Using aboveground vegetation attributes as proxies for mapping peatland belowground carbon stocks

Article *in* Remote Sensing of Environment · June 2019

DOI: 10.1016/j.rse.2019.111217

CITATIONS
2

READS
263

5 authors, including:

Javier Lopatin
Karlsruhe Institute of Technology
29 PUBLICATIONS **127** CITATIONS
SEE PROFILE

Teja Kattenborn
Karlsruhe Institute of Technology
42 PUBLICATIONS **404** CITATIONS
SEE PROFILE

Mauricio Galleguillos
University of Chile
63 PUBLICATIONS **519** CITATIONS
SEE PROFILE

Jorge Perez-Quezada
University of Chile
59 PUBLICATIONS **230** CITATIONS
SEE PROFILE

This is a pre-print version of the following article: Lopatin, J. Teja Kattenborn, Mauricio Galleguillos, Jorge F. Perez-Quezada and Sebastian Schmidtlein (2019). Using aboveground vegetation attributes as proxies for mapping peatland belowground carbon stocks. Remote Sensing of Environment.

Using aboveground vegetation attributes as proxies for mapping peatland belowground carbon stocks

Javier Lopatin[a,*], Teja Kattenborn[a], Mauricio Galleguillos[b,c], Jorge F. Perez-Quezada[b,d] and Sebastian Schmidtlein[a]

[a]Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany.

[b]Department of Environmental Science and Renewable Natural Resources, University of Chile, Casilla 1004, 8820808 Santiago, Chile.

[c]Center for Climate Resilience Research (CR)[2], University of Chile, 8370449 Santiago, Chile.

[d]Institute of Ecology and Biodiversity, Las Palmeras 3425, 7800003 Santiago, Chile.

*Corresponding author: javier.lopatin@kit.edu

Abstract

Peatlands are key reservoirs of belowground carbon (C) and their monitoring is important to assess the rapid changes in the C cycle caused by climate change and direct anthropogenic impacts. Frequently, information of peatland area and vegetation type estimated by remote sensing has been used along with soil measurements and allometric functions to estimate belowground C stocks. Despite the accuracy of such approaches, there is still the need to find mappable proxies that enhance predictions with remote sensing data while reducing field and laboratory efforts. Therefore, we assessed the use of aboveground vegetation attributes as proxies to predict peatland belowground C stocks. First, the ecological relations between remotely detectable vegetation attributes (i.e. vegetation height, aboveground biomass, species richness and floristic composition of vascular plants) and belowground C stocks were obtained using structural equation modeling (SEM). SEM was formulated using expert knowledge and trained and validated using *in-situ* information. Second, the SEM latent vectors were spatially mapped using random forests regressions with UAV-based hyperspectral and structural information. Finally, this enabled us to map belowground C stocks using the SEM functions parameterized with the random forests derived maps.

This SEM approach resulted in higher accuracies than a direct application of a purely data-driven random forests approach with UAV data, with improvements of r^2 from 0.39 to 0.54, normalized RMSE from 31.33% to 20.24% and bias from -0.73 to 0.05. Our case study showed that: (1) vegetation height, species richness and aboveground biomass are good proxies to map peatland belowground C stocks, as they can be estimated using remote sensing data and hold strong relationships with the belowground C gradient; and (2) SEM is facilitates to incorporate theoretical knowledge in empirical modeling approaches.

Keywords: UAV, hyperspectral, SEM, PLS path modeling, belowground carbon stocks, vegetation attributes, random forests.

1. Introduction

Peatlands are important for the regulation of carbon (C) cycling and store one third of the world soil C stocks (Parish et al., 2008), from which belowground C accounts for ~95% of the total peatland C pool (Smith et al., 2004). These are fragile ecosystems and if the conservation of their peat properties is degraded either due to anthropogenic impact or climate change, they release accumulated soil C to the atmosphere (in the forms of CH_4 and CO_2) faster than it is being sequestered (Fenner & Freeman, 2011). This accelerates the greenhouse effect and global warming (Phillips & Beeri, 2008; Schaepman-Strub et al., 2008). Although concerns about the consequences of human-induced changes to the C cycle have generated many international initiatives to quantify peatland C stocks (Hribljan et al., 2017), the contribution of C stocks in small isolated peatlands in comparison to larger systems is still uncertain (McClellan et al., 2017).

Remote sensing provides a rapid and economical approach to supplement traditional direct methods such as coring and probing (Rudiyanto et al., 2018). Frequently, *in-situ* measurements include peat thickness, dry bulk density and carbon concentration, which has to be obtained by laboratory analysis (e.g. Dargie et al., 2017), while remote sensing products include the peatland area and vegetation types. These variables are then used to predict belowground C stocks by applying allometric functions to the obtained vegetation classes. For example, Dargie et al. (2017) have used a maximum likelihood classification with radar, optical and DEM-derived data to classify different land covers including *terra firme* and peatland classes. Afterwards, they have used *in-situ* data to create allometric functions with peat and soil properties as input to estimate the total C pools of the central Congo basin. Other studies have used remote sensing to map peatland attributes, such as peat thickness, and derived continuous estimations of belowground C stocks by allometric functions using these spatially explicit estimations (e.g. Rudiyanto et al., 2018). Despite the accuracy of such approaches, their application is hampered by high field and laboratory costs. Hence, linkages between belowground C stocks and aboveground peatland attributes derived by remote sensing, such as vegetation attributes, could be used to decrease monitoring efforts.

