Peat depth and carbon storage of the Hudson Bay Lowlands, Canada

Alemu Gonsamo¹, Yiyao Li¹, Daorui Han¹, Cheryl A Rogers¹,², Sarah A Finkelstein³, Oleksandra Hararuk⁴, James M Waddington¹, Carlos Barreto⁵, James W Mclaughlin⁵, and James Snider⁶

¹School of Earth, Environment & Society, McMaster University
²Department of Geography and Environmental Studies, Toronto Metropolitan University
³Department of Earth Sciences, University of Toronto
⁴Northern Forestry Centre, Canadian Forest Service, Natural Resources Canada
⁵Ontario Ministry of Natural Resources and Forestry
⁶World Wildlife Fund Canada

June 13, 2024
Peat depth and carbon storage of the Hudson Bay Lowlands, Canada

Yiyao Li¹, Daorui Han¹, Cheryl A. Rogers¹,², Sarah A. Finkelstein³, Oleksandra Hararuk⁴, James M. Waddington¹, Carlos Barreto⁵, James W. McLaughlin⁶, James Snider⁶, and Alemu Gonsamo¹

¹School of Earth, Environment & Society, McMaster University, Hamilton, ON, Canada,
²Department of Geography and Environmental Studies, Toronto Metropolitan University, Toronto, ON, Canada,
³Department of Earth Sciences, University of Toronto, Toronto, ON, Canada,
⁴Northern Forestry Centre, Canadian Forest Service, Natural Resources Canada, Edmonton, Canada,
⁵Ontario Ministry of Natural Resources and Forestry, Sault Ste. Marie, ON, Canada, and
⁶World Wildlife Fund Canada, Toronto, Ontario, Canada

Correspondence to: gonsamo@mcmaster.ca

Key points:
- Peat depth increases with distance further away from the coast in Hudson Bay Lowlands (HBL)
- The average peat depth of HBL is 184 (±48) cm with 90% of depths falling within 87-264 cm
- HBL stores 30 (±6) Pg carbon (C) with average value of 86 (±35) kg m⁻²

Abstract
The Hudson Bay Lowlands (HBL) are recognized as the second largest peatland complex in the world. Due to variability in peat thickness across a large and heterogeneous landscape, the existing carbon (C) storage estimates for the HBL may contain large uncertainty. Here, we use geospatial variables that are associated with HBL peat formation, age, accumulation, and occurrence to understand the driving factors for peat depth variability and map peat depth and total C storage at 30m spatial resolution. The estimated average peat depth of HBL is 184 (±48) cm with 90% of values falling between 87 and 264 cm. Based on the spatially explicit peat depth, the HBL total C storage is estimated to be 30 (±6) Pg. Distance to the coastline is the most important indicator of peat depth where the depth increases with distance further away from Hudson Bay coastline, suggesting that the time since peat formation is closely related to peat depth.

Plain Language Summary
The Hudson Bay Lowlands (HBL) contain the second largest peatland complex in the world. We used spatial data sourced from satellite observations and geospatial information that are associated with peat occurrence, age, formation, and accumulation to estimate peat depth and carbon storage at 30mx30m spatial details for the entire HBL. The estimated average peat depth was 183.5 cm while the entire HBL stores 30 billion tonnes of carbon. The peat depth and carbon storage presented in this study will help improve our understanding of the developmental processes and ecosystem functions of the HBL peatland.
1. Introduction

Northern peatlands store large terrestrial carbon (C) stocks, contributing substantially to the global C cycle and influencing feedback mechanisms in the Earth’s climate system (Frolkin et al., 2007; 2011; Helbig et al., 2020). Approximately 85% of global peatland C storage is concentrated in the northern high-latitude regions with an estimated 415 ± 150 Pg C, where favorable conditions of low temperatures and high precipitation promote peat accumulation (Hugelius et al., 2020; Xu et al., 2018; Yu et al., 2010). High-latitude regions are warming much faster than other areas of the world, resulting in permafrost thaw, more frequent and severe wildfires, altered microbial dynamics, and hydrological and vegetation changes (Seward et al., 2020; Sim et al., 2023; Treat et al., 2021; Zhang et al., 2018). In turn, these changes will affect the peatland C storage and sink capacity (Wilkinson et al., 2023). Therefore, accurate methodologies to map, and monitor peatland C storage and its driving factors have great importance to understanding how northern high latitude ecosystems may respond to climate change.

