

# Wall-to-wall mapping of peat depth from Lidar terrain and airborne radiometrics in Norwegian landscapes

Julien Vollering<sup>1</sup>, Naomi Gatis<sup>2</sup>, Mette Kusk Gillespie<sup>1</sup>, Karl-Kristian Muggerud<sup>1</sup>, Sigurd Daniel Nerhus<sup>1</sup>, Knut Rydgren<sup>1</sup>, and Mikko Sparf<sup>1</sup>

<sup>1</sup>Department of Civil Engineering and Environmental Sciences, Western Norway University of Applied Sciences, Norway

<sup>2</sup>Department of Geography, University of Exeter, United Kingdom

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

**Abstract.** The abstract goes here. It can also be on *multiple lines*.

## 1 Introduction

Peat soils are a terrestrial carbon stock of global importance. They store 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). In other words, they are extremely carbon dense – denser than any other ecosystem per square meter (Temmink et al., 2022). This makes peatlands crucial to climate change mitigation. Intact peatlands sequester carbon and produce a negative temperature forcing, overall (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 lies in the third dimension: their depth. Peat soils range from zero to over ten meters deep, so deep peats comprise a large volume. Their depth results from the accumulation of organic matter 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. There it remains sequestered as long as the water table stays high and prevents oxidation.

Peatlands are unevenly distributed globally, 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. However, other land uses also displace peatlands. In Norway, where 9% of the land area is peatland (Bryn et al., 2018), construction and development have become important drivers of loss, after lawmakers restricted conversion to forest and farmland in recent decades.

The spatial distribution of peat depth is often overlooked in land use planning and carbon accounting because peat depth is not mapped with sufficient coverage, resolution, or accuracy. Maps are crucial because they link high-level targets to specific management decisions, unlike spatially aggregated estimates (OECD, 2022). For example, Norwegian regulations restrict conversion to farmland more strictly where peat depth exceeds 1 m, and the distribution of peat depth on a farmer's property determines whether he is allowed to convert peatland (Forskrift om nydyrkning, 1997, § 5a). Maps 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 over short distances – as in most peatlands. 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. This approach has grown important with the availability of remotely sensed data and the advancement of methods for identifying pattern, especially through machine learning (Minasny et al., 2019; Wadoux et al., 2020).

The crux for DSM of peat depth are the relationships between peat depth and the other, mapped variables. For DSM to work, 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. 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).

The most practical and widespread *scorpan* factors for peat depth 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. For example, most of mainland Norway has at least one elevation measurement per square meter from airborne LiDAR surveys. Moreover, mechanisms of peat formation are linked to topography. For example, a steep slope is unlikely to have a high water table, thereby limiting accumulation.

Studies have consistently shown relationships between peat depth and topography, although 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 (TWI) and the Multi-Resolution Valley Bottom Flatness (MRVBF) index, have shown associations with peat depth in some studies (e.g. Rudiyanto et al., 2018; Koganti et al., 2023; Li et al., 2024). Some of the variation between studies in quantifying these relationships is undoubtedly attributable to issues of spatial scale – both the scaling of the topographic variables and the resolution of the peat depth analysis.

Another set of variables 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 the earth's crust: potassium-40, uranium-238, and thorium-232 (Reinhardt and Herrmann, 2019). These exist in bedrock (and mineral soils) and a peat overburden attenuates the radiation intensity at the surface. The degree of attenuation relates to properties of the overburden, especially composition and depth. Deep soil with high water content attenuates most (Beamish, 2013; Reinhardt and Herrmann, 2019). Thus, gamma-ray radiometrics integrate the *scorpan* factors soil, parent material, and age.

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, presenting a big opportunity insofar as they are predictive. Airborne surveys are used for various applications, and some countries have high spatial coverage of such data (Minasny et al., 2019; Baranwal and Rønning, 2020).

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. There are many mire massif types, with different hydrology, formation, and development – from topogeneous or soligenous fens to blanket or raised bogs (Lyngstad et al., 2023). These types have fundamentally different geomorphology, suggesting that peat depth relationships with terrain or radiometric variables will vary by landscape.

The little we know about the depths of Norwegian peatlands comes from surveys meant to identify arable land. After scattered surveys in the early 20th century, a comprehensive round of surveying was completed 1964-2001 as part of a wider land cover mapping in Norway (Bjørndal, 2007). Because of its agricultural and silvicultural focus, it 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 one-meter probe, so they measured peat depth as shallow (< 1 m), or deep (> 1 m). These classes were assigned to whole mire polygons, so spatial resolution is on the order of hectares.

Soil science needs more studies of relationships between peat depth and topographic or radiometric data to determine how predictive they are, which variables are most predictive, and how consistent their associations. The push for nature-based climate solutions motivates for broad mapping of peat depth, to identify rich peatland carbon stocks and avoid their conversion. Although small areas can be mapped accurately on the ground, having landscape-scale maps *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, but by that time, the route of the road is already set. With a digital peat depth map, planners could better compare routes and climate mitigation measures would have more leverage.

We assess how well remotely sensed topographic and radiometric data can predict peat depth at the landscape scale, with a view towards revising regional and national maps in Norway. Specifically, we: (1) quantify the accuracy of predictions from topographic and radiometric variables, (2) identify key predictive variables, and (3) compare these relationships across two different Norwegian landscapes. To our knowledge, this is the first study to predict peat depth from airborne radiometric data using machine learning algorithms. Where airborne radiometric data have been used to predict 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 algorithms have been used on airborne radiometric data, it has been to predict peat extent rather than depth (e.g., O’Leary et al., 2022).

## 2 Materials and methods

### 2.1 Sites

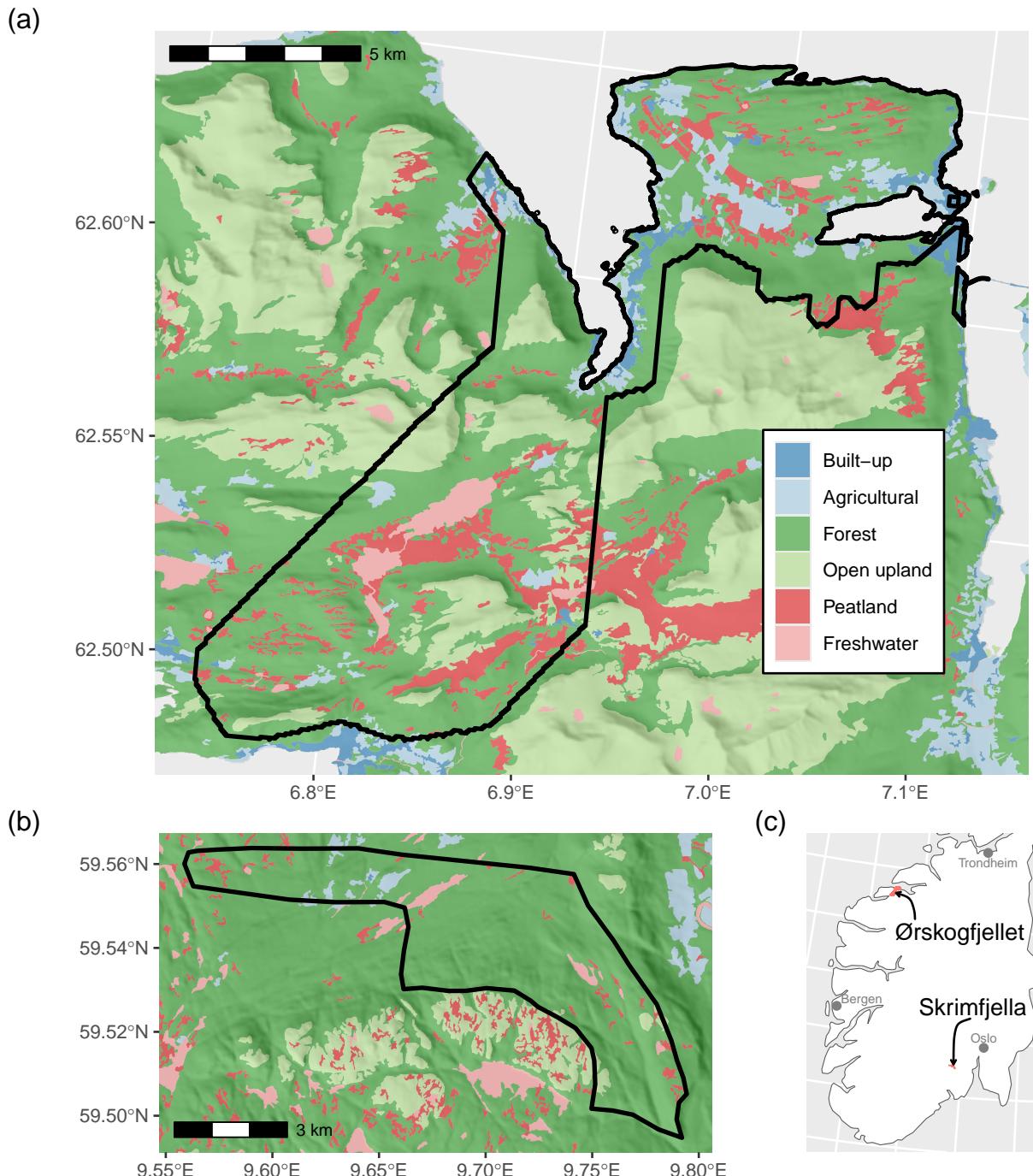
We assessed how well we could predict peat depth at two sites with conspicuously different physical geography: Skrimfjella in eastern Norway and Ørskogfjellet in western Norway (Fig. 1c). These sites were chosen because they were covered by radiometric data from airborne surveys, relatively little built-up area, and road access.

