTNU Vitenskapsmuseet, 2023.

Ma, L., Zhu, G., Chen, B., Zhang, K., Niu, S., Wang, J., Ciais, P., and Zuo, H.: A Globally Robust Relationship between Water Table Decline, Subsidence Rate, and Carbon Release from Peatlands, Communications Earth & Environment, 3, 1–14, https://doi.org/10.1038/s43247-022-00590-8, 2022.

815 Ma, T., Brus, D. J., Zhu, A.-X., Zhang, L., and Scholten, T.: Comparison of Conditioned Latin Hypercube and Feature Space Coverage Sampling for Predicting Soil Classes Using Simulation from Soil Maps, Geoderma, 370, 114 366, https://doi.org/10.1016/j.geoderma.2020.114366, 2020.

Malone, B. P., Minansy, B., and Brungard, C.: Some Methods to Improve the Utility of Conditioned Latin Hypercube Sampling, PeerJ, 7, e6451, https://doi.org/10.7717/peerj.6451, 2019.

820 Marchant, B. P.: Using Remote Sensors to Predict Soil Properties: Radiometry and Peat Depth in Dartmoor, UK, Geoderma, 403, 115 232, https://doi.org/10.1016/j.geoderma.2021.115232, 2021.

Maxwell, A. E. and Shobe, C. M.: Land-Surface Parameters for Spatial Predictive Mapping and Modeling, Earth-Science Reviews, 226, 103 944, https://doi.org/10.1016/j.earscirev.2022.103944, 2022.

McBratney, A. B., Mendonça Santos, M. L., and Minasny, B.: On Digital Soil Mapping, Geoderma, 117, 3–52, https://doi.org/10.1016/S0016-825 7061(03)00223-4, 2003.

Meinshausen, N.: Quantile Regression Forests, Journal of Machine Learning Research, 7, 983–999, 2006.

Meyer, H. and Pebesma, E.: Machine Learning-Based Global Maps of Ecological Variables and the Challenge of Assessing Them, Nature Communications, 13, 2208, https://doi.org/10.1038/s41467-022-29838-9, 2022.

Meyer, H., Ludwig, M., Milà, C., Linnenbrink, J., and Schumacher, F.: The CAST Package for Training and Assessment of Spatial Prediction 830 Models in R, https://doi.org/10.48550/arXiv.2404.06978, 2024.

Milà, C., Mateu, J., Pebesma, E., and Meyer, H.: Nearest Neighbour Distance Matching Leave-One-Out Cross-Validation for Map Validation, Methods in Ecology and Evolution, 13, 1304–1316, https://doi.org/10.1111/2041-210X.13851, 2022.

Miller, B. A., Koszinski, S., Wehrhan, M., and Sommer, M.: Impact of Multi-Scale Predictor Selection for Modeling Soil Properties, Geoderma, 239–240, 97–106, https://doi.org/10.1016/j.geoderma.2014.09.018, 2015.

835 Minasny, B. and McBratney, A. B.: A Conditioned Latin Hypercube Method for Sampling in the Presence of Ancillary Information, Computers & Geosciences, 32, 1378–1388, https://doi.org/10.1016/j.cageo.2005.12.009, 2006.

Minasny, B., Berglund, Ö., Connolly, J., Hedley, C., de Vries, F., Gimona, A., Kempen, B., Kidd, D., Lilja, H., Malone, B., McBratney, A., Roudier, P., O'Rourke, S., Rudiyanto, Padarian, J., Poggio, L., ten Caten, A., Thompson, D., Tuve, C., and Widyatmanti, W.: Digital Mapping of Peatlands – A Critical Review, Earth-Science Reviews, 196, 102 870, https://doi.org/10.1016/j.earscirev.2019.05.014, 2019.