Ecological studies have already established empirical knowledge about linkages between aboveground vegetation attributes and belowground C stocks. For example, vegetation properties such as aboveground biomass, species richness and plant functional type cover are related to the decrease of water-logging and peat mineralization processes in peatlands (Dorrepaal et al., 2005; Jonsson & Wardle, 2009). The abundance and biomass of certain vascular species have proven to be good proxies of belowground C stocks, while species richness has been shown to be positively related to soil microbial activity (Chen et al., 2017, 2018; Cong et al., 2014; Lange et al., 2015) and hence to organic matter decomposition rate (Fenner & Freeman, 2011). Previous studies have retrieved these proxies in peatlands using remote sensing data directly. For example, Castillo-Riffart et al. (2017) and Cabezas et al. (2016) estimated species richness in anthropogenic peatlands using a combination of satellite optical and textural data, while Turetsky et al. (2011) successfully estimated aboveground biomass. Likewise, species composition, plant functioning and growth forms have been mapped successfully by hyperspectral data with reasonable accuracies (Harris et al., 2015; Schmidtlein et al., 2012; Schmidtlein & Sassin, 2004).

To assess the relationship between vegetation attributes and belowground C stocks, we propose to use structure equation models (SEM). SEMs are a family of multivariate methods that allows the definition of preexisting relationships between variables (expert knowledge). Accordingly, SEMs are well-suited for hypothesis testing and understanding underlying processes. Furthermore, SEMs can be constrained using prior knowledge (e.g. causal relationships) of the system at hand and they thus feature a higher transferability than pure data driven modeling approaches (Grace et al., 2010).

The main objective of this study was to assess the potential of vegetation attributes as proxies to predict and map belowground C stocks in a small isolated peatland in Chiloé Island, Chile. First, linkages among vegetation attributes and belowground C stocks were investigated using SEM. Second, we combined hyperspectral and photogrammetric information derived from an Unmanned Aerial Vehicle (UAV) to spatially extrapolate the previously selected vegetation attributes with random forests regressions. Finally, we fed the SEM with the random forests extrapolations to obtain the belowground C stock maps.

Using this general setup, we addressed the following research questions:

- Does the structural equation model outperform a purely data-driven random forests approach?
- Which mappable aboveground vegetation attributes are reliable indicators of belowground C stocks?

2. Materials and methods

2.1. Study site

The study site is an anthropogenic acidic peatland located at the FLUXNET site Senda Darwin Peatland (CL-SDP; 418520 S, 738400 W) in the north of Chiloé Island, in the Los Lagos region of Chile (Fig. 1a,b). This peatland was formed due to the burning of a temperate rainforest, which produced a poorly drained soil that was later colonized by moss species of the genus *Sphagnum* (Díaz et al., 2007). The belowground C stocks of this peatland were estimated as ~11.56 ± 1.57 kg C m^{-2} (Cabezas et al., 2015).

The study area is about 16 ha and covers two types of land use (Fig. 1b): a conservation area (~5.5 ha; southern part of the area), which has been protected and used for scientific investigations for the last 20 years, and a managed area (~10.5 ha; northern part of the area) where *Sphagnum* mosses are artisanally harvested for commercial purposes (~10 kg of dry moss per month) and for grazing of four oxen (see Cabezas et al., 2015 for a full description of the area).

The floristic composition of the area is dominated by *Sphagnum magellanicum*, but also by the bryophytes *Dicranoloma imponens* and *Campylopus introflexus*. These species can be found in sites affected by water logging. With less water availability, dominant graminoids species are *Danthonia chilensis*, *Schoenus rhynchosporoides* and *Uncinia tenuis*. On poorly-drained and water-saturated areas dominant species are *Juncus procerus*, *Juncus planifolius* and *Juncus stipulatus*.

Common forb species are *Anagallis alternifolia*, *Centella asiatica* and *Lotus pedunculatus*, while common shrub species are *Baccharis patagonica*, *Gaultheria mucronata* and *Myrteola nummularia* (list of species in Appendix A of Cabezas et al., 2015).