At Skrimfjella we delineated a study area of 34 km<sup>2</sup> based on radiometric coverage and accessibility (Fig. 1b). The study area has a diverse bedrock, with 32 % alkali feldspat granite, 26 % mergelstein, 10 % granite, and eight other types with > 1 % coverage (NGU, 1:250 000 dataset). The landscape within our delineation is classified as *inland hills and mountains* (Simensen et al., 2021). It is almost without human infrastructure, dominated by forest, and borders on a large nature reserve. The study area has a mean elevation of 438 m above sea level (range 223–711, IQR 351–509), and its mean slope at 10 m resolution is 10.8° (IQR 4.6–15.1°). In Norway's AR5 national land capability dataset (Ahlstrøm et al., 2019), 1.5 km<sup>2</sup> (4.5 %) of the study area is classified as mire – defined as areas with mire vegetation and at least 30 cm of peat depth.

At Ørskogfjellet we defined a study area of 124 km<sup>2</sup> which basically followed the footprint of the radiometric survey (Fig. 1a). According to the Geological Survey of Norway, bedrock in the area is 84 % granitic gneiss, 11 % granite, and 5 % aluminium silicate gneiss (NGU, 1:250 000 dataset). This study area comprises a wide range of major landscape types: *coastal plains*, *coastal fjord*, *inland valleys*, as well as *inland hills and mountains* (Simensen et al., 2021). It 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, IQR 73–310), and its mean slope at 10 m resolution is 13.0° (IQR 4.7–18.3°). The AR5 dataset counts 15.3 km<sup>2</sup> (12.4 %) of the study area as mire.

### 2.2 Peat depth measurements

At both study sites, our measurements of peat depth were made for the purpose of training a Random Forest (RF) model of peat depth, and we designed our sampling with this in mind (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). We chose for our sampling and modelling a spatial resolution of 10 m. We considered this a reasonable compromise between digital terrain model (DTM) resolution (1 m) and small mires on the one hand, and airborne radiometric resolution (50 m) on the other. The point measurements of peat depth described below were ultimately aggregated to 10 m resolution by taking the mean of point values within each cell, inversely weighted by their distances to the cell center.



**Figure 1.** Study areas at Ørskogfjellet (a) and Skrimfjella (b) within southern Norway (c). Land cover shown here is from the AR50 national land resource database and has simplified geometry with respect to the AR5 database used in the study.

### 2.2.1 Skrimfjella

We measured peat depth in selected locations (10 m raster cells) at Skrimfjella. The locations were chosen only from areas delineated as mire in the AR5 national land capability dataset. Within this mire area, we stratified our sample across values of elevation, slope, and potassium ground concentration (from processed airborne gamma ray spectrometry, Baranwal et al., 2013). Specifically, we used the *eSample* function in the *iSDM* R package (v.1.0) to chose an environmentally systematic sample. This function defines the environmental space as a two-dimensional convex hull around the 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). Elevation was extracted from the 10 m national DTM, slope calculated in degrees, and potassium ground concentration downsampled with bilinear resampling. We set a target sample size of 100, excluded the top and bottom percentile from the convex hull, and with these parameters *eSample* returned 105 raster cells.

In addition to the peat depth locations, we had another arm of our sampling design for measuring peatland occurrence, as a binary variable. We wanted to measure peatland occurrence outside of mapped mire areas because the AR5 dataset is known to underestimate peatland coverage (especially in forests, Bryn et al., 2018), and because airborne radiometrics may help identify unmapped peatland (Gatis et al., 2019; O’Leary et al., 2022). The occurrence locations were sampled from the part of the study area that (1) was mapped as something other than mire in the AR5 database and (2) had a slope < 20°. We performed environmentally systematic sampling of this population with the same procedure as for the depth locations, and *eSample* returned 106 raster cells.

Field work at Skrimfjella was conducted in August 2020. We navigated to the centers of the raster cells in the depth and occurrence samples by handheld GPS, checking that positional error was below 3 m. For each depth sample location, we measured peat depth three times (at the vertices of a triangle with 2 m sides) to get a more representative value for the 10 m raster cell, and to dampen the effect of outlying measurements (Parry et al., 2014). We used a metal probe pushed downward until resistance indicated 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. A single probe measurement was right-censored because the peat column was deeper than the probe. For each occurrence sample location, we recorded the presence or absence of peatland – primarily by digging and examining the top 20 cm of soil (where this was possible). We judged whether the soil was a peat soil based on its density, texture, and color. Occasionally, when the soil itself was difficult to judge, we made our determination also based on the presence or absence of mire vegetation. Although peat soil is strictly defined by organic content (which we did not analyse), we believe our protocol produced reasonable determinations of presence or absence that would generally satisfy most of the varying definitions of peatland (Minasny et al., 2023).

Besides the depth and occurrence measurements described above, we also measured peat depth in three subjectively-chosen, individual mires, using ground-penetrating radar (GPR). We used the Malå ProEx GPR system (Guideline Geo AB, Sweden) with its 500 MHz shielded antenna mounted in a plastic sledge, and its control unit connected to a GNSS receiver. At each of the three mires we recorded GPR traces along walking transects that covered the extent of the mire, mostly in traversing, zigzag patterns with between 5 m and 20 m spacing at their vertices. Along the GPR transects we also probed peat depth at marked

trace locations, to be able to calibrate the GPR wave speed velocity. We processed the GPR data with Reflex2DQuick software (v.3.0, Sandmeier Scientific Software, Germany), applying a time-zero correction, a dewow filter, and a gain filter based on observed energy decay. Then we picked strong reflectors in the radargrams that we interpreted as the base of the peat column. We used picks at marked trace locations to calibrate wave speed velocity; we pooled calibration points across the three mires and fitted a linear regression of depth on one-way travel time with the intercept fixed at zero. In total we had 46 calibration points along 3.5 km of GPR transects. Finally, we used the calculated wave velocity ( $0.0387 \text{ m ns}^{-1}$ ,  $R^2 = 0.874$ ) to convert the travel times of all picks to calibrated peat depths.

### 2.2.2 Ørskogfjellet

At Ørskogfjellet we also measured peat depth in a sample of 10 m raster cells, selected from the part of the study area classified in the AR5 dataset as mire. Before selecting locations, we determined a minimal sample size that would adequately capture the terrain and radiometric properties of the entire mire area. Specifically, we aimed to identify the size at which adding locations produced diminishing decreases in divergence between sample and population distributions – i.e. the elbow point in a curve of similarity between sample and population (Malone et al., 2019). This approach has been found to identify sample sizes that correspond with diminishing returns in predictive model performance on external evaluation data (Saurette et al., 2023). We defined a sequence of sample sizes (50–500) and for ten replicate samples at each size (drawn by conditioned latin hypercube sampling, Minasny and McBratney, 2006; Roudier, 2011), we calculated the mean Kullback-Leibler divergence between sample and population distributions (Malone et al., 2019; Saurette et al., 2023). The predictors in the divergence calculation were terrain slope and four radiometrics: potassium, thorium, uranium, and total count. Next, we fitted a asymptotic regression of mean divergence on sample size, and identified the sample size at which the curve reached 95 % of the fitted asymptote. Through this procedure we found that we could adequately capture the population distribution with a sample of 160 locations.