The Hudson Bay Lowlands (HBL) are recognized as one of the largest peatland complexes globally (Yu et al., 2012). More than 90% of the HBL by area is covered by peatland (Martini, 1989), containing an estimated range of 30 Pg to 38 Pg C in peat soils (Hugelius et al., 2014; Packalen et al., 2014; Sothe et al., 2022a). The HBL are a key component of the northern peatlands, and the large C storage can have a considerable impact on the global climate by affecting C fluxes through both its carbon sink function and its source functions, the latter through methane emissions (Davies et al., 2021; Packalen et al., 2014). The estimates of the total C storage within HBL need further investigation due to low spatial coverage of ground measurements of peat depth and C densities across most of this large and heterogeneous region.

Peat depth is a crucial determinant of the developmental process, ecosystem functions, and the C storage capacity of peatlands, as deeper peat layers contain up to 90% of the soil C reserves that are missed in surveys that often sample the upper 20 cm to 50 cm depth (Clymo et al., 1998; Yu et al., 2010). Gaining a better understanding and accurately measuring peatland C stocks in the HBL are vital given the expected increases in road networks, resource extraction, development, drainage, as well as climate change and increased wildfires (Wilkinson et al., 2023; Finkelstein et al., 2023). However, the vast remote landscape contributes to high field sampling costs and presents a great challenge for data collection in the HBL.

Traditional soil mapping proves to be expensive and time-consuming due to the need for numerous field observations or manual mapping through aerial photography, and the outcomes tend to be often subjective (Minasny et al., 2019; Rudiyanto et al., 2018; Rathnayaka & Vitharana, 2016). In past studies, other methods such as Kriging interpolation (Akumu & McLaughlin, 2014), ground-penetrating radar (GPR, Gerber et al., 2010), and LiDAR (Chasmer et al., 2016; Millard & Richardson, 2013) have been used for peatland mapping in Canada, but they still require large amounts of field data for calibration, and have limitations in terms of areal coverage and performance in complex terrain (Gerber et al., 2010; Rudiyanto et al., 2018; Minasny et al., 2019). Compared to traditional soil mapping, Digital Soil Mapping (DSM) combines limited field observations with spatial and non-spatial information to generate spatially explicit soil property maps (McBratney et al., 2003). This may include developing and verifying statistical models using georeferenced soil data and environmental variables to predict and map
soil properties with associated uncertainties at various spatial resolution (Béguin et al., 2017; Hugelius et al., 2014; Sothe et al., 2022a; Sothe et al., 2022b; Xu et al., 2018). DSM can help produce large-scale peatland maps more efficiently, cost-effectively, and accurately, especially over large geographical areas and regions as inaccessible to humans as HBL (Kempen et al., 2012; Minasny et al., 2019). Although there are some past studies that mapped peatland through DSM techniques in Canada, few studies have focused on peat depth, which is a key variable to accurately estimate the total C stock (Hugelius et al., 2014, 2020; Sothe et al., 2022a). The peat depth map generated by Hugelius et al. (2020) for northern peatlands lacks adequate field measurements for the HBL and the 10km x 10km spatial resolution is too coarse to capture the spatial heterogeneity of the HBL landscape which is a mosaic of interconnected bogs, fens, swamps, marshes, limited uplands, rivers, water tracks and pools (Riley, 2011; Westbrook, 2014).

In this study, we use the largest available dataset of peat depth measurements, and geospatial variables that are associated with peat formation, age, and accumulation as well as satellite observations that indicate peat occurrence to map peat depth and total C storage of HBL peatland at 30m spatial resolution. Considering the huge size of HBL, the difficulty to access and the high cost of field surveys, our ground measurements covered most of the past research and is the most detailed dataset in the study of HBL so far. Specifically, we investigate the driving factor of peat depth variability, the spatial distribution of peat depth and C stock, and the total C storage of the HBL peatland complex. We use a “stacked regression” (or “stacking”) (Breiman, 1996; Fryda et al., 2023; Wolpert, 1992), a multi-model ensemble machine learning strategy, to improve model performance and prediction accuracy for generalization. This study fills major gaps in our understanding of peat depth distribution, their driving factor and peat formation processes and contributes to refined methodologies for mapping and monitoring the C stock of the globally important HBL peatland complex.