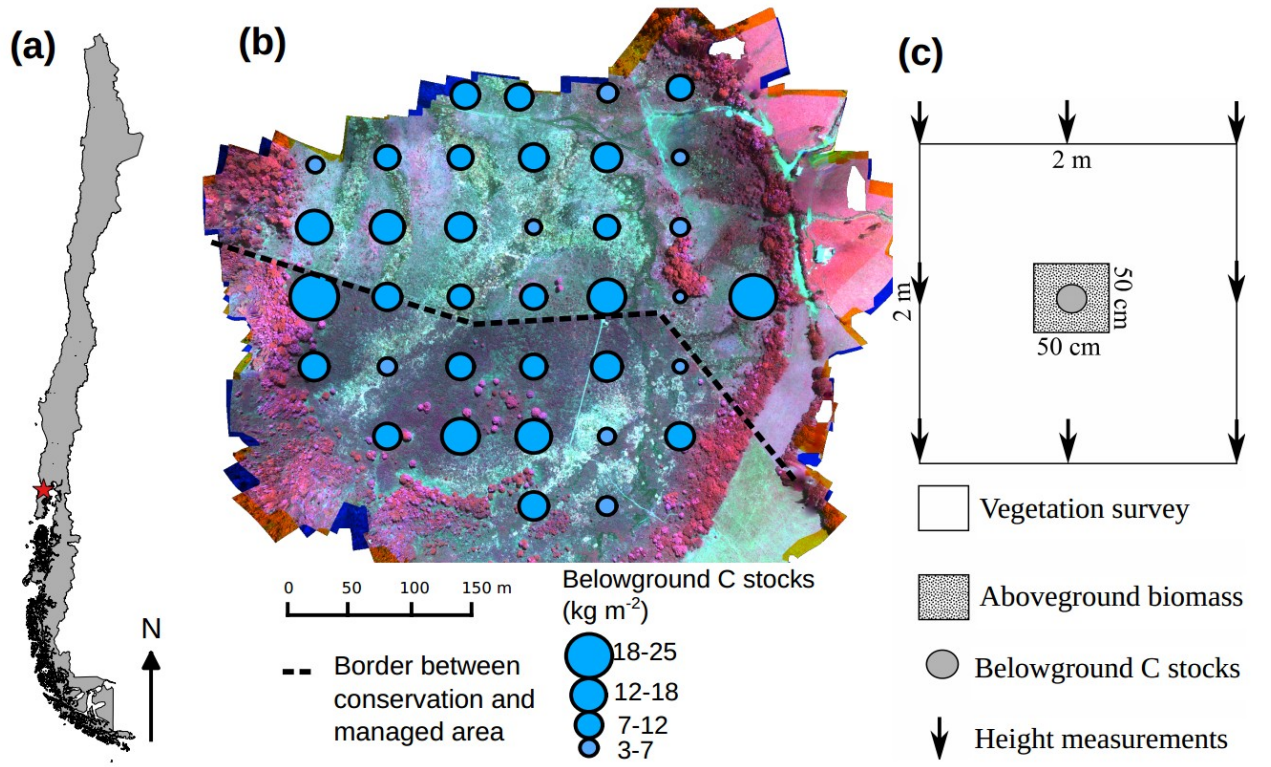

Fig. 1. (a) Location of the study area in Chile; (b) UAV false color composite (R = 765 nm, G =665 nm, B = 565 nm) of the study area with plot locations scaled with the measured belowground C stocks (blue points). The dashed line shows the border between the conservation area (bottom) and the managed area (upper); (c) detailed sampling design.

2.2 *In-situ* / reference data

A vegetation assessment was conducted between January and April 2014, in which 36 plots were established in a systematic grid with 60 m × 60 m spacing. Each plot was a 2 m × 2 m quadrat where the species presence, cover, biomass, vegetation height, soil depth and the C stocks were estimated.

Species coverages were obtained by averaging visual estimates of two observers. The biomass and C reservoirs were divided into two type of stocks: the aboveground stocks included the vascular flora growing on the substrate, and the underground stocks included the peat, live moss, debris (fine and coarse) and buried trunks (remnants from burning the forest). A detailed description of the sampling methods is presented in Cabezas et al. (2015); the aboveground biomass stocks for each plot was estimated by harvesting the vascular flora in a 0.25 m^2 sub-plot located in the center of the plot. Belowground C stocks, consisting of peat, living moss, woody debris (fine and coarse) and buried burnt trunks, were sampled using a peat profile sampler (Eijkelkamp, Giesbeek, Netherlands) at the central point of the plot (Fig. 1c). Soil and aboveground C stocks were dried at 70°C for 72 h to obtain the weight and density of the material, and to estimate their values in kg m^{-2}. Five composite samples were generated from random sampling points to obtain the C fraction for each stock (Cabezas et al., 2015). From these composite samples a sub-sample of 10 g was extracted, ground mixed and processed in an elemental analyzer (NA2500, Carlo Erba, Milan, Italy). Soil depth was measured by summing up all these composites.

2.3. UAV data acquisition

Hyperspectral data were collected in February 2016 using a UAV (octocopter) based on a NAZA-M V2 flight controller (DJI, Shenzhen, China). The UAV carried two small snapshot mosaic cameras (Gamaya, Lausanne, Switzerland), one covering the visible spectral region (VIS) with 16 bands and the other the near-infrared (NIR) spectral region with 25 bands. The flight plan aimed for an average of 80% of forward and 70% of sided overlap, and an altitude of 100 m above the ground. The image frames were processed in a Structure-from-Motion (SfM) pipeline (Agisoft Photoscan, Russia) to obtain a single hyperspectral brick of 41 bands (450 mn−950 mn, and 10 nm bandwidth) and a point cloud with elevation information (Kattenborn et al., 2018; Lopatin et al., 2019). Reflectance data were obtained by calibrating the raw hyperspectral data with a reference panel with known reflectance placed in the field during the flight. The resulting pixel size of the hyperspectral orthomosaic was ~0.1 m while the point cloud densities resulted in ~1,000 points/m^2 (i.e. an average of 1 point every 0.003 m).

The height values of the point cloud were transformed from meters above sea level to meters above the ground using TreesVis (Weinacker et al., 2004). Retrieving absolute vegetation height information from SfM-based point clouds can be challenging in peatlands as there is a low chance that points representing the terrain will be reconstructed (Mercer and Westbrook, 2016). As we considered peat as part of the belowground stocks, only vascular vegetation height was measured. We used the measured *in-situ* vascular vegetation height to correct for possible shifts in the SfM-based elevations. A bilinear interpolation was applied to the average height values of the point cloud falling inside the plots using the average plot measurement as targets. The vegetation height accuracy resulted in root mean square errors (RMSE) of ~0.02 m.

Finally, both hyperspectral data and the point cloud were georectified with a bilinear interpolation using a Pleiades panchromatic image (0.5 m × 0.5 m) collected on January 28, 2014. We selected 23 steady features (e.g. tree crowns and corners of an elevated footbridge located in the conservation area) as ground control points. The positional accuracy resulted in RMSE of ~0.40 m. Moreover, we placed markers without georeference in the corners of all 36 plots to make the plots directly visible from the hyperspectral orthomosaic. From the 36 plots, only 31 were finally visible in the orthomosaic. Unseen markers were located mainly in the conservation area, where the visibility was hampered by the high shrub biomass. For plots where the markers were not visible in the orthomosaic, we retrieved the spectral and vascular vegetation height information from the GPS position.

2.4. Selected aboveground vegetation attributes

We reviewed the literature to determine the main vegetation characteristics related to peatland belowground C stocks. We used this information to develop a structural equation model that assimilates this knowledge into a network of multivariate interactions, whichwere trained with and tested against our data.