To choose 160 locations, we performed feature space coverage sampling. 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 aims to disperse samples as uniformly as possible in multidimensional predictor space, and is implemented by choosing locations that are closest to cluster centers in a k-means clustering of the standardized predictor space (Brus, 2019). 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: terrain slope and four radiometrics. The radiometric predictors were downscaled to 10 m resolution with cubic B-spline resampling in QGIS (v.3, QGIS Development Team, 2024). We adjusted the feature space coverage sampling to ensure that locations were accessible within time constraints, and 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 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). 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 % (with respect 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.

Field work at Ørskogfjellet was conducted in August 2023. We navigated to the centers of the raster cells in the sample using real time kinematic differential GNSS (Topcon Positioning Systems, USA), to ensure sub-meter positional accuracy. At each location we measured peat depth three times by manual probing, with probe locations spaced approximately 2.5 m apart. For a tiny fraction of these measurements (5 total), it was not possible to reach the base of the peat column, and a right-censored depth was recorded. In areas with dense sampling locations, we also measured peat depth with GPR along snaking transects passing through the centers of the sampling cells (seven transects, 6.2 km total length). We used the same GPR system as at Skrimfjella, but with a 100 MHz Malå rough terrain antenna (Guideline Geo AB, Sweden) at some transects. To navigate the GPR transects, we placed flags at the cell centers of the sample locations, and used a handheld GNSS receiver to guide the GPR operator. At sampling locations crossed by a GPR transect, we arranged the manual probe positions along the transect (for better calibration of the GPR wave speed velocity), while other locations were probed in a triangular pattern around the cell center like at Skrimfjella.

We processed the GPR data with Reflexw software (v.8.5, Sandmeier Scientific Software, Germany), applying a dewow filter, time-zero correction, bandpass filter, gain filter, and a dynamic correction that accounts for the non-vertical wave path between offset transmitter and receiver antennae. The last correction is important for the rough terrain antenna, which has an antenna separation (2.2 m) – comparable to typical peat depths. As with the data from Skrimfjella, we picked the base of the peat column from strong reflectors in the radargram, and calibrated wave velocity with manual probe measurements in a linear regression. The points in the regression were created by joining to each probe measurement the travel time of the nearest pick, but only if these were within 2 m of each other. In total we had 78 calibration points along 7.8 km of interpretable GPR traces (transect length exceeded because of extra GPR data). Finally, we used the calculated wave velocity ( $0.0427 \text{ m ns}^{-1}$ ,  $R^2 = 0.946$ ) to convert the travel times of all picks to calibrated peat depths.

We also used two sets of existing depth measurements from Ørskogfjellet. The first set was provided by the Norwegian Public Roads Administration, who commissioned GPR surveys of particular peatland areas in 2020 and 2021. The surveys were conducted with a dual channel system (70 MHz and 300 MHz; ImpulseRadar AB, Sweden), connected to GNSS with CPOS correction. We used interpreted and calibrated traces from these surveys, and discarded some data where multiple depths were interpreted for the same locations. This summed to 7.4 km of interpreted traces. The second set of existing depth data we extracted from a paper map made by the Norwegian Soil and Mire Company in 1984. This map presents 44 borehole depths (in decimeters) across a 9 ha peatland area. We georeferenced the map and digitized the borehole locations and depths.

## 2.3 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 surveyors (Geological Survey of Norway): ground concentration of Potassium, Thorium, Uranium, 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. Complete descriptions of predictors follow below.

### 2.3.1 Radiometric

The radiometric survey covering Skrimfjella was conducted in 2008–2011. The survey was flown at an average altitude of 75 m and average speed of  $108 \text{ km h}^{-1}$ , with flight lines spaced 200 m apart. Spectrometer count rates were calibrated annually to known concentrations of Potassium, Thorium, and Uranium in mobile pads. The Geological Survey of Norway processed data from the spectrometer following standard procedures outlined by the International Atomic Energy Association, and the 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. Further details about the survey and data processing are provided in Baranwal et al. (2013). We downscaled the processed data to 10 m resolution by cubic spline resampling, using the *terra* package in R.

A very similar radiometric survey covering Ørskogfjellet was conducted in December 2014 and January 2015. This survey was flown at an average altitude of 80 m and average speed of  $88 \text{ km h}^{-1}$ , with flight lines also spaced 200 m apart. Spectrometer count rates were calibrated in 2013 to known concentrations of Potassium, Thorium, and Uranium in mobile pads. The spectrometer data were processed following the same procedure as for the survey at Skrimfjella, except that a convolution filter was added to smooth the gridded data. Further details about the survey and data processing are provided in Ofstad (2015). The 10 m resolution predictors from this survey were identical to the layers used in the sampling design (resampled from 50 m resolution with cubic B-splines).

### 2.3.2 Terrain

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

We used the *terra* R package 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 from 1 m DTM resolution and then aggregated, or aggregated to 10 m DTM resolution and then calculated the indices. This kind of multiscale

**Table 1.** Candidate predictors of peat depth.

Name	Description
radK	Potassium ground concentration
radTh	Thorium ground concentration
radU	Uranium ground concentration
radTC	Total count of gamma radiation
elevation	Mean elevation
slope1m	Mean of 1 m slope
TPI1m	Mean of 1 m topographic position index
TRI1m	Mean of 1 m terrain ruggedness index
roughness1m	Mean of 1 m roughness
slope10m	10 m slope
TPI10m	10 m topographic position index
TRI10m	10 m terrain ruggedness index
roughness10m	10 m roughness
MRVBF	Multi-resolution valley bottom flatness
TWI5m	Mean of 5 m topographic wetness index
TWI10m	10 m topographic wetness index
TWI20m	Bilinear interpolation of 20 m topographic wetness index
TWI50m	Bilinear interpolation of 50 m topographic wetness index
DTW2500	Depth-to-water index, flow initiation area of 0.25 ha
DTW5000	Depth-to-water index, flow initiation area of 0.5 ha
DTW10000	Depth-to-water index, flow initiation area of 1 ha
DTW20000	Depth-to-water index, flow initiation area of 2 ha
DTW40000	Depth-to-water index, flow initiation area of 4 ha
DTW80000	Depth-to-water index, flow initiation area of 8 ha
DTW160000	Depth-to-water index, flow initiation area of 16 ha
DMK	DMK peat depth class, categorical with 3 levels

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 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, Conrad et al., 2015) with default parameters.

Next, we calculated the Topographic Wetness Index (Quinn et al., 1991). This index 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, 2016), 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 last terrain-based predictor we included was the depth-to-water index (Murphy et al., 2007). This index approximates a location's vertical height above the surface water feature 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 the Whitebox software. Also 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 catchments area layer was then thresholded at seven different levels (*flow initiation areas* from 0.25 ha to 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 the depth-to-water index 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.

### 2.3.3 Peat depth class

We prepared one categorical predictor – peat depth class – from a historical national map dataset called *DMK* (Ahlstrøm et al., 2019). The DMK peat depth classes are: < 1 m (*shallow*), > 1 m (*deep*), and *unknown*. Mappers 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.

## 2.4 Predictive models of peat depth

### 2.4.1 Modelling approach

We used RF to predict peat depth at both sites. RF is a tree-based 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 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 data sets and its predictions can be interrogated to learn about predictor importance (Khaledian and Miller, 2020). Evaluating variable importance in a maximally-predictive model aligns with the aim of this study.

RF by itself is not a spatial model, and it will only predict spatial structure in the outcome to the degree that the structure is captured by predictors. We considered using regression kriging – a hybrid between non-spatial and spatial techniques that would be achieved by adding 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 compared to a non-spatial model (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 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 rather regression kriging 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 (v.0.16, Wright and Ziegler, 2017). RFs were fit with 1000 trees, minimum node size of 5, and the number of predictors 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).

#### 2.4.2 Model performance

For both sites we compared the performance of models with five different configurations of three groups of predictors. Specifically, we trained models with:

1. only DMK peat depth class (2 indicator variables from 3 levels)
2. only terrain (21 continuous variables)
3. terrain and DMK peat depth class (23 variables)
4. terrain and radiometric predictors (25 continuous variables)
5. all predictors (27 variables)

These different configurations simulate different scenarios of data availability that are common in Norway or internationally. Comparing the different configurations allowed us to isolate the added value of each of the predictor groups. We did not explore configurations comprising radiometrics without terrain, because Lidar terrain surveys typically precede airborne radiometric surveys. The models with only DMK peat depth class were simple linear models rather than RFs, and 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.