2. Materials and methods

2.1. Study area

The HBL is a vast plain covering 376,880 km² area, located southwest of James Bay and Hudson Bay in the center of Canada (Martini, 2006; Packalen et al. 2014) (Figure 1). The HBL is about 1,400 km in length and 540 km wide. The northern part of the lowlands experiences a maritime subarctic climate, while the southern and inland areas belong to the boreal climate zone (Dredge & Dyke, 2020). HBL is marked by short, cool summers and long, very cold winters with mean annual temperatures ranging from −7°C in northwest to −1°C in southeast and mean precipitation varying from about 400 mm in the northwest to 800 mm in the southeast. Because of climatic conditions and relatively low topographic variations, peat has accumulated in nearly 90% of the HBL, which accounts for nearly 10% of the pan-Arctic peatland area (Dredge & Dyke, 2020; Packalen et al., 2016). The two major peatland classes in the HBL are bog and fen where fens are more dominant in the coastal areas (Packalen et al., 2016).
Figure 1. Hudson Bay Lowlands (HBL) and distribution of peat depth ground measurements. The background image is a digital elevation model (DEM) from the Shuttle Radar Topography Mission (SRTM). Inset: the distribution of peat depth from ground measurements where vertical lines indicate the mean and ± one standard deviation.

2.2. Data
In this study, ground-measured peat depth data serve as the dependent variable, while three groups of geospatial covariates related to peat depth variations are the independent variables (Table S1). Group 1 includes two covariates associated with peat formation and age: Distance to coastline (Distbay003) and Distance to nearest river (Distriver003). Group 2 includes two covariates associated with peat accumulation: Shuttle Radar Topographic Mission Digital Elevation Model (SRTM DEM) and Topographic Wetness Index (TWI). Group 3 includes thirteen covariates related to peat occurrence and abundance, derived from long-term satellite observations of land surface temperature, greenness, and polarization signatures in Synthetic-Aperture Radar (SAR). Details of each dataset are provided below, and example maps of covariates are provided in Figure S1.

2.2.1 Field measurements of peat depth
Field measurements of peat depth data were obtained from the Ontario Ministry of Natural Resources and Forestry data archive that is compiled from previously published literature (e.g., Glaser et al., 2004; Gorham et al., 2012; O’Reilly et al., 2014; Packalen et al. 2014; Riley, 2011). We retained 495 data records that have location, peat depth and C stock information out of 504 total data records. This data, with a peat depth range of 24 cm to 450 cm, mean depth of 205 cm and standard deviation of 82.8 cm (Figure 1) was used for model training. For model validation, we collected independent peat depth information using a Russian-type peat corer for 32 sites spanning from 51°N to 55°N in July and September 2022. The mean peat depth of validation data was 129.2 cm and standard deviation was 53.3 cm.

2.2.2 Geospatial proxies for peat formation, age, and accumulation

The distance to the nearest open water bodies such as rivers and coastal areas in HBL is related to ongoing glacial isostatic adjustment since ice retreat began and ongoing uplift that is continually producing new land areas, rivers and peats (O’Reilly et al., 2014; Packalen et al., 2016; Rudiyanto et al., 2018). The distances to open water bodies are important indicators of peat development processes that affect peat depth such as peat formation, age and accumulation through isostatic adjustment since the last glacial period (O’Reilly et al., 2014; Packalen et al., 2014). We combined all HBL rivers into one spatial dataset and calculated minimum Euclidean distance to the nearest river for each pixel in the study area at 30m spatial resolution. Additionally, we calculated the minimum Euclidean distance to Hudson and James Bay coastlines.

Since the surface of bogs are often convex where a peat dome is present (Clymo et al., 1998), geospatial information such as elevation, and other terrain attributes have been previously reported as indicators of peat accumulation that determine the distribution of peat depth (Packalen et al., 2016; Rudiyanto et al., 2018). We use the 1 arc-second digital elevation model (DEM) derived from the Shuttle Radar Topography Mission (SRTM) observations. As our study area is covered by scattered trees only in the southern portion, the SRTM DEM provides accurate information about the surface topography of HBL. Topographic Wetness Index (TWI) was calculated from the DEM data as an additional terrain attribute that may influence peat accumulation. TWI is a function of slope and the upstream contributing area and indicates the direction of water flow and soil moisture pattern (Gatis et al., 2019; Qin et al., 2011).