The key aboveground vegetation characteristics influencing belowground C stocks were related to the decrease of water-logging and peat mineralization, which are causing the decline of *Sphagnum* cover and promoting the colonization by vascular plants. Hence, these characteristics relate to processes that alter belowground fluxes and gas-exchanges, resulting in a slow decreases in C stocks. Moss species (mainly *Sphagnum* spp) exude inhibitory polyphenol compounds by the rhizoids that increase the accumulation of belowground C due to a low decomposition rate of

organic matter (i.e. the decomposition rate is lower than the accumulation rate; Fenner & Freeman, 2011). On the contrary, vascular plants tend to stimulate belowground microbial activity through increased labile carbon, accelerating the decomposition rate of the organic matter. This is followed by an expected alteration of soil nutrient content, pH conditions and a water table that further reinforce the vascular plant colonization. This will cause the decrease of peat abundance and eventually belowground C stocks (Fenner & Freeman, 2011). For this reason, the cover and biomass of growth forms (i.e. bryophytes, graminoids, forbs and shrubs) and certain vascular species have been shown to be proxies for C stocks (Dorrepaal et al., 2005; Ma et al., 2017). Likewise, a higher plant diversity has been demonstrated to increase rhizosphere C inputs and belowground microbial activity (Chen et al., 2017, 2018; Cong et al., 2014; Lange et al., 2015), hence increasing the organic matter decomposition rate on peatlands (Fenner & Freeman, 2011).

Following these relations, we selected four aboveground vegetation characteristics from the literature related to belowground C stocks that can be estimated from remote sensing data with moderate accuracies: vegetation height (e.g. Rudiyanto et al., 2018), species richness (e.g. Castillo-Riffart et al., 2017), aboveground biomass (e.g. Turetsky et al., 2011) and the assemblages of vascular species communities or floristic gradients estimated by an ordination algorithm (Nonmetric Multidimensional Scaling, NMDS, see Supplementary data; e.g. Schmidtlein & Sassin, 2004). The range observed for these variables and their indicators are presented in Table 1. Furthermore, we included soil depth as a key intermediate mediator (Akumu & McLaughlin, 2014).

Table 1. Plot-based variables included in the modeling.

Variables	Indicators (unit)	Min	Mean	Max
Vascular vegetation height	Vegetation height (cm)	2.83	30.30	130.55
Floristic composition	NMDS 1	-0.78	0	1.18
	NMDS 2	-1.13	0	0.71
	NMDS 3	-0.66	0	1.13
Vascular aboveground biomass	Shrub biomass (kg m^{-2})	0	0.38	5.36
	The sum of graminoid, forb and fern biomass (kg m^{-2})	0.03	0.13	1.86
Vascular species richness	Shrub richness (N)	0	4	7
	Graminoid richness (N)	0	1.81	4
	Forb richness (N)	0	1.43	5
	Ferns richness (N)	0	2.01	5
Soil depth	Soil depth of peat, live moss, debris (fine and coarse) and buried trunks (remnants from burned forest) (cm)	18.50	38.47	91
Belowground C stocks	Carbon stocks of peat, live moss, debris (fine and coarse) and buried trunks (remnants from burned forest) (kg m^{-2})	3.07	11.40	24.91

2.5. Algorithms used

2.5.1. Random forests

The ensemble regression tree method random forests (RF; Breiman, 2001) has been reported to be an efficient regression algorithm when the numbers of observations is comparably low in relation to the number of predictors (Svetnik et al., 2003). RF requires two parameters to be set: 1) mtry, the number of predictor variables used for the data partitioning at each split and 2) ntree, the total number of trees to be grown in the model run. We set ntree to 500 based on literature recommendations, whereas mtry was tuned using leave-one-out cross-validation. The importance of predictor variables was measured by the Gini decrease in node impurity measure, which is computed by permuting the predictor variables with the out-of-bag data in the RF validation approach (for details see Liaw & Wiener, 2002).

2.5.2. Structural equation modeling

Structural equation modeling (SEM) is a multivariate family of methods that can model a large number of interactions simultaneously, providing a framework for inferring cause-effect relationships and estimating direct and indirect relations among variables (Grace et al., 2010).

We used partial least squares path modeling (PLS-PM or PLS-SEM; Tenenhaus et al., 2005), a non-parametric composite-based SEM, whichis mainly used in social science, but has already shown potential in ecological (Ferner et al., 2018) and remote sensing applications (Lopatin et al., 2015). PLS-PM is validated at different levels (i.e. loading, weights, R^2), and has an overall model goodness-of-fit (GoF; ranging from 0 to 1, where 1 is a perfect representation) proposed by Tenenhaus et al. (2005). The significance ($\alpha = 0.05$) of each interaction and model outputs were obtained by means of bootstrapping (1,000 iterations).

PLS-PM can create latent variables (LVs) comprised of one or several variables in a supervised manner. Thus, PLS-PM can create components to separate predictors according to their correlations and reduce overfitting. We used all LVs in reflective mode or type ‘A’ (see Tenenhaus et al. (2005) for information on the LVs types). The interactions among LVs are explained by Linear Ordinary Least Squares (OLS). We standardized variables to normalize path coefficients and intercepts (i.e. turn variables with different raw units into standard deviation units; Grace & Bollen, 2005). The PLS-PM was tuned using the Cronbach’s alpha index (check for unidimensionality among indicators) and the loading values (correlation within the indicators of a LV; variables with loadings below 0.5 were dropped).

The theoretical construct of the model is presented in Fig. 2 and the indicators used in Table 1. We used Moran’s I index and spatial correlograms to check for spatial autocorrelation on the residuals of the bootstrapped LVs.