840 Minasny, B., Adetsu, D. V., Aitkenhead, M., Artz, R. R. E., Baggaley, N., Barthelmes, A., Beucher, A., Caron, J., Conchedda, G., Connolly, J., Deragon, R., Evans, C., Fadnes, K., Fiantis, D., Gagkas, Z., Gilet, L., Gimona, A., Glatzel, S., Greve, M. H., Habib, W., Hergoualc'h, K., Hermansen, C., Kidd, D. B., Koganti, T., Kopansky, D., Large, D. J., Larmola, T., Lilly, A., Liu, H., Marcus, M., Middleton, M., Morrison, K., Petersen, R. J., Quaife, T., Rochefort, L., Rudiyanto, Toca, L., Tubiello, F. N., Weber, P. L., Weldon, S., Widyatmanti, W., Williamson, J., and Zak, D.: Mapping and Monitoring Peatland Conditions from Global to Field Scale, Biogeochemistry, 167, 383–425, 845 https://doi.org/10.1007/s10533-023-01084-1, 2024a.

Minasny, B., Bandai, T., Ghezzehei, T. A., Huang, Y.-C., Ma, Y., McBratney, A. B., Ng, W., Norouzi, S., Padarian, J., Rudiyanto, Sharififar, A., Styc, Q., and Widyastuti, M.: Soil Science-Informed Machine Learning, Geoderma, 452, 117 094, https://doi.org/10.1016/j.geoderma.2024.117094, 2024b.

Moen, A.: National Atlas of Norway: Vegetation, Norwegian Mapping Authority, Hønefoss, 1999.

850 Molla, A., Zhang, W., Zuo, S., Ren, Y., and Han, J.: A Machine Learning and Geostatistical Hybrid Method to Improve Spatial Prediction Accuracy of Soil Potentially Toxic Elements, Stochastic Environmental Research and Risk Assessment, 37, 681–696, https://doi.org/10.1007/s00477-022-02284-1, 2023.

Mulder, V. L., de Bruin, S., Schaepman, M. E., and Mayr, T. R.: The Use of Remote Sensing in Soil and Terrain Mapping — A Review, Geoderma, 162, 1–19, https://doi.org/10.1016/j.geoderma.2010.12.018, 2011.

855 Murphy, P. N. C., Ogilvie, J., Connor, K., and Arp, P. A.: Mapping Wetlands: A Comparison of Two Different Approaches for New Brunswick, Canada, Wetlands, 27, 846–854, https://doi.org/10.1672/0277-5212(2007)27[846:MWACOT]2.0.CO;2, 2007.

Murphy, P. N. C., Ogilvie, J., and Arp, P.: Topographic Modelling of Soil Moisture Conditions: A Comparison and Verification of Two Models, European Journal of Soil Science, 60, 94–109, https://doi.org/10.1111/j.1365-2389.2008.01094.x, 2009.

Murphy, P. N. C., Ogilvie, J., Meng, F.-R., White, B., Bhatti, J. S., and Arp, P. A.: Modelling and Mapping Topographic Variations in Forest
860 Soils at High Resolution: A Case Study, Ecological Modelling, 222, 2314–2332, https://doi.org/10.1016/j.ecolmodel.2011.01.003, 2011.

Natural England: Data Exchange Standard for Peat Surveys v1.0, Tech. rep., Natural England, 2023.

Newman, D. R., Saurette, D. D., Cockburn, J. M. H., Dragut, L., and Lindsay, J. B.: Assessing Spatially Heterogeneous Scale Representation with Applied Digital Soil Mapping, Environmental Modelling & Software, 160, 105 612, https://doi.org/10.1016/j.envsoft.2022.105612, 2023.

865 Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, L., Schaepman, M. E., and Papritz, A.: Evaluation of Digital Soil Mapping Approaches with Large Sets of Environmental Covariates, SOIL, 4, 1–22, https://doi.org/10.5194/soil-4-1-2018, 2018.

OECD: OECD Environmental Performance Reviews: Norway 2022, Tech. rep., OECD, https://doi.org/10.1787/59e71c13-en, 2022.

Ofstad, F.: Helicopter-Borne Magnetic and Radiometric Geophysical Survey in Romsdalsfjorden, Møre Og Romsdal, Tech. Rep. 2015.015, Geological Survey of Norway, 2015.

870 O'Leary, D., Brown, C., and Daly, E.: Digital Soil Mapping of Peatland Using Airborne Radiometric Data and Supervised Machine Learning – Implication for the Assessment of Carbon Stock, Geoderma, 428, 116 086, https://doi.org/10.1016/j.geoderma.2022.116086, 2022.