2.2.3 Satellite proxies for peat occurrence

Surface vegetation type and abundance can be affected by the depth of the water-saturated substrate, hydrology, nutrient availability, and substrate chemistry, which are inherently related to peat depth (Rydin and Jeglum, 2013). Additionally, peat depth can affect land surface temperature directly through moisture retention and indirectly through its effects on vegetation communities, hydrology, and thus land surface albedo (Brown and Wilson, 2009). To represent these processes, we compiled a large satellite dataset that captures vegetation amount and type, surface and vegetation structure, and land surface thermal properties. To represent vegetation amount and type, we used median and standard deviation values of Enhanced Vegetation Index (EVI) from Sentinel-2 satellite sensor data, collected at a 10 m spatial resolution during the growing season months (May to September) and the full year. To represent peat moisture,
surface and vegetation structure, we use data acquired by Sentinel-1A Synthetic-Aperture Radar (SAR) between 2015 and 2022. From Sentinel-1A SAR 10 m data with vertical (V) and horizontal (H) polarization, we extracted the long-term mean values of VV, VH, the ratio of VH to VV, the difference between VV and VH, and the average of VV and VH. The interaction of a SAR wave with peatland can result in changes in backscatter type and polarization depending on the amount of peat moisture and the structure and density of surface vegetation. The ratio and difference of VV and VH backscatter is proven to respond to changes in vegetation height and surface roughness very well and insensitive to changes in water content in vegetation and peat (Greifeneder et al., 2018). To represent surface thermal properties, we used land surface temperature (LST) observations from the Landsat 8 satellite sensor. From long-term (2013 to 2021) Landsat 8 30 m observations, we extracted median and standard deviation LST values for the growing season months (May to September) and the full year. Missing LST data due to cloud cover were filled with temperature observations from the Tier I Landsat 8 data archive. Standard deviations from long-term satellite observations are used to represent variability as a function of peat depth. For example, deeper peats generally have higher moisture content, resulting in a higher heat capacity and lower thermal conductivity than shallower peats, leading to lower variations in long-term LST. This means deeper peats have lower LST standard deviations than shallower peats.

2.3 Model training and prediction

In summary, we compiled 17 spatial covariates that are related to peat depth (Table S1). The spatial resolution for all raster datasets is 30 m or less. Raster datasets with 10 m resolution such as Sentinel 2 and Sentinel 1 were resampled to 30 m. Prior to modelling peat depth using the 17 covariates, we performed variable selection using the Boruta algorithm, an all relevant feature selection wrapper to identify the important variables in the dataset and remove variables that are less relevant after the statistical test (Kursa and Rudnicki, 2010). While the TWI variable was found to be not relevant for peat depth mapping, the remaining 16 covariates showed statistically significant (p<0.05) importance for peat depth mapping (Figure S2).

Data was trained using multi machine learning algorithms. Based on the Root Mean Squared Error (RMSE) derived from 10-fold cross-validation, four models were selected for further prediction, including Gradient Boosting Machine (GBM), Deep Learning, Distributed Random Forest (DRF) and Extreme Gradient Boosting (XGBoost). For the final peat depth estimation, we further applied a second-level “meta-learner” called stacked regression, an ensemble learning strategy aiming to find an optimal combination of the individual base models to get a final prediction with better performance and generalization ability (Breiman, 1996; Fryda et al., 2023; Wolpert, 1992). Here we used generalized linear model (GLM) during the stacking process to map peat depth for the entire HBL (Fryda T et al., 2023). The uncertainty was estimated as ± one standard deviation around the mean estimates of all base models. Model training and prediction was conducted using the ‘h2o’ package (v3.40.0.4) in R programming language platform (Fryda T et al., 2023).

Previous studies have shown that there is a strong linear relationship between C stock and peat depth in HBL peatlands (Packalen et al., 2016; McLaughlin et al., 2021). Pixel level C stock is estimated based on empirical relationship between the estimated peat depth and C stock ($C_{mass}$ (kg m$^{-2}$)) from the 495 model training ground measurements that resulted in the following regression equation:
C_{mass} (\text{kg m}^{-2}) = 0.38 \times \text{peat depth (cm)} + 17.14

This linear equation ($R^2 = 0.96$; p-value $<0.000$) was applied to the estimated HBL peat depth and its uncertainty maps to estimate the HBL C stock pixel-by-pixel. Finally, open water bodies, coastal sand banks, tidal flats and exposed rock outcrops were masked from the HBL peat depth and C stock maps.

3. Results

3.1 Model performance for peat depth estimation

For the final peat depth prediction, we used 16 covariates (see Table S1) and the stacked ensemble from the four base models, namely, GBM, Deep Learning, DRF and XGBoost. Among the four base models, the GBM shows the best performance on model training data with the lowest Root Mean Squared Error (RMSE) and highest $R^2$ followed by XGBoost. DRF and XGBoost have a comparable performance on training data (Table S2 and Figure S4). The ranking of variable importance for the four base models is shown in Figures S2 and S3. Based on frequency, the three most important variables are Distance to coastline, elevation, and ratio of VH to VV (Table S2 and Figure S3). Other variables such as mean annual land surface temperature and EVI, difference between VV and VH, annual standard deviation of EVI and Distance to the nearest river also appeared as the top five important variables for the four base models (Table S2).