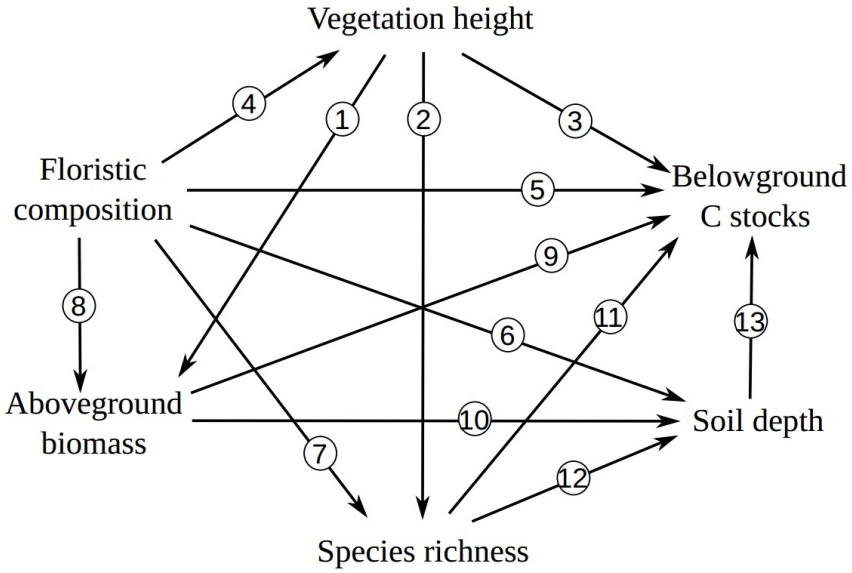

Fig. 2. The hypothesized model for the plot-based PLS-PM. Solid black lines represent expected effects obtained from literature. Examples are (1), (2), (3), (4), (5) and (6) Lawson et al. (2014); (7) Rocchini et al. (2018); (8) Dorrepaal (2007); (9) and (10) Dorrepaal et al. (2005); (11) and (12) Lange et al. (2015) and (13) Lawson et al. (2014), Akumu & McLaughlin (2014) and Draper et al. (2014).

2.6. Modeling and validation

Three types of belowground C stocks predictive models were tested (summarized in Fig. 3):

a) using random forests and PLS-PM independently with plot-based predictors (Fig. 3a);

b) using random forests and PLS-PM independently with UAV-based predictors (Fig. 3b);

c) using a hybrid model combining PLS-PM plot-based information with random forests UAV estimations (Fig. 3c).

In the models depicted in Fig. 3 a) and b), PLS-PM and RF were used independently to compare their performances. For the plot-based models we used the variables presented in Table 1, while the remote sensing models (Fig. 3b) were parameterized with UAV-based spectral and height information. We applied a brightness normalization (Feilhauer et al., 2010) to the hyperspectral data to compensate for high heterogeneity in illumination owed due to the high spatial resolution and a minimum noise fraction transformation (MNF; Green et al., 1988) algorithm to reduce noise. We selected the first three MNF components (~99% of the original information) for the analysis. The canopy height information was obtained from the point cloud using FUSION (McGaughey, 2018). The point cloud variables included common vegetation height metrics, such as the minimum, maximum, mean, median, mode and standard deviation of the vegetation height (compare Supplementary data for the full list of variables). The 3D metrics were rasterized to pixel size of 2 m × 2 m to match the field plots size. The PLS-PM structural model for the remote sensing estimation is presented in Fig. A1.

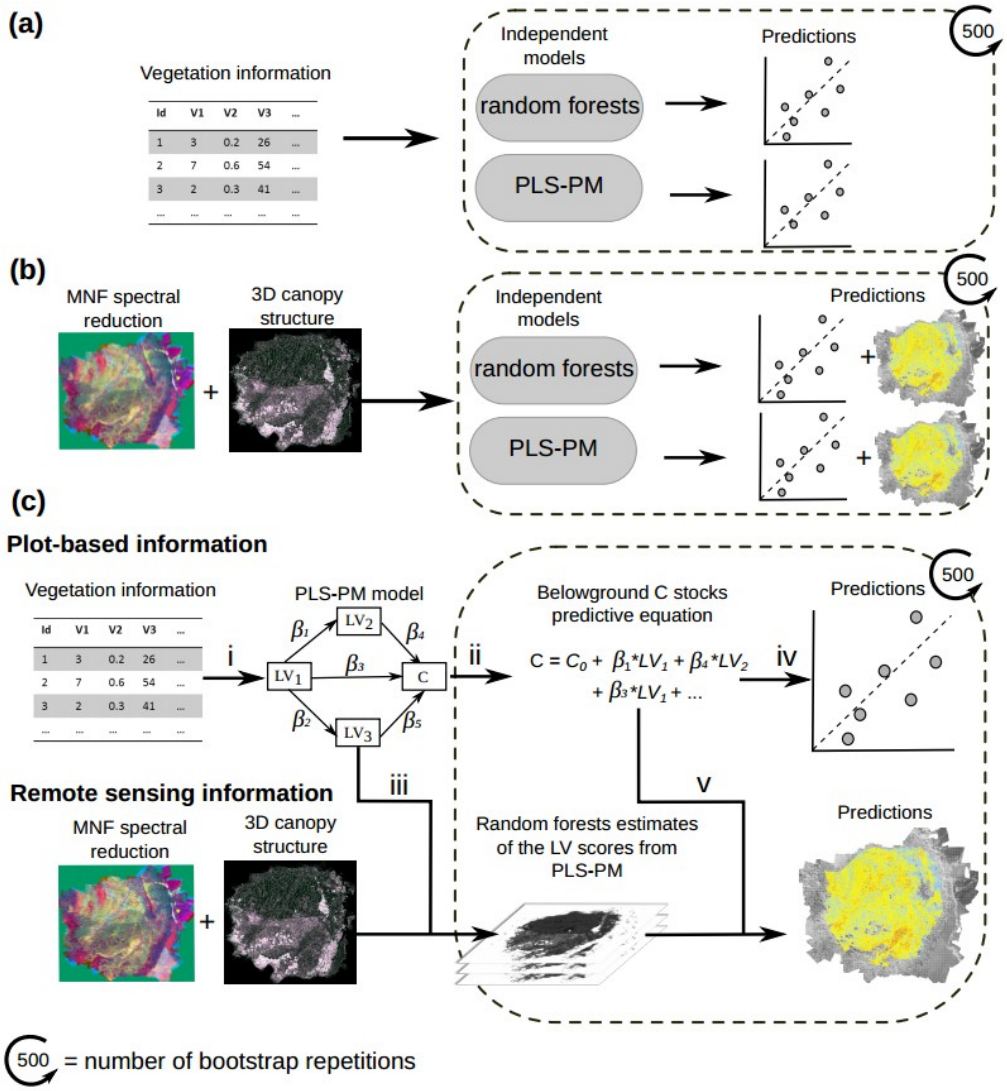

Fig. 3. Methodological workflow of the predictive models. In (a) and (b) belowground C stocks independent predictive models using RF and PLS-PM were applied using the plot-based and UAV-based predictors, respectively; and (c) the hybrid models combining the SEM plot-based information with the RF UAV estimations.