Parry, L., West, L., Holden, J., and Chapman, P.: Evaluating Approaches for Estimating Peat Depth, Journal of Geophysical Research: Biogeosciences, 119, 567–576, https://doi.org/10.1002/2013JG002411, 2014.

Parry, L. E., Charman, D. J., and Noades, J. P. W.: A Method for Modelling Peat Depth in Blanket Peatlands, Soil Use and Management, 28,
875 614–624, https://doi.org/10.1111/j.1475-2743.2012.00447.x, 2012.

Parsekian, A. D., Slater, L., Ntarlagiannis, D., Nolan, J., Sebesteyen, S. D., Kolka, R. K., and Hanson, P. J.: Uncertainty in Peat Volume and Soil Carbon Estimated Using Ground-Penetrating Radar and Probing, Soil Science Society of America Journal, 76, 1911–1918, https://doi.org/10.2136/sssaj2012.0040, 2012.

Pelletier, R., Davis, J., and Rossiter, J.: Peat Analyses In The Hudson Bay Lowlands Using Ground Penetrating Radar,
880 in: [Proceedings] IGARSS'91 Remote Sensing: Global Monitoring for Earth Management, vol. 4, pp. 2141–2144, https://doi.org/10.1109/IGARSS.1991.575463, 1991.

Pohjankukka, J., Räsänen, T. A., Pitkänen, T. P., Kivimäki, A., Mäkinen, V., Väänänen, T., Lerssi, J., Salmivaara, A., and Middleton, M.: Digital Mapping of Peat Thickness and Extent in Finland Using Remote Sensing and Machine Learning, Geoderma, 455, 117 216, https://doi.org/10.1016/j.geoderma.2025.117216, 2025.

885 Price, J., Evans, C., Evans, M., Allott, T., and Shuttleworth, E.: Peatland Restoration and Hydrology, in: Peatland Restoration and Ecosystem Services, edited by Bonn, A., Allott, T., Evans, M., Joosten, H., and Stoneman, R., pp. 77–94, Cambridge University Press, Cambridge, ISBN 978-1-139-17778-8, https://doi.org/10.1017/CBO9781139177788.006, 2016.

Probst, P., Wright, M. N., and Boulesteix, A.-L.: Hyperparameters and Tuning Strategies for Random Forest, WIREs Data Mining and Knowledge Discovery, 9, e1301, https://doi.org/10.1002/widm.1301, 2019.

890 Proulx-McInnis, S., St-Hilaire, A., Rousseau, A. N., and Jutras, S.: A Review of Ground-Penetrating Radar Studies Related to Peatland Stratigraphy with a Case Study on the Determination of Peat Thickness in a Northern Boreal Fen in Quebec, Canada, Progress in Physical Geography: Earth and Environment, 37, 767–786, https://doi.org/10.1177/0309133313501106, 2013.

Quinn, P., Beven, K., Chevallier, P., and Planchon, O.: The Prediction of Hillslope Flow Paths for Distributed Hydrological Modelling Using Digital Terrain Models, Hydrological Processes, 5, 59–79, https://doi.org/10.1002/hyp.3360050106, 1991.

895 Reinhardt, N. and Herrmann, L.: Gamma-Ray Spectrometry as Versatile Tool in Soil Science: A Critical Review, Journal of Plant Nutrition and Soil Science, 182, 9–27, https://doi.org/10.1002/jpln.201700447, 2019.

Riihimäki, H., Kemppinen, J., Kopecký, M., and Luoto, M.: Topographic Wetness Index as a Proxy for Soil Moisture: The Importance of Flow-Routing Algorithm and Grid Resolution, Water Resources Research, 57, e2021WR029 871, https://doi.org/10.1029/2021WR029871, 2021.

900 Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., and Dormann, C. F.: Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure, Ecography, 40, 913–929, https://doi.org/10.1111/ecog.02881, 2017.