![Figure 2](image.png)

Figure 2. Comparison of the predicted and measured peat depth (cm) using independent validation dataset. (a) Scatter plot of the predicted and measured peat depth (cm) from ensemble model output. (b) Box plot showing the comparison of measured and estimated peat depth from ensemble and 4 base models. The performance of the four base models on the training dataset is given in Figure S5.
To evaluate the performance and accuracy of the peat depth model, we used the independent validation dataset collected from 32 locations in HBL. The results indicate that the ensemble model performs better (highest $R^2$, Figure 2) than the four base models (Figure S5) indicating the stacking approach is indeed the best for applying trained model on independent data. A fundamental goal of stacking in machine learning is generalization itself, the ability to draw accurate inferences on unseen data from a model or models trained on a given. Although we observe better model performance on training data from the four base models (Figure S4) compared to ensemble estimate (Figure 3), the ensemble estimate (Figure 2) performs better than the four base models (Figure S5) when applied on an independent validation data that was not used for model training.

Figure 3. Comparison of the predicted and measured peat depth (cm) using all training data. (a) Scatter plot of the predicted and measured peat depth (cm) from ensemble model output. (b) Box plot showing the comparison of measured and estimated peat depth from ensemble and 4 base models. The performance of the four base models on the training dataset is given in Figure S4.

3.2 Peat depth and carbon storage of the Hudson Bay Lowlands

We mapped the entire HBL peatland area by excluding no-peat pixels, covering 350,459 km$^2$, using the final stacked ensemble model. The range of peat thickness was from 0 to 600 cm with a mean of 184 ($\pm$48) cm and standard deviation of 53.93 cm (Figure 4A). The thickness was described by five classes: 0-40 cm, 41-100 cm, 101-200 cm, 200-300 cm and above 300 cm (Anda et al., 2021). The group of 0-40 cm indicates non-peat pixels, which accounts for 1,422 km$^2$ (0.4%) of the entire HBL. These are mostly located along the coasts of James and Hudson Bays where land recently emerged and are initially dominated by mineral soil wetlands, gradually transitioning to peatlands several km inland (Packalen et al., 2016; Riley, 2011). Shallow peat (41-100 cm) constitutes 7.4% of HBL with 25,770 km$^2$ area while moderate peat (101-200 cm) dominates much of HBL peatlands with 183,971 km$^2$ which is 52.5% of the entire HBL. Deep
peat (201-300 cm) concentrated in the inland area covering 38.7% of HBL with 135,632 km². The very deep peat (301-600 cm) constituting 1% of HBL with 3,663 km² is found near the margin of the Boreal Shield where the oldest peatlands are located (Davies et al., 2022).

The very deep peat (301-600 cm) constituting 1% of HBL with 3,663 km² is found near the margin of the Boreal Shield where the oldest peatlands are located (Davies et al., 2022).

The estimated average C stock based on peat depth is 86 (±35) kg m⁻² with standard deviation of 20 kg m⁻². Since the relationship between peat depth and C stock is linear, the spatial distribution of C stocks follows that of peat depth (Figure 4). High C stocks are found in inland areas with the highest stocks concentrated in the southwest part of HBL near the margin of the Boreal Shield where the oldest peatlands are located. The estimated total peat C stock for the entire HBL is 30 (±6) Pg C. High uncertainty in both peat depth and carbon (C)
stock estimates often coincides, albeit not exclusively, with deeper peats in inland areas of the HBL (Figure 4 C and D). We also find that areas with higher elevation (Figure 1) and farthest areas from coastline and rivers (Figure S1 A and B) show relatively higher uncertainty in peat depth and C stock estimations. Note that the uncertainty estimates in this study are related to disagreements among models and do not encompass all sources of uncertainty.

4. Discussion and Conclusions

This study presents the use of a stacked ensemble modeling approach, integrating multiple geospatial and satellite observations, to estimate peat depth and consequently carbon (C) stock of the Hudson Bay Lowlands (HBL). By incorporating long-term and multi-source covariates pertinent to the study area and reflective of peat formation processes, coupled with a choice of modeling algorithms, our aim was to enhance prediction performance and accuracy. The ensemble predictions derived from the stacked regression, a technique involving the amalgamation of multiple base models, were anticipated to match or surpass the performance of the most proficient base model included in the ensemble (Breiman, 1996; Wolpert, 1992). However, the primary advantage of stacking lies in its ability to mitigate overfitting risks while simultaneously making the stacked ensemble robust across diverse data types. Our analysis corroborates this notion, as evidenced by the superior performance of the ensemble estimate when compared against independent peat depth data that was not used during model training. Notably, the ensemble outperformed any individual base model (Figure 2 and S5) on validation data.