For the hybrid model c), we trained and validated PLS-PM with plot-based information (Fig. 3c i) to find significant vegetation characteristics influencing belowground C stocks and to use their path coefficients (β) to build a predictive function (Fig. 3c ii) as:

$$C = \sum_{i=1}^{n} \beta_1 \times LV_i + \varepsilon_1 (1)$$

where C is the estimated belowground C stocks, LV_i is a significant vegetation LV, β_i is the path coefficient of the LV_i,, ε_i is the error and n is the number of significant vegetation characteristics detected by PLS-PM. We then used RF to extrapolate the selected PLS-PM LV scores using the UAV-based predictors (Fig. 3c iii). Finally, we used equation (1) to predict and map belowground C stocks using the RF map extrapolations (Fig. 3c iv and 3c v respectively).

To avoid overfitting and to assess model accuracy and map extrapolation stability, all three types of models (Fig. 3a, b and c) were embedded in a bootstrapping procedure with 500 repetitions. In each iteration, 36 observations were randomly selected with replacement from the 36 available samples, from which on average 36.8% (~13 samples) were not selected. We used this observations as

holdout samples for the validation (Kohavi, 1995). Model performances were compared based on differences in the coefficients of determination (r^2; calculated as the squared Pearson's correlation coefficient), the normalized root mean square error (%RMSE) and the bias between predicted and observed variables of the holdout samples in the bootstrap. The normalized root mean square error was calculated as:

$$\%RMSE = \left(\sqrt{\frac{1}{2} \sum_{i=1}^{n} (y_j - \hat{y}_j)^2} / [max(C) - min(C)] \right) \times 100 \quad (2)$$

where C is the measured belowground C stocks, while the bias of prediction was measured as one minus the slope of a regression without intercept of the predicted versus observed values (Lopatin et al., 2016, 2017).

We applied a one-sided bootstrapping test to check for significant differences (in terms of r^2, %RMSE and bias) among models (Araya-López et al., 2018; Lopatin et al., 2016). We tested for differences between PLS-PM and RF in the three type of models presented in Fig. 3.

Finally, it is worth-mentioning that still *in-situ* values of belowground C stocks are needed to calibrate and validate the relationships obtained by the vegetation characteristics. Hence, we did not obviate soil sampling completely for the analysis. We used the R-packages 'plspm' (Sanchez et al., 2017) and 'randomForest' (Liaw & Wiener, 2002) for the analyses, while the image processing was accomplished using Python 3.6 with the Scikit-learn (Pedregosa et al., 2011) and Scikit-image libraries (van der Walt et al., 2014; scripts available in Appendix C).

2.7. Maps

We calculated the median and the coefficient of variation (CV, in percentage, calculated as $CV_p = [\frac{SD_p}{mean(C)_0}] \times 100$, where CV_p is the pixel's coefficient of variation, SD_p is the pixel's standard deviation, and $mean(C)_o$ is the mean belowground C stock value of all reference measurements; Araya-López et al., 2018) of the 500 predicted belowground C stock maps produced by the models in Fig. 3b and c during the iterative validation. Pixels with high CV indicate higher uncertainties of the predictive model, while low values depict areas with stable predictions.

We masked out areas that were not represented in the plot-based information to avoid predictions outside the training range (non-vegetation: NDVI < 0.3; and areas with presence of trees: vegetation height > 2 m).

Finally, we used RGB maps of the PLS-PM LV scores extrapolated by RF (Fig. 3c iii) to see the spatial distribution of the plot-based information and support the analysis of the predicted belowground C stock gradients. The components were normalized between 0-255 and plotted with a 20–80% percentiles linear stretching for visual interpretation of variable interactions.

3. Results

3.1. Model performances

The PLS-PM plot-based model yielded an overall goodness-of-fit of 0.65. Moran´s I test showed that LVs (inner models) and their residuals were not significantly affected by spatial autocorrelation (aboveground biomass: I = -0.02 and P = 0.72; species richness: I = -0.02 and P = 0.71; soil depth:

I = -0.04 and P = 0.42; and belowground C stocks: I = -0.04 and P = 0.41). Details on the construction of the LVs (outer model) are presented in Table A1. The model path diagram with significant path feedbacks and total variable importance on belowground C stocks are presented in Fig. 4. Only soil depth and vegetation height showed a significant direct relation to the belowground C stocks. The relation of species richness with belowground C stocks became significant when the indirect influences were included (mainly caused by the relation between vascular species richness and soil depth; Fig. 4b).

Soil depth appeared to be negatively related to aboveground biomass and species richness. No direct link was observed to floristic composition. Hence, a linear equation was created using vegetation height and soil depth to predict belowground C stocks. To use only vegetation attributes in the equation, we used aboveground biomass and species richness path coefficients to predict soil depth, resulting in the following equation:

$$Belowground\,C\,stocks = -(H \times 0.34) + \left(\underbrace{[-(BM \times 0.37) - (SR \times 0.37)]}_{Soil\,depth} \times 0.88 \right) \quad (3)$$

where H is the vegetation height [cm], BM is the aboveground biomass [kg m^{-2}] and SR is the species richness.