Rosa, E., Larocque, M., Pellerin, S., Gagné, S., and Fournier, B.: Determining the Number of Manual Measurements Required to Improve Peat Thickness Estimations by Ground Penetrating Radar, Earth Surface Processes and Landforms, 34, 377–383,
905 https://doi.org/10.1002/esp.1741, 2009.

Roudier, P.: Clhs: A R Package for Conditioned Latin Hypercube Sampling., 2011.

Rudiyanto, Minasny, B., Setiawan, B. I., Arif, C., Saptomo, S. K., and Chadirin, Y.: Digital Mapping for Cost-Effective and Accurate Prediction of the Depth and Carbon Stocks in Indonesian Peatlands, Geoderma, 272, 20–31, https://doi.org/10.1016/j.geoderma.2016.02.026, 2016.

910 Rudiyanto, Minasny, B., Setiawan, B., Saptomo, S., and McBratney, A.: Open Digital Mapping as a Cost-Effective Method for Mapping Peat Thickness and Assessing the Carbon Stock of Tropical Peatlands, Geoderma, 313, 25–40, https://doi.org/10.1016/j.geoderma.2017.10.018, 2018.

Ruols, B., Baron, L., and Irving, J.: Development of a Drone-Based Ground-Penetrating Radar System for Efficient and Safe 3D and 4D Surveying of Alpine Glaciers, Journal of Glaciology, 69, 2087–2098, https://doi.org/10.1017/jog.2023.83, 2023.

915 Rydin, H., Sjörs, H., and Löfroth, M.: 7. Mires, Acta Phytogeographica Suecica, 84, 91–112, 1999.

Saurette, D. D., Heck, R. J., Gillespie, A. W., Berg, A. A., and Biswas, A.: Divergence Metrics for Determining Optimal Training Sample Size in Digital Soil Mapping, Geoderma, 436, 116 553, https://doi.org/10.1016/j.geoderma.2023.116553, 2023.

Schönauer, M. and Maack, J.: R-Code for Calculating Depth-to-Water (DTW) Maps Using GRASS GIS, https://doi.org/10.5281/zenodo.5718133, 2021.

Schönauer, M., Väätäinen, K., Prinz, R., Lindeman, H., Pszenny, D., Jansen, M., Maack, J., Talbot, B., Astrup, R., and Jaeger, D.: Spatio-Temporal Prediction of Soil Moisture and Soil Strength by Depth-to-Water Maps, International Journal of Applied Earth Observation and Geoinformation, 105, 102 614, https://doi.org/10.1016/j.jag.2021.102614, 2021.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., and Brenning, A.: Hyperparameter Tuning and Performance Assessment of Statistical and Machine-Learning Algorithms Using Spatial Data, Ecological Modelling, 406, 109–120, https://doi.org/10.1016/j.ecolmodel.2019.06.002, 2019.

Shrestha, D. L. and Solomatine, D. P.: Machine Learning Approaches for Estimation of Prediction Interval for the Model Output, Neural Networks, 19, 225–235, https://doi.org/10.1016/j.neunet.2006.01.012, 2006.

Siemon, B., Ibs-von Seht, M., and Frank, S.: Airborne Electromagnetic and Radiometric Peat Thickness Mapping of a Bog in Northwest Germany (Ahlen-Falkenberger Moor), Remote Sensing, 12, 203, https://doi.org/10.3390/rs12020203, 2020.

Strack, M., Davidson, S. J., Hirano, T., and Dunn, C.: The Potential of Peatlands as Nature-Based Climate Solutions, Current Climate Change Reports, 8, 71–82, https://doi.org/10.1007/s40641-022-00183-9, 2022.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., and Zeileis, A.: Conditional Variable Importance for Random Forests, BMC Bioinformatics, 9, 307, https://doi.org/10.1186/1471-2105-9-307, 2008.

Szabó, B., Szatmári, G., Takács, K., Laborczi, A., Makó, A., Rajkai, K., and Pásztor, L.: Mapping Soil Hydraulic Properties Using Random-Forest-Based Pedotransfer Functions and Geostatistics, Hydrology and Earth System Sciences, 23, 2615–2635, https://doi.org/10.5194/hess-23-2615-2019, 2019.