The performance of DSM must be evaluated with reference to a specific purpose (i.e., map vs. model validation, interpolation vs. extrapolation, Roberts et al., 2017; Milà et al., 2022), and here we aimed to evaluate maps of peat depth across the study

areas. In the absence of additional field work to collect a design-based independent validation set, we used a spatial cross-validation scheme to evaluate model performance (Wadoux et al., 2021; Meyer and Pebesma, 2022). Specifically, we used k-Means Nearest Neighbor Distance Matching (kNNDM), which creates cross-validation folds that mimic the spatial prediction task that is defined as the goal (Linnenbrink et al., 2024). In particular, kNNDM looks for the spatial assignment of training data to folds 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 (v.1.0.2, Meyer et al., 2024), setting prediction locations to all AR5 mire cells in the study area, and choosing a number of folds (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).

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). Moreover, the quality of uncertainty estimates should be evaluated, just as the quality of predictions are (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 predictor configuration that showed the highest performance at each site (under the assumption that these models would be put into production) and we extracted 90 % prediction intervals.

While the primary aim of the DSM was to predict peat depth within mapped peatlands (where peat depth > 30 cm), we also performed a minimal evaluation of our ability to estimate peat depth or occurrence outside of mapped peatland. In other words, we tested whether our peatland-trained models of peat depth could indicate which other locations were most likely to have peat. It is generally impractical to obtain good data for this purpose, because locations with non-zero peat depth outside of mapped peatland are unknown and relatively rare. For Skrimfjella, we had independent data from the arm of our sampling design that measured peatland occurrence outside of mapped mire areas. For Ørskogfjellet we did not have such data. Therefore, we first – for both sites – performed an evaluation with just our peat depth measurements, which included a fraction of locations not mapped as peatland. For this evaluation we used the same spatial cross-validation folds as previously, but now we trained the models only on locations mapped as peatland (in the training folds), and evaluated them only on locations not mapped as peatland (in the test fold). In this way, we simulated the situation where a model trained on known peatland is used to predict peat depth outside the known peatland. From the cross-validation we quantified mean absolute error. Second – for Skrimfjella – we simply trained a model on all peat depth measurements and then used the independent occurrence dataset to calculate the area under the curve of the receiver operating characteristic, to evaluate the model's ability to discriminate between peat presence and absence. In this analysis, the model's prediction of peat depth is treated as a rank index of peatland

likelihood. Since DMK peat depth class is always undefined outside known peatlands, we evaluated models with two predictor configurations: terrain predictors alone or terrain and radiometric predictors together.

#### 2.4.3 Model interpretation

We quantified global variable importance and examined partial dependence plots for the best-performing predictor configuration at each site. Global variable importance measures the influence of a given predictor on the output of the model, aggregated across all locations. Partial dependence plots depict the shape of the fitted relationship between a given predictor and the outcome. 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 measures (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 above 0.7. Thus, highly-correlated sets of predictors are represented by a single variable for the purposes of model interpretation.

We calculated variable importance with the *vip* R package (v.0.4.1), 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). *Permutation* values were obtained from ten iterations, with RMSE as the performance measure.

We calculated partial dependence with the *pdp* R package (v.0.8.1, Greenwell, 2017). For the top six most important variables, we plotted both partial dependence and individual conditional expectation (ICE), to show the average effect of the predictor on the outcome and the variation in the effect across observations, respectively (Goldstein et al., 2015). Non-parallel ICE lines indicate the presence of interactions between predictors.

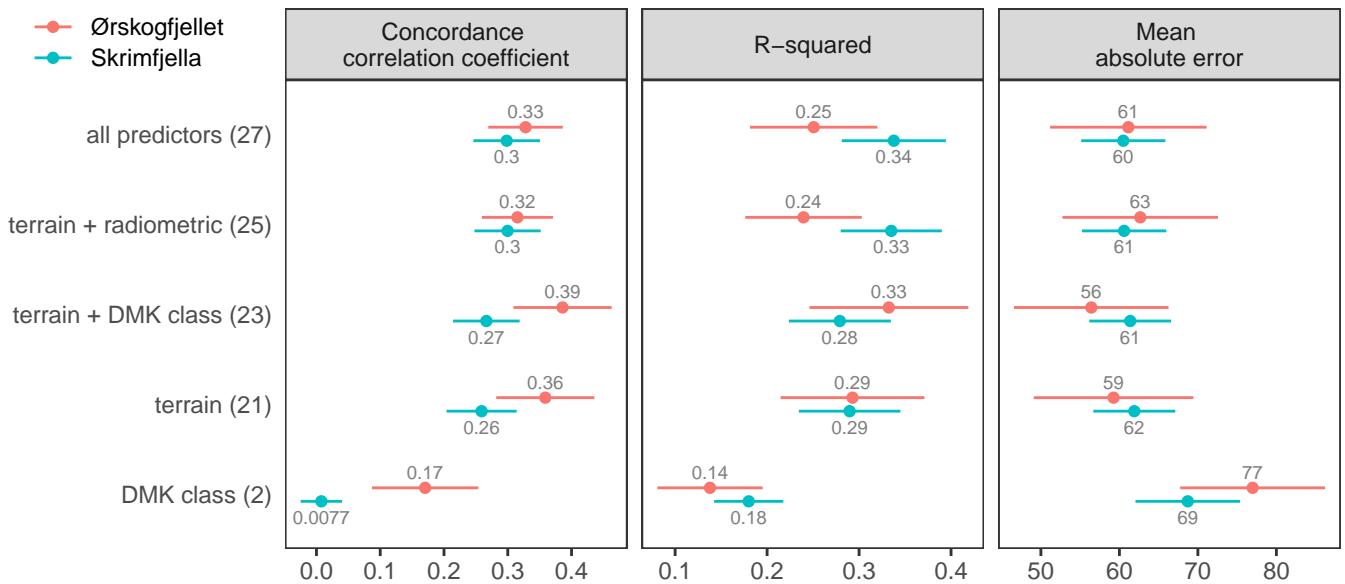
### 3 Results

We obtained depth measurements for 372 cells at Skrimfjella (area equivalent to 2.4 % of mapped peatland) and 1878 cells at Ørskogfjellet (area equivalent to 1.2 % of mapped peatland). Roughly 80 % of these 10 m cells were within mapped peatland, and the remainder in forest, open upland, or farmland (Table 2). Coverage of DMK peat depth classes was higher at Ørskogfjellet than at Skrimfjella (79 % versus 27 % of cells), and at Ørskogfjellet the *deep* and *shallow* classes showed a larger difference in measured depth. Overall mean peat depths were similar at Skrimfjella and Ørskogfjellet: 119 cm and 126 cm, respectively.

None of the models were able to predict peat depth across the study areas with high accuracy (Fig. 2). For Skrimfjella the best model achieved a concordance correlation coefficient of 0.3, an  $R^2$  of 0.34, and a mean absolute error of 60 cm. For

**Table 2.** Mean peat depth (cm) in 10 m cells at Skrimfjella and Ørskogfjellet. The cells are also shown stratified by AR5 land class and DMK peat depth class.