We found that the distance to the coastline is the primary factor explaining peat depth variability in the HBL — peat depth increases as the distance from the Hudson Bay coastline increases. This is explained by the fact that peat ages are much younger in coastal than inland areas due to ongoing glacial isostatic adjustment since ice retreat in the Middle Holocene where ongoing uplift is continually producing new land areas and new peatlands (Glaser et al., 2004). Packalen et al. (2014) reported that coastal shallow peats formed approximately 1,900±1,830 cal years BP, while inland deeper peats formed around 5,220±1,450 cal years BP. Therefore, inland peatlands have been accumulating organic matter for much longer than the coastal areas resulting in deeper peats. This is further supported by the fact that elevation appeared as an important contributor for peat depth variability, which ranked second in importance in most models (Table S2). Although elevation has been shown to have a linear relationship with peat depth, especially domed structures in previous studies (Dredge & Dyke, 2020; Rudiyanto et al., 2015; Silvestri et al., 2019), the HBL is a flat mosaic of heterogeneous landscapes without prominent peat domes. Rather in the HBL, elevation increases with distance from the coastline and related to peat age and peat accumulation, both of which affect peat depth.

The high contribution of the ratio of VH and VV from Sentinel 1 SAR observations is due to its ability to capture surface roughness and local topography as bare flat surfaces have a weak depolarizing effect, while vegetation canopies and sloping terrain are highly depolarizing (Greifeneder et al., 2018; Rudiyanto et al., 2018). Similarly, the standard deviation of long-term land surface temperature (LST) was also found to be an important variable, particularly for deep learning model (Table S2). Peat depth influences water retention and surface moisture levels (see Waddington et al., 2015) and deeper peat layers typically exhibit higher thermal inertia, leading to lower LST variability than shallow peat layers. Long-term mean and standard deviation of EVI, a proxy of photosynthetic biomass are also important indicators of peat depth variability.
(see Table S2) because peat accumulation depends greatly on the dominant vegetation community types (e.g., Sphagnum moss peat, herbaceous fen peat, treed peat) which influences plant productivity (Laine et al., 2021) and peat decomposition rate (Turetsky et al., 2008).

The estimated average peat depth of the HBL is 184 cm with 90% of peat depths falling between 87 cm and 264 cm, and 99% of peat depths falling between 0 cm and 300 cm. These results are plausible for HBL as the range of peat depth from the 495 ground measurements used for model training is between 24 cm and 450 cm. We also observed very deep peat that ranged between 300 cm and 600 cm in depth, the latter being the maximum depth mapped but this range only made up 1% of the HBL. The average peat depth also depends on the specific peatland class with bogs generally having a higher peat depth than fens (Packalen et al., 2016). For example, the average peat depth in the HBL is typically 230 cm for bogs and 160 cm for fens (Bysouth and Finkelsein, 2021; Packalen et al., 2016).

Peat deposits increase from the coast to the inland and from northwest to southeast (Figure 4). The northwest part of HBL belongs to a continuous permafrost area with the coldest and driest climate within the HBL, which tends to experience lower net primary productivity as well as slower decomposition of soil organic matter due to lower temperatures (Packalen et al., 2016). The southwest part of HBL is the most continental region with the strongest seasonality, warmest surface temperature, greatest precipitation, and the highest rates of biomass production, which contribute to peat development. Coupled with time since peat initiation, which gets older with distance from the coastline due to isostatic uplift, these factors contribute to the deepest peat present along the HBL and Canadian Shield interface (Davies et al., 2023). This distribution is also consistent with the drivers for peatland development reported by Glaser et al. (2004) in that in addition to climatic factors, geological influences, such as differential rates of regional isostatic uplift, drive the distribution of peatlands and an increase in peat depth from coast to inland. The land surface started to rise during glacial isostatic adjustment with the highest rates of uplift occurring around the mouth of James Bay, which resulted in a decrease in river slope, an increase in water level, spread of peatland and the longest peat accumulation time.