Takoutsing, B. and Heuvelink, G. B.: Comparing the Prediction Performance, Uncertainty Quantification and Extrapolation Potential of Regression Kriging and Random Forest While Accounting for Soil Measurement Errors, Geoderma, 428, 116 192, https://doi.org/10.1016/j.geoderma.2022.116192, 2022.

Tarboton, D. G.: A New Method for the Determination of Flow Directions and Upslope Areas in Grid Digital Elevation Models, Water Resources Research, 33, 309–319, https://doi.org/10.1029/96WR03137, 1997.

Temmink, R. J. M., Lamers, L. P. M., Angelini, C., Bouma, T. J., Fritz, C., Van De Koppel, J., Lexmond, R., Rietkerk, M., Silliman, B. R., Joosten, H., and Van Der Heide, T.: Recovering Wetland Biogeomorphic Feedbacks to Restore the World's Biotic Carbon Hotspots, Science, 376, eabn1479, https://doi.org/10.1126/science.abn1479, 2022.

Torppa, J.: A Statistical Approach to Determining the Uncertainty of Peat Thickness, Mires and Peat, 8, 1–22, 2011.

UNEP: Global Peatlands Assessment – The State of the World's Peatlands: Evidence for Action toward the Conservation, Restoration, and Sustainable Management of Peatlands. Main Report., Tech. rep., United Nations Environment Programme, Nairobi, 2022.

Vaysse, K. and Lagacherie, P.: Using Quantile Regression Forest to Estimate Uncertainty of Digital Soil Mapping Products, Geoderma, 291, 55–64, https://doi.org/10.1016/j.geoderma.2016.12.017, 2017.

Wadoux, A. M. J.-C, Brus, D. J., and Heuvelink, G. B. M.: Sampling Design Optimization for Soil Mapping with Random Forest, Geoderma, 355, 113 913, https://doi.org/10.1016/j.geoderma.2019.113913, 2019.

Wadoux, A. M. J. C., Minasny, B., and McBratney, A. B.: Machine Learning for Digital Soil Mapping: Applications, Challenges and Suggested Solutions, Earth-Science Reviews, 210, 103 359, https://doi.org/10.1016/j.earscirev.2020.103359, 2020.

Wadoux, A. M. J. C., Heuvelink, G. B. M., de Bruin, S., and Brus, D. J.: Spatial Cross-Validation Is Not the Right Way to Evaluate Map Accuracy, Ecological Modelling, 457, 109 692, https://doi.org/10.1016/j.ecolmodel.2021.109692, 2021.

Wadoux, A. M. J.-C., Walvoort, D. J. J., and Brus, D. J.: An Integrated Approach for the Evaluation of Quantitative Soil Maps through Taylor and Solar Diagrams, Geoderma, 405, 115 332, https://doi.org/10.1016/j.geoderma.2021.115332, 2022.

Wang, L. and Liu, H.: An Efficient Method for Identifying and Filling Surface Depressions in Digital Elevation Models for Hydrologic Analysis and Modelling, International Journal of Geographical Information Science, 20, 193–213, https://doi.org/10.1080/13658810500433453, 2006.

Widyastuti, M. T., Minasny, B., Padarian, J., Maggi, F., Aitkenhead, M., Beucher, A., Connolly, J., Fiantis, D., Kidd, D., Ma, Y., Macfarlane, F., Robb, C., Rudiyanto, Setiawan, B. I., and Taufik, M.: PEATGRIDS: Mapping Thickness and Carbon Stock of Global Peatlands via Digital Soil Mapping, https://doi.org/10.5194/essd-2024-333, 2024.

Wright, M. N. and Ziegler, A.: Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R, Journal of Statistical Software, 77, 1–17, https://doi.org/10.18637/jss.v077.i01, 2017.

Wu, Q. and Brown, A.: 'whitebox': 'WhiteboxTools' R Frontend, 2022.

Xu, J., Morris, P. J., Liu, J., and Holden, J.: PEATMAP: Refining Estimates of Global Peatland Distribution Based on a Meta-Analysis, CATENA, 160, 134–140, https://doi.org/10.1016/j.catena.2017.09.010, 2018.