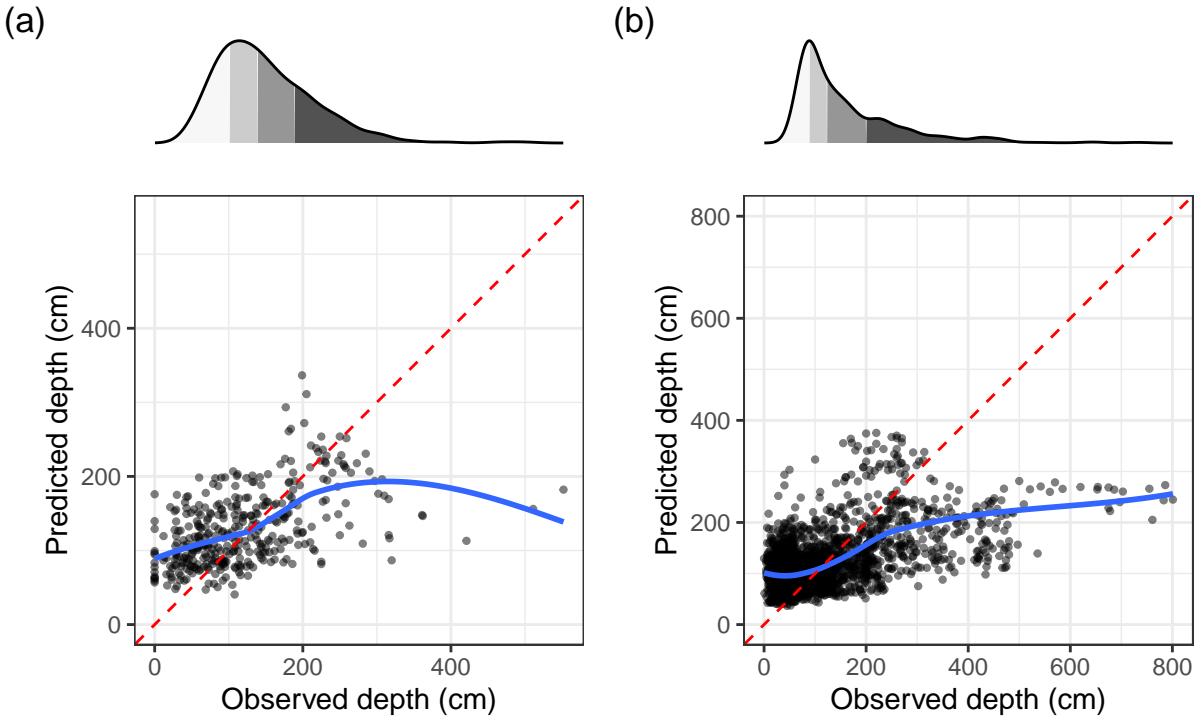
		Skrimfjella			Ørskogfjellet		
		n	%	depth	n	%	depth
		372	100	119	1878	100	126
AR5	Agricultural				21	1	30
	Forest	52	14	65	272	15	71
	Open upland	19	5	98	134	7	39
DMK	Peatland	301	81	130	1451	77	145
	deep (>100 cm)	94	25	100	659	35	219
	shallow (<100 cm)	6	2	50	838	45	82
	unknown	272	73	127	381	20	60



**Figure 2.** 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.

Ørskogfjellet the same values were 0.39, 0.33, and 56 cm, so the best model at Ørskogfjellet was slightly more accurate than the best model at Skrimfjella. These values were derived from kNNMDM spatial cross-validation with 20 folds at Skrimfjella and 10 folds at Ørskogfjellet.

Although the difficulty of predicting peat depth caused prediction intervals to be wide, these uncertainty estimates were well-calibrated. At Skrimfjella the prediction interval coverage probability was 91 %, and at Ørskogfjellet it was 84 % (both



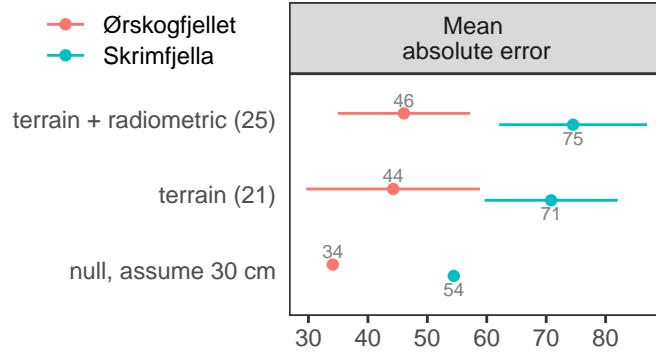
**Figure 3.** Calibration plots for the best-performing models at Skrimfjella (a) and Ørskogfjellet (b), with predictions from spatial cross-validation. Blue smoother lines are local polynomial regressions. Marginal distributions are shaded by quartile.

compared to the target value of 90 %). Observations outside of the prediction intervals were evenly distributed across each of the study areas.

For Skrimfjella, the best predictor configuration was *all predictors*, followed closely by *terrain and radiometric*. The performance gap between the *terrain and DMK* configuration and the *terrain-only* configuration was similarly small. Compared to the *terrain-only* configuration, the *terrain and radiometric* configuration improved concordance correlation by 0.04,  $R^2$  by 0.04, and mean absolute error by 1 cm. *DMK class* alone – although calibrated to measured depths – was a very poor predictor of peat depth, with a concordance correlation coefficient of 0.008.

For Ørskogfjellet, the best predictor configuration was *terrain and DMK*, followed by *terrain-only*. Adding radiometric predictors to these configurations worsened model performance, especially in terms of concordance correlation and  $R^2$ . By itself, *DMK class* produced a concordance correlation coefficient of 0.17 (compared to 0.008 at Skrimfjella), but the worst mean absolute error of any model at either site: 77 cm.

The best models at both sites overpredicted shallow peats and strongly underpredicted very deep peats (Fig. 3). The mean error (bias) of these models was 10 cm at Skrimfjella and -4 cm at Ørskogfjellet.



**Figure 4.** Performance of models that extrapolate from training data inside of mapped peatland to test data outside of mapped peatland, evaluated via spatial cross-validation. Parentheses denote the number of variables in each model, and point estimates are shown +/- their standard error.

When we tested how well models extrapolated from within to outside of mapped peatland, we found that the models produced worse mean absolute error than just assuming a constant 30 cm depth (Fig. 4). With the independent occurrence data at Skrimfjella, we found that neither the *terrain-only* nor *terrain and radiometric* configurations were able to discriminate between peat presence and absence (area under the curve of the receiver operating characteristic 0.44 and 0.52 respectively, where 0.5 indicates random guessing).

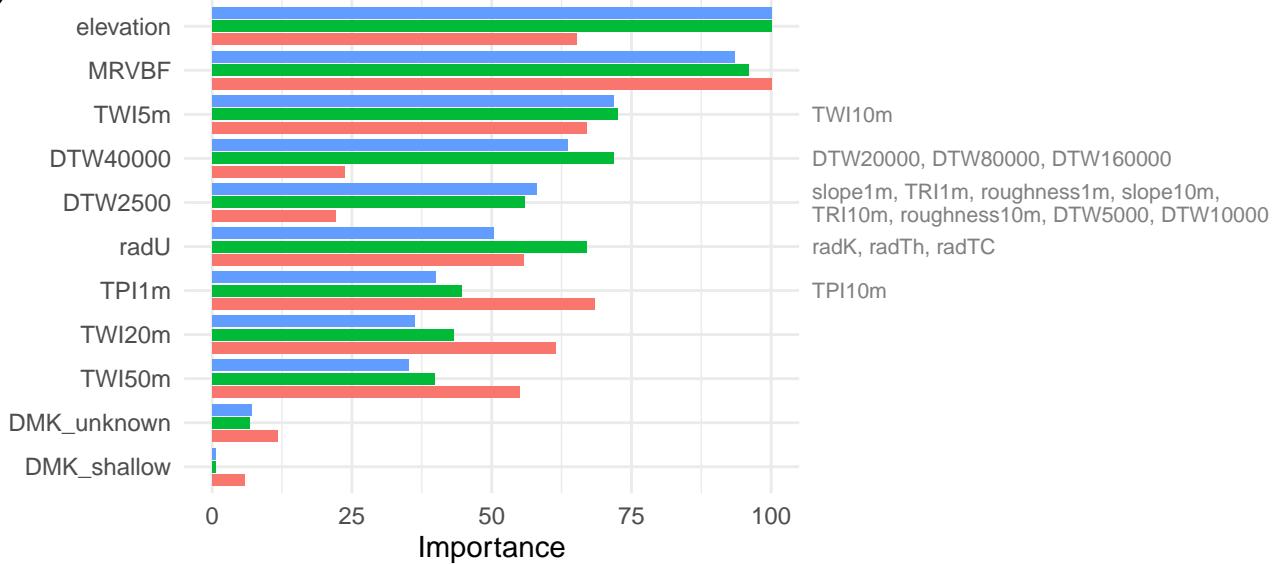
For the purpose of model interpretation, the *all predictors* configuration for Skrimfjella was reduced from 27 variables to 11 non-collinear variables. Similarly, the *terrain and DMK* configuration for Ørskogfjellet was reduced from 23 variables to 11 non-collinear variables. At both sites, *elevation* and *MRVBF* were important predictors (Fig. 5). At Skrimfjella these two predictors were of similar importance, while at Ørskogfjellet *elevation* was more important than *MRVBF*. *DMK* was also important – the shallow class in particular – but only at Ørskogfjellet. Some realizations of the hydrological predictors *TWI* and *DTW* showed considerable importance, while others showed little – with no consistency between sites. For example, *TWI5m*, *DTW40000*, and *DTW2500* rounded out the top five most important variables at Skrimfjella, while *TWI20m* and *TWI50m* were least important, other than *DMK*. At both sites, the most important realizations of hydrological predictors were more important than the simple terrain indices *slope*, *TRI*, *TPI*, and *roughness*. The radiometric variable *radU* showed moderate importance at Skrimfjella.