Our estimated total C stock of 30 (±6) Pg C in the HBL is in close agreement with the 30 (±1) Pg C reported by Packalen et al. (2014). Packalen et al. (2014) used a subset of our training data, comprising 100 out of 495 data records, and estimated the total C stock by multiplying the HBL area with the mean C stock derived from the relationship between C content in individual samples and peat basal age. Despite the differing methodologies for upscaling the variable number of ground measurements to total HBL C stock estimates, our results also confirm a close relationship between the formation process and age of peatland, peat depth, and C stocks in the HBL. In contrast, our 30 (±6) Pg C estimate differs from that of Sothe et al. (2022a), who reports 38 Pg C in the top one meter of peat depth in the HBL. This discrepancy likely arises from the coarse pixel size (250m) used by Sothe et al. (2022a), which also failed to exclude open water areas, and the insufficient bulk density measurements from HBL and other peatlands in Canada in their national C stock mapping. Significant disparities between large-scale and local soil C stock estimates have been observed in previous studies, attributed primarily to uncertainties in bulk density measurements, particularly in boreal high-latitude regions (Tifafi et al., 2018). This underscores the critical need for additional ground measurements, specifically bulk density measurements, in areas with high peat depth uncertainty, which often coincide, albeit not exclusively, with deeper peats in inland areas of the HBL (see Figure 4 C and D).
While our study compiled a large set of ground measurements and covered a substantial portion of the HBL, certain limitations are acknowledged. Despite distance from the coastline and time since peat initiation being key drivers of peat C stocks in the HBL, variations in plant communities reflecting differing hydrology and peat depth within areas of similar peat age may not be adequately captured by satellite-derived EVI or land surface temperature data. Moreover, Sothe et al. (2022a) highlight that bulk density typically increases beyond 30 cm soil depth, especially in areas with high concentrations of soil organic C. Given that peat core samples are often collected from locations with deep peat profiles, there exists potential for overestimation of soil organic C stocks in shallow peats. Furthermore, the influence of microtopographic features such as open-water pools has been insufficiently addressed in estimating peat volume and C content. Despite more than 45% of the peatland surface area in the HBL being covered by small open-water pools, the overall coverage remains unclear, contributing to a 30-40% uncertainty in the estimation of total C (Loisel et al., 2017). Nevertheless, this study addresses a significant gap in peat depth distribution mapping and contributes to a more accurate estimation of C stock in the HBL. Historically, digital soil mapping studies have predominantly focused on C stored in shallow soil layers, such as the top 2 m depth (Sothe et al., 2022a) or 2–3m (Hugelius et al., 2013). Therefore, our estimation of 30 Pg C stored in the HBL with vertical depths ranging from 0 to 6 m holds great significance for understanding the dynamics of multi-scale peatland C cycles. However, depth and bulk density heterogeneity in Canadian peatlands warrants further investigation, and additional ground samples remain indispensable for improving model accuracy and verification purposes.

Acknowledgments

This research was funded by World Wildlife Fund Canada (WWF-Canada), Natural Sciences and Engineering Research Council of Canada (NSERC) Alliance grant (application No: ALLRP 566310-21), and Canada Research Chairs Program (CRC2019-00139). S. Finkelstein acknowledges research funding from NSERC, the Ontario MNRF, Parks Canada and ECCC for past and ongoing work in the HBL and adjacent areas. We thank the Mushkegowuk Council, Attawapiskat, Fort Albany, Kashechewan, Moose Cree, and Weenusk First Nations, Calvin Lincez, Vern Cheechoo, and Colleen McLeod for supporting our research in their homelands. We thank Clare Wark, Michelle Kalamandeen, David Hunter, Jack Williams, Jonathan Rickard and the Mushkegowuk community members for support during data collection. We also thank the Ontario Ministry of Natural Resources and Forestry for the external data sharing agreement on the project “Peat carbon storage in the Hudson Bay Lowlands”.

References


Zhang, H., Piilo, S., Amesbury, M., Charman, D., Gallego-Sala, A., & Välimäki, M. (2018). The role of climate change in regulating arctic permafrost peatland hydrological and vegetation change over the last millennium. Quaternary Science Reviews, 182, 121-130.
Supporting Information for

Peat depth and carbon storage of the Hudson Bay Lowlands, Canada

Yiyao Li¹, Daorui Han¹, Cheryl A. Rogers¹², Sarah A. Finkelstein³, Oleksandra Hararuk⁴, James M. Waddington¹, Carlos Barreto⁵, James W. McLaughlin⁶, James Snider⁶, and Alemu Gonsamo¹

¹School of Earth, Environment & Society, McMaster University, Hamilton, ON, Canada, ²Department of Geography and Environmental Studies, Toronto Metropolitan University, Toronto, ON, Canada, ³Department of Earth Sciences, University of Toronto, Toronto, ON, Canada, ⁴Northern Forestry Centre, Canadian Forest Service, Natural Resources Canada, Edmonton, Canada, ⁵Ontario Ministry of Natural Resources and Forestry, Sault Ste. Marie, ON, Canada, and ⁶World Wildlife Fund Canada, Toronto, Ontario, Canada

Correspondence to: gonsamoa@mcmaster.ca

Contents of this file

Table S1 – S3
Figures S1 – S5
Table S1. Geospatial and satellite derived variables selected to map peat depth