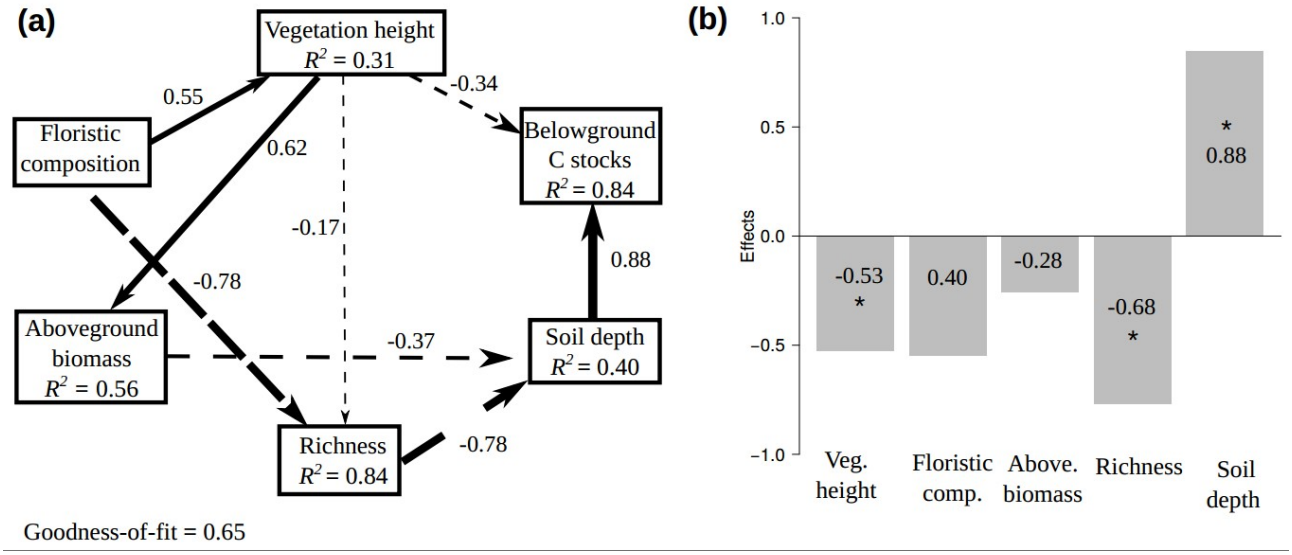

Fig. 4. (a) Resulting PLS-PM model using plot-based information. Arrows represent unidirectional relationships among LVs. Solid and dashed arrows denote positive and negative relationships, respectively. Arrows with non-significant coefficients (α = 0.05) were not drawn. The thickness of the paths is scaled based on the magnitude of path coefficient (β). Path coefficients and internal R^2 correspond to mean bootstrap values of the internal PLS-PM validation. (b) Bar plot showing the variable importance (direct + indirect effects) of the vegetation attributes to predict belowground C stocks. Asterisks (*) indicate a significant (α = 0.05) influence over belowground C stocks.

Table 2 shows the results of belowground C stocks accuracy prediction using RF and PLS-PM independently with plot-based predictors (models of Fig. 3a). Both algorithms yielded high median accuracies with r^2 over 0.70 and %RMSE below 25%. However, PLS-PM showed more parsimony as fewer predictors were used and presented significant improvements (α = 0.05) in terms of %RMSE and bias. RF variable importance is presented in Table 3, showing that soil depth obtained

the highest model scores, whereas all vegetation attributes were non-important. Table 4 presents the modeling results using UAV-based predictors (models of Fig. 3b), where it is shown that RF outperformed PLS-PM significantly in all models. The canopy 3D variables derived from the point cloud showed higher importance than the spectral MNF components for all models except for the model of belowground C stocks (compare Supplementary data). Finally, Fig. 5 depicts the results of the hybrid model coupling the plot-based information (PLS-PM) and UAV data (RF). The hybrid model resulted in higher accuracies than the purely data-driven random forests approach, with improvements from of r^2 from 0.39 to 0.54, normalized RMSE of from 31.33% to 20.24%, and bias from -0.73 to 0.05.

Table 2. Prediction accuracies of belowground C stocks applying plot-based information with random forests (RF) and PLS path modeling (PLS-PM). The median values of the iterative validation are shown. Asterisks (*) indicate a significant difference ($\alpha = 0.05$) between PLS-PM and RF.

Algorithm	r^2	%RMSE	Bias	N° predictors
PLS-PM	0.79	14.48*	0.12*	3
RF	0.72	22.35	0.70	5

Table 3. Variable importance of the random forests plot-based model (Fig. 3a).

Variables	Gini purity index
Soil depth	100.00
Vascular species richness	16.98
Floristic composition	10.41
Vascular vegetation height	1.01
Vascular aboveground biomass	0.00

Table 4. Prediction accuracies of peatland characteristics (latent variables) using PLS-PM and RF with solely remote sensing information. Vegetation height was obtained directly from the UAV photogrammetric point cloud. All values are the median of the bootstrapping validation.

	PLS-PM			RF		
Latent variables	r^2	%RMSE	bias	r^2	%RMSE	bias
Vascular vegetation height	-	-	-	-	-	-
Floristic composition	0.22	26.77	-0.78	0.57***	22.57**	-0.59**
Vascular aboveground biomass	0.30	38.13	-0.69	0.67**	20.18**	-0.27*
Vascular species richness	0.32	26.44	-0.69	0.59**	25.30***	-0.58***
Soil depth	0.40	30.93	-0.62	0.45*	28.83**	-0.79
Belowground C stocks	0.34	36.67	-0.62	0.39*	31.44**	-0.73**

Significant differences: *$\alpha = 0.1$, **$\alpha = 0.05$, ***$\alpha = 0.001$.