Many of the 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 meters above sea level, after which a further increase was predictive of shallower peat. At Skrimfjella the partial dependence on *elevation* had the opposite shape, with the shallowest peats predicted at intermediate elevations, around 350 meters above sea level. *TWI50m* at Ørskogfjellet and *DTW4000* and *radU* at Skrimfjella were other predictors that showed considerable fluctuations in their predictive effects on peat depth. The radiometric predictor in particular displayed an idiosyncratic effect, with a marked dip in predicted depth at intermediate values of *radU*. On the other hand, the partial effects of some important predictors were

(a)



(b)



**Figure 5.** Global variable importance in the best-performing models at Ørskogfjellet (a) and Skrimfjella (b), as measured by three different metrics. Variables removed due to collinearity are shown to the right of that with which they are most correlated.

more straightforward. The partial dependence on *MRVBF* was quite similar across sites, with the deepest peats predicted in the very flattest valley bottoms. Also, *TWI5m* and *DTW2500* at Skrimfjella showed monotonically positive and negative predictive effects, respectively.

Individual conditional expectation lines indicated some interactions between predictors (Fig. 6). For example, the magnitude of the increase in depth with *elevation* that the model expected at Ørskogfjellet was different for different observations; some depth predictions increased by only 40 cm while others increased by more than 100 cm, over the same elevation gain. Similarly, ICE lines of *radU* at Skrimfjella were non-parallel, with some locations showing monotonically increasing peat depth predictions with uranium concentration (unlike the average effect). Nevertheless, most ICE lines were generally parallel – indicating that the average effects of the predictors were good representations of their overall effects.

## 4 Discussion

### 4.1 Can we improve Norway's peat depth maps?

#### 4.1.1 Remotely sensed variables are weak (but not useless!) predictors of peat depth

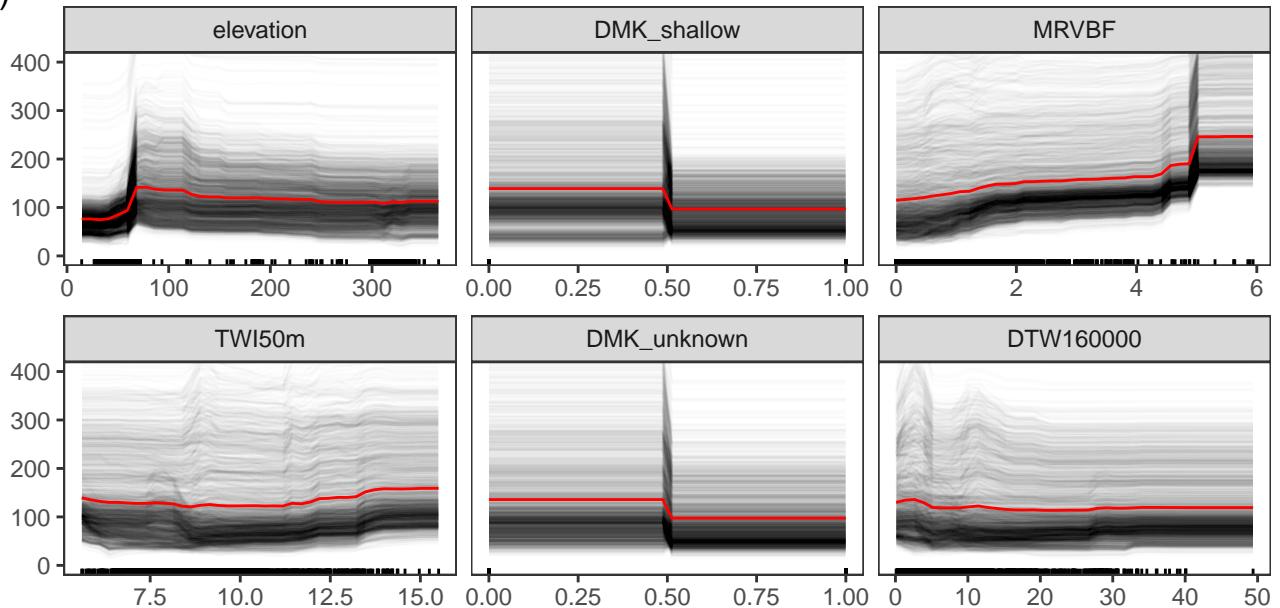
Our ability to predict peat depth in the study sites based on terrain and radiometric data was limited. Mean absolute errors of 60 and 56 cm at the two sites — relative to mean depths of 119 and 126 cm — illustrate the practical limitations of these maps. Since any given 10 m cell will miss by about 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.

On the other hand, we improved on the best available map of peat depth (DMK depth class), which 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. Since remotely sensed data are widely available, improvements to soil maps as shown here are low hanging fruit. This point is recognized and reflected in the rise of DSM.

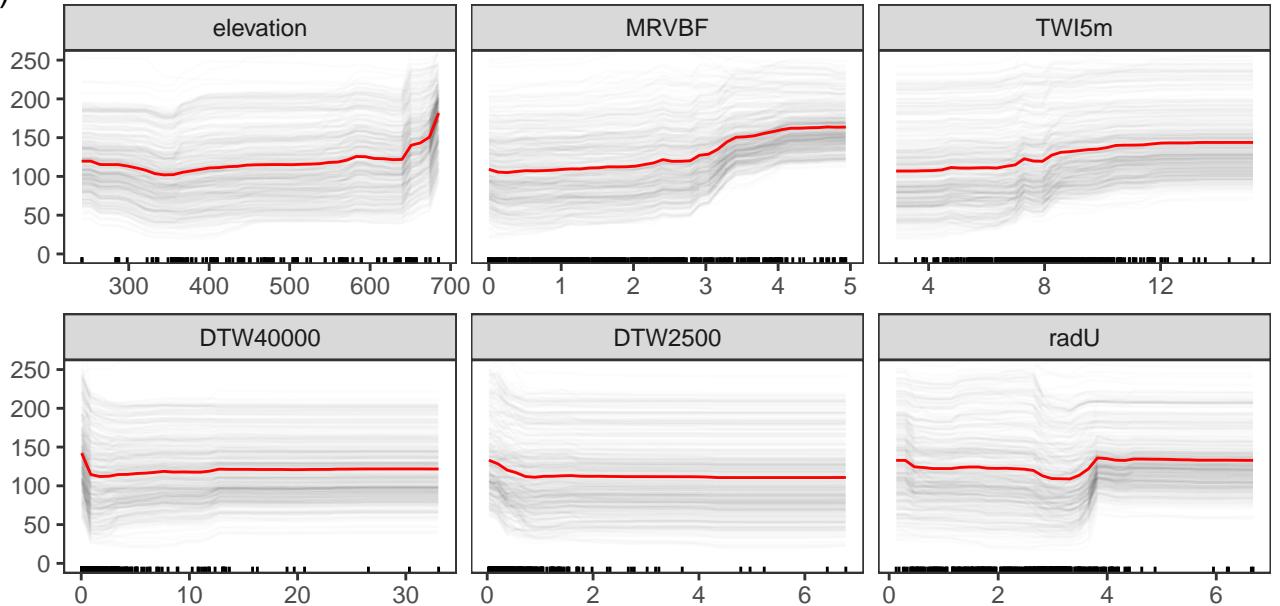
DMK peat depth classes were a worse predictor of peat depth than our models even though we calibrated them with the same data. That is: we gave DMK and the DSM equal footing for a fair comparison. If we had taken the DMK peat depth classes at face value and assumed depths according to their class definitions (< 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 varying from 0-8 m at fine spatial scales.

The performance of our terrain- and radiometric-based maps could have been improved with a spatially explicit mapping approach like regression kriging. By ignoring spatial autocorrelation in peat depth (i.e. the  $n$  in scorpan), we have not extracted all of the information about peat depth in the study area out of the training data. If we were to put our DSM approach into production for published maps, we would harness the spatial component, but the actual maps for Ørskogfjellet and Skrimfjella are not of primary interest in this study. Moreover, residuals of the RF predictions showed weak spatial structure at Skrimfjella,

(a)



(b)



**Figure 6.** Partial dependence plots of the six most important variables in the best-performing models at Ørskogfjellet (a) and Skrimfjella (b). The average effect of the predictor on the outcome (red line) overlays the variation in the effect across observations (grey lines).