<table>
<thead>
<tr>
<th>Process</th>
<th>Indicator variable name</th>
<th>Abbreviation</th>
<th>Model selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peat formation and age</td>
<td>Distance to coast</td>
<td>DistBay003</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Distance to nearest river</td>
<td>DistRiver003</td>
<td>Confirmed</td>
</tr>
<tr>
<td>Accumulation</td>
<td>SRTM DEM</td>
<td>SRTM</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Topographic Wetness Index</td>
<td>TWI</td>
<td>Rejected</td>
</tr>
<tr>
<td>Peat occurrence</td>
<td>Median of LST for growing season</td>
<td>LST05</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>SD of LST for growing season</td>
<td>LSTsd05</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Median of LST for whole year</td>
<td>LST12</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>SD of LST for whole year</td>
<td>LSTsd12</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Median VH/VV for whole year</td>
<td>vvvhRatio</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Median (VV + VH)/2 for whole year</td>
<td>vvvhPlus2</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Median VH-VH for whole year</td>
<td>vvvhMinus</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Mean VH for whole year</td>
<td>VH</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Mean VV for whole year</td>
<td>VV</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Median of EVI for whole year</td>
<td>EVI12</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>SD of EVI for whole year</td>
<td>EVIsd12</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>Median of EVI for growing season</td>
<td>EVI05</td>
<td>Confirmed</td>
</tr>
<tr>
<td></td>
<td>SD of EVI for growing season</td>
<td>EVI05</td>
<td>Confirmed</td>
</tr>
</tbody>
</table>


Table S2. Top 5 variables selected by each model and their performance

<table>
<thead>
<tr>
<th>Model</th>
<th>Top 5 important variables</th>
<th>RMSE</th>
<th>RMSLE</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gradient Boosting Machine (GBM)</td>
<td>DistBay, SRTM, vvvhRatio, vvvhMinus, DistRiver</td>
<td>60.34</td>
<td>0.38</td>
<td>3641.16</td>
<td>44</td>
</tr>
<tr>
<td>Deep Learning</td>
<td>EVI05, EVI05, LSTsd12</td>
<td>63.87</td>
<td>0.39</td>
<td>4079.17</td>
<td>47.84</td>
</tr>
<tr>
<td>Distributed Random Forest (DRF)</td>
<td>DistBay, SRTM, vvvhRatio, vvvhMinus vvvhPlus</td>
<td>62.26</td>
<td>0.39</td>
<td>3875.83</td>
<td>45.55</td>
</tr>
<tr>
<td>XGBoost</td>
<td>DistBay, SRTM, vvvhMinu vvvhRatio, EVI05</td>
<td>61.31</td>
<td>0.37</td>
<td>3759</td>
<td>45.24</td>
</tr>
<tr>
<td>Ensemble</td>
<td></td>
<td>59.79</td>
<td>0.38</td>
<td>3574.76</td>
<td>44.76</td>
</tr>
</tbody>
</table>

RMSE = Root Mean Squared Error, RMSLE = Root Mean Squared Logarithmic Error, MSE = Mean Squared Error = MSE, and MAE = Mean Absolute Error.
Table S3. Area and proportion of total mapped peatland of different depth class in HBL.

<table>
<thead>
<tr>
<th>Peat Depth Class (cm)</th>
<th>Area (km²)</th>
<th>Total Mapped Peatland Area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-40</td>
<td>1422.39</td>
<td>0.41%</td>
</tr>
<tr>
<td>41-100</td>
<td>25770.21</td>
<td>7.35%</td>
</tr>
<tr>
<td>101-200</td>
<td>183971.5</td>
<td>52.49%</td>
</tr>
<tr>
<td>201-300</td>
<td>135632</td>
<td>38.7%</td>
</tr>
<tr>
<td>301-600</td>
<td>3663.55</td>
<td>1.05%</td>
</tr>
</tbody>
</table>

Figure S1. (a) Distance to coastline (b) Distance to river (c) The standard deviation of whole year land surface temperature for 2013 – 2021. (d) The standard deviation of whole year EVI for 2015 to 2022.
Figure S2. Importance of the 17 covariates from ensemble model produced by Boruta algorithm, an all relevant feature selection wrapper. The green colour means the covariates are confirmed as important, red colour represents rejected variables, while blue colour represents shadow features with randomness created by Boruta algorithm. Note that the shadow features, shadowMin, shadowMean and shadowMax are random variables introduced by ensemble algorithm. Covariates that have lower importance than any of the random variables contribute negligibly to prediction of the target variable.
Figure S3. Importance rank of the 16 covariates for each model.
Figure S4. Comparison of the predicted and measured peat depth (cm) from the four models using all training dataset.
Figure S5. Comparison of the predicted and measured peat depth (cm) from the four models using independent validation dataset.