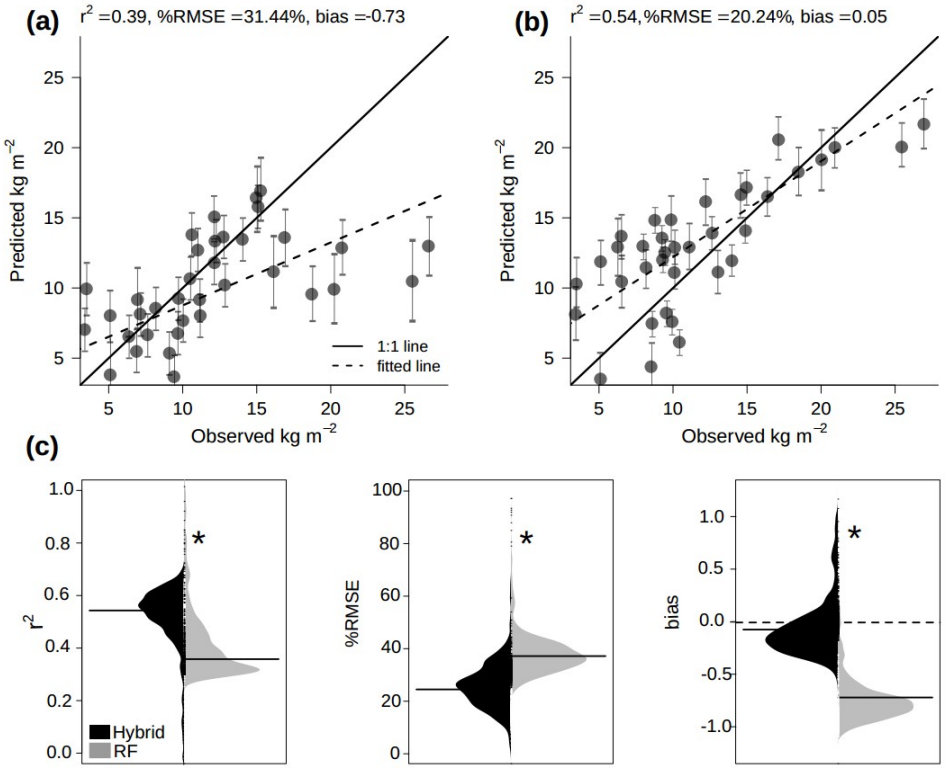

Fig. 5. Scatterplots of observed versus predicted values of belowground C stocks for: (a) the direct use of random forests using UAV-based predictors, and (b) the hybrid model combining PLS-PM plot-based information and RF UAV predictions. The dots and the error bars represent the median and the standard deviation of the values generated in the iterative validation, respectively. (c) Bootstrapping distribution of accuracies. Asterisks (*) indicate a significant differences (α = 0.05).

3.2. Spatial extrapolation of belowground C stocks

The prediction maps (Fig. 6) show that the use of RF with UAV-based predictors resulted in an underestimation of belowground C stocks (see also Fig. 5). Furthermore, the hybrid model depicted lower model uncertainties in prediction, with coefficient of variation (CV) values < ~30%.

The extrapolated PLS-PM plot-based LV scores by the RF models (Fig. 3c iii) showed clear differences between land use types (Fig. 7). The conservation area (to the south) featured more interactions (color combinations), with higher LV score values of aboveground biomass, species richness, floristic composition and soil depth, while the managed area (to the north) was characterized by simpler and smoother patterns of species richness and floristic composition scores. Areas where the bootstrap iteration procedure showed higher CV agree with areas of higher (LV) biomass and species richness.

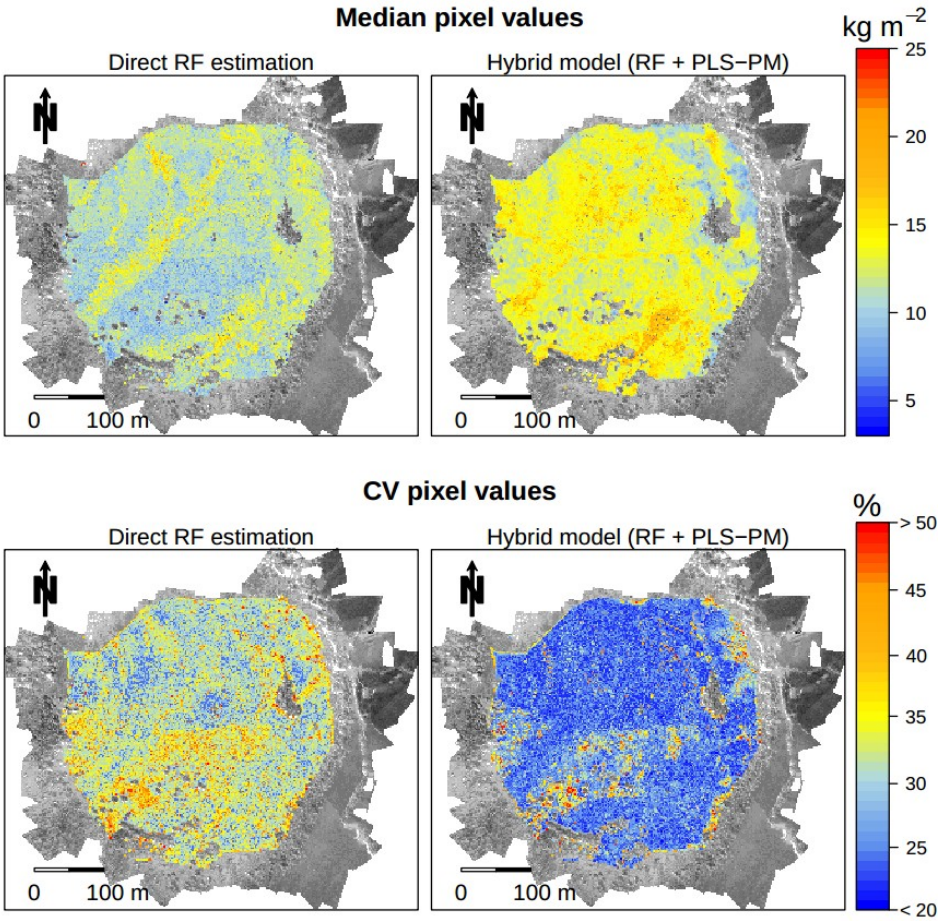

Fig. 6. Belowground C stocks prediction maps using random forests (RF) with UAV-based predictors (left) and the hybrid model combining PLS-PM plot-based information and RF UAV estimations (right). The maps represent the median (top) and the coefficient of variation (CV; below) of the 500 bootstrapping iterations.