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 but generally produced better or comparable error magnitude ( $R^2$  vs. MAE/RMSE; 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. More standardized reporting would help but not eliminate this consideration. 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).

Gatis et al. (2019) used similar predictors and the same spatial grain, finding a much stronger correlation 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. 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 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 radK alone predicted peat depth with much higher concordance than our models ( $CCC = 0.76$ ) and that elevation alone produced comparable performance ( $CCC = 0.27$ ). The RMSE 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 (RMSE = 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.

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

Our models erred most for the deepest peats. Where overprediction occurred, it was smaller. This is not unexpected for the right-skewed distributions of peat depth, but it has management implications. Identifying the deepest peats will require additional field work in candidate areas, which could be defined by an upper quantile of predicted depth. Map users should

not trust the maps to identify all large carbon stocks, but they can trust that identified large stocks really are large. That makes the map more suited for “red-lighting” than “green-lighting” peatland conversion, for example. Where it does not prohibit conversion (i.e. predicts shallow peat), a ground survey should be done before conversion is allowed (assuming depth is the only consideration). Although this recommendation aligns with the precautionary principle, here we make it on technical grounds based on the maps’ characteristics.

#### 4.1.4 Depth predictions do not necessarily extrapolate to peat occurrence

That we measured more than 30 cm of peat in areas not mapped as peatland was expected, because AR5 underrepresents peatland extent by about one third (Bryn et al., 2018). For that reason, maps of peatland occurrence also need revising. Unfortunately, our models did not predict peat occurrence outside mapped peatlands. In these areas, assuming a constant 30 cm of peat (the depth threshold used in AR5) produced less error than our models. Moreover, the performance gap would likely have been larger had our evaluation data been representative of the non-peatland areas. Instead they were near mapped peatlands. Most importantly, the independent occurrence data at Skrimfjella showed the model completely failing to predict occurrence, and these data were suited for testing this ability. Although the Ørskogfjellet model may have done better than the Skrimfjella model on a similar independent evaluation set (since it predicted depth better), the improvement would probably be small.

Why wasn’t peat occurrence predicted well? The models were trained on fundamentally different populations of locations (~80 % mapped peatland in training vs. 0 % in prediction), so it is not surprising that the associations they learned did not transfer well. Moreover, the ~20 % of training data from outside mapped peatland were incidentally collected (near peatland) and not representative of other land cover classes. In short, the models were blind to the fact that most of both study areas have no peat.

Peatland extent is probably best mapped using different remotely sensed data than we used here (Bakkestuen et al., 2023), and this study’s purpose was not to map extent. Nevertheless — as long as the relevant peatland definition includes a depth component — depth predictions (or predictors with subsurface information) should help delineate peatland extent (O’Leary et al., 2022; Beamish and White, 2024). We return to this point in our discussion of implications for digital soil mapping.

### 4.2 Which variables predict peat depth?

#### 4.2.1 Airborne radiometrics do not predict Norwegian peat depth

Radiometric data had no predictive value at Ørskogfjellet, while at Skrimfjella they had minor influence in a relatively weak model. The bedrock is more homogeneous at Ørskogfjellet than at Skrimfjella, so uneven radiogenesis is not a viable explanation for the differences between sites nor the poor performance in general (Beamish, 2014; Reinhardt and Herrmann, 2019). To the extent that radiometrics had predictive value at Skrimfjella, it appears that they were most valuable near the extremes of the depth distribution, since their inclusion improved  $R^2$  more than mean absolute error. All four 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 radTC was a much better predictor than radK.

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 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 like peatlands (Kock and Samuelsson, 2011; Karjalainen et al., 2024). Karjalainen et al. (2024) 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), likely due to the extensive blanket bog landscape with more gradual changes in depth (Lindsay, 1995; Lyngstad et al., 2023). We are unsure whether short-range depth changes explain the weak associations that Siemon et al. (2020) found in a large raised bog.

Weather conditions varied during the Ørskogfjellet radiometric survey and affected its data (Ofstad, 2015). Thus, uneven snow cover and air moisture may also have masked the soil signal in these data.

We do not believe that the poor predictive value of the radiometric data in this study was caused by fully attenuated radioactivity — at least not in large part. 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 radU at Skrimfjella, the expected negative relationship between depth and radU cannot be found by ignoring the left (highly-attenuated) side of the distribution. About a quarter of the peats in our study were less than a meter 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. Elevation was the most important predictor at both sites, and peat depth showed non-monotonic responses to changes in elevation. 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. Indeed, elevation at Ørskogfjellet seems to interact with other variables (non-parallel ICE lines), complicating its interpretation. Moreover, 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 upper marine limit after the last ice age. At Ørskogfjellet, the upper 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 upper 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. We cannot evaluate this effect at Skrimfjella, where the upper marine limit is below our study area (at 175 m.a.s.l.).

Another influential terrain-based predictor was MRVBF. Unlike elevation, it showed a monotonic effect on peat depth: greater valley bottom flatness was always associated with increases in peat depth. Delineating a valley bottom involves ambiguity, but the MRVBF 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 multiresolution 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 MRVBF 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 index). Notably, terrain slope was inferior to (Ørskogfjellet) or highly correlated with (Skrimfjella) these hydrological indices. Mappers of peat depth should not assume that slope is the best predictor in its class, despite its prevalence in the literature. Wetter locations (high TWI and low DTW) were generally associated with deeper peat, but these relationships were not as strong or consistent as with MRVBF. The optimal scale for Topographic Wetness Index and Depth-To-Water index 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 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. At Ørskogfjellet, it increased performance more, and both indicator variables were among the most important in the model. 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. 73 % of cells measured at Skrimfjella had unknown depth in DMK, compared to 20 % at Ørskogfjellet. Still, the 25 % of cells classified as deep at Skrimfjella were not meaningfully deeper than the rest, according to our model. Interactions between DMK and other variables

underline the inconsistency of DMK depth maps, even within a site. That a peat at Ørskogfjellet was classified as shallow rather than deep did not uniformly lower the model's expectation of depth there. Rather, the veracity of the deep-indication was contingent on other factors.

### 4.3 Implications for digital soil mapping of peat depth

The performance gap between the best models and the DMK-only models shows that peat depth in Norway should be mapped digitally. Anywhere we have some calibrating measurements, we can get better maps than DMK peat depth classes, at low cost. Moreover, DSM methodology can align map products with open science principles by making their production transparent, reproducible, and updatable. The large difference we found between the coverage and quality of DMK peat depth at Skrimfjella versus Ørskogfjellet underlines these advantages. With DSM we can apply the same approach across different areas and make maps with full spatial coverage, continuous values, and validated uncertainty. The rest of this section discusses recommendations and needs for more extensive peat depth DSM.

#### 4.3.1 Peat depth measurements should be organized

High-quality digital terrain models are available for mainland Norway, making peat depth measurements the critical training data need. For a given area, some minimum number of measurements is necessary to create meaningful improvement over DMK peat depth. The proportion of peatlands sampled at Skrimfjella was twice that at Ørskogfjellet, but the Skrimfjella model performed worse, which shows that the size of the depth dataset is not all-important. We had the luxury of stratifying our measurements over candidate predictors, and performance may suffer where locations are opportunistic or tied to a sampling design with a different purpose. The magnitude of this penalty will depend on the dataset characteristics, but having enough depth measurements is probably more consequential (Wadoux et al., 2019).

Depth measurements are foundational, and better infrastructure to collect, preserve, and make FAIR these data would help DSM and other applications. Geoportal access and data exchange standards (like Natural England's for peat surveys, 2023) are important. Peat data often fall through the cracks between geology-oriented and ecology-oriented archives, but increased awareness of peatland importance is a good impetus to remedy this situation. Peat depth is quick and easy to measure, so integrating its measurement into existing national field programs, like Norway's spatially representative nature monitoring (ANO) or national forest inventory, would be helpful (although not sufficient for regional or local mapping). Municipalities and local actors in Norway are increasingly measuring peat depth and can contribute to a growing data foundation (Kyrkjeeide et al., 2023). 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). All of the above can lay the groundwork for renewed peatland maps.

### 4.3.2 Spatial scale affects model performance and utility

The scarcity of depth measurements and their short spatial autocorrelation mean that mapping peat depth at broad scales is not an exercise in spatial interpolation (Hengl et al., 2004). Peat depths in our GPR data showed spatial autocorrelation to a range of 50-100 m, and getting measurements at such fine grain is only realistic for small areas, not across landscapes. Therefore, we anticipate that spatially-explicit DSM is of limited value in all but the most intensively sampled landscapes (currently nowhere in Norway). Fine-scale variation in peat depth raises the possibility of mapping at very fine resolution (1 m) — even if land use planning and carbon accounting do not operate at this grain. Terrain-depth relationships might be stronger in 1m cells than in our 10 m cells, considering the hummock-hollow microtopography of many peatlands.

Choosing a spatial extent for DSM can be tricky. A natural starting point is the bounding area around a spatial cluster of depth measurements. For a given cluster, kNNDM can define how far the boundary can extend beyond the point measurements; if a cross-validation nearest neighbor distribution cannot match the sample-to-prediction nearest neighbor distribution, the boundary is too expansive (Meyer and Pebesma, 2022; Linnenbrink et al., 2024). However, it is unclear how big (N) any cluster of depths should be, and 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, where the composition of peatland types is constant (e.g. Moen, 1998). Depth varies strongly between peats formed by terrestrialization versus paludification (Buffam et al., 2010), and probably along other axes of peatland typology. Although most Norwegian landscapes contain multiple, unknown types, it makes sense to limit their diversity within a single DSM.

### 4.3.3 Machine learning approaches can build on success

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, we may move from pixel-based learners to deep learning approaches like convolutional neural networks (CNNs). Bakkestuen et al. (2023) successfully predicted peatland extent using CNNs, and their ability to automatically learn multi-scale spatial features would reduce the need to manually engineer these. MRVBF's success as a predictor of peat depth demonstrates that multi-scale spatial patterns matter for peat depth and CNNs are designed to learn such patterns. 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 CNN could learn. However, since this approach is data-hungry, we should build soil knowledge into the DSM where we can (Minasny et al., 2024). For example, if further research confirms the effect of the upper marine boundary that we found, then it is better to include the marine boundary as a predictor than to make the algorithm learn this pattern from elevation independently.

#### 4.3.4 Peat occurrence and depth should be mapped together

Finally, we want to highlight the need for research on peatland occurrence and peat depth to be better integrated. Since peatland occurrence/extent requires non-zero peat depth (the specific threshold varies by definition, Minasny et al., 2023), they are fundamentally linked. The goal, therefore, should be a unified prediction framework for occurrence and depth. We caution against reducing continuous depth predictions to arbitrary classes (as in Ivanovs et al., 2024; Karjalainen et al., 2024), since classes can be derived from continuous predictions. The distribution of peat depths across landscapes is zero-inflated, and research is needed to determine whether it is more efficient to parameterize a single model of peat depth (with a larger, generalized dataset) or to break down the problem by classifying zero depth and then regressing non-zero depths (with smaller, specialized datasets). Coupling occurrence and depth will reduce the prevalence of incoherence that we found: deep peat outside the peatland extent and zero depth inside it. A key challenge going forward will be obtaining training data that represents both zero and non-zero components of the depth distribution, since sampling designs often focus on known peatlands.

- . *Code and data availability.* Use this to add a statement when having data sets and software code available
- . *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.
- . *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 Institute of Bioeconomy Research (NIBIO), © Geovest, and (3) the Norwegian License for Public Data (NLOD), made available by the Geological Survey of Norway (NGU). Large language models have 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

- 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, e00255, <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, Hønefoss, Kongsberg and Numedalen., Tech. Rep. 2013.029, Geological Survey of 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](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 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.
- Beamish, D. and White, J. C.: On the Detailed Mapping of Peat (Raised Bogs) Using Airborne Radiometric Data, *Journal of Environmental Radioactivity*, 277, 107462, <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.
- 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 von Moorböden, in: *Jahrestagung der DBG 2015: Unsere Böden - unser Leben*, München, 2015.
- Bjørndal, 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, 116781, <https://doi.org/10.1016/j.geoderma.2024.116781>, 2024.
- Breiman, L.: Random Forests, *Machine Learning*, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>, 2001.

- 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, <https://doi.org/10.1029/2009JG001034>, 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 Plain of Montreal, *Canadian Journal of Soil Science*, 103, 103–120, <https://doi.org/10.1139/cjss-2022-0031>, 2023.
- Dornik, A., Chețan, M. A., Drăguț, L., Dicu, D. D., and Iliuță, A.: Optimal Scaling of Predictors for Digital Mapping of Soil Properties, *Geoderma*, 405, 115453, <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.
- 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.
- 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., Grätzl, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., 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.
- Gallant, J. C. and Dowling, T. I.: A Multiresolution Index of Valley Bottom Flatness for Mapping Depositional Areas, *Water Resources Research*, 39, <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.
- 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., Schmidlein, 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.
- Heuvelink, G. B. M. and Webster, R.: Spatial Statistics and Soil Mapping: A Blossoming Partnership under Pressure, *Spatial Statistics*, 50, 100639, <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, 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.
- 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., 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*, 0, 1–14, <https://doi.org/10.1080/14942119.2024.2398947>, 2024.
- 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 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.
- Kyrkjeeide, M. O., Fandrem, M., Kolstad, A. L., Bartlett, J., Cretois, B., and Silvennoinen, H. M.: A Calculator for Local Peatland Volume and Carbon Stock to Support Area Planners and Decision Makers, *Carbon Management*, 14, 2267018, <https://doi.org/10.1080/17583004.2023.2267018>, 2023.
- Lamichhane, S., Kumar, L., and Wilson, B.: Digital Soil Mapping Algorithms and Covariates for Soil Organic Carbon Mapping and Their 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, 117009, <https://doi.org/10.1016/j.geoderma.2024.117009>, 2024.

- Lindsay, J. B.: Whitebox GAT: A Case Study in Geomorphometric Analysis., *Computers & Geosciences*, 95, 75–84, 2016.
- Lindsay, R.: Bogs: The Ecology, Classification and Conservation of Ombrotrophic Mires, Technical report, Scottish Natural Heritage, Edinburgh, ISBN 9781853971006, 1995.
- 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., Fialkiewicz-Koziel, B., Finkelstein, S. A., Galka, 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., 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., Välimäki, 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.
- 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.
- 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, 114366, <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.
- Marchant, B. P.: Using Remote Sensors to Predict Soil Properties: Radiometry and Peat Depth in Dartmoor, UK, *Geoderma*, 403, 115232, <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, 103944, <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/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(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 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.

- 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.
- 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*, <https://doi.org/10.1007/s10533-023-01084-1>, 2023.
- 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 Widystuti, M.: Soil Science-Informed Machine Learning, *Geoderma*, 452, 117094, <https://doi.org/10.1016/j.geoderma.2024.117094>, 2024.
- Moen, A.: Vegetasjonstyper, in: *Vegetasjon, Norges geografiske oppmåling*, Hønefoss, ISBN 978-82-90408-26-3, 1998.
- 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.
- 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](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 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.
- 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.
- 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, 614–624, <https://doi.org/10.1111/j.1475-2743.2012.00447.x>, 2012.
- Pelletier, R., Davis, J., and Rossiter, J.: Peat Analyses In The Hudson Bay Lowlands Using Ground Penetrating Radar, 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.
- 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.
- QGIS Development Team: QGIS Geographic Information System, QGIS Association, 2024.
- 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.
- 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, e2021WR029871, <https://doi.org/10.1029/2021WR029871>, 2021.
- 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.
- Roudier, P.: Clhs: A R Package for Conditioned Latin Hypercube Sampling., 2011.
- Rudiyanto, Minasny, B., Setiawan, B. I., Arif, C., Sapomo, 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.
- Rudiyanto, Minasny, B., Setiawan, B., Sapomo, 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*, pp. 1–12, <https://doi.org/10.1017/jog.2023.83>, 2023.
- 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.
- Simensen, T., Erikstad, L., and Halvorsen, R.: Diversity and Distribution of Landscape Types in Norway, *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography*, 75, 79–100, <https://doi.org/10.1080/00291951.2021.1892177>, 2021.
- 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, 116192, <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.
- 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, 113913, <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, 103359, <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, 109692, <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, 115332, <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.

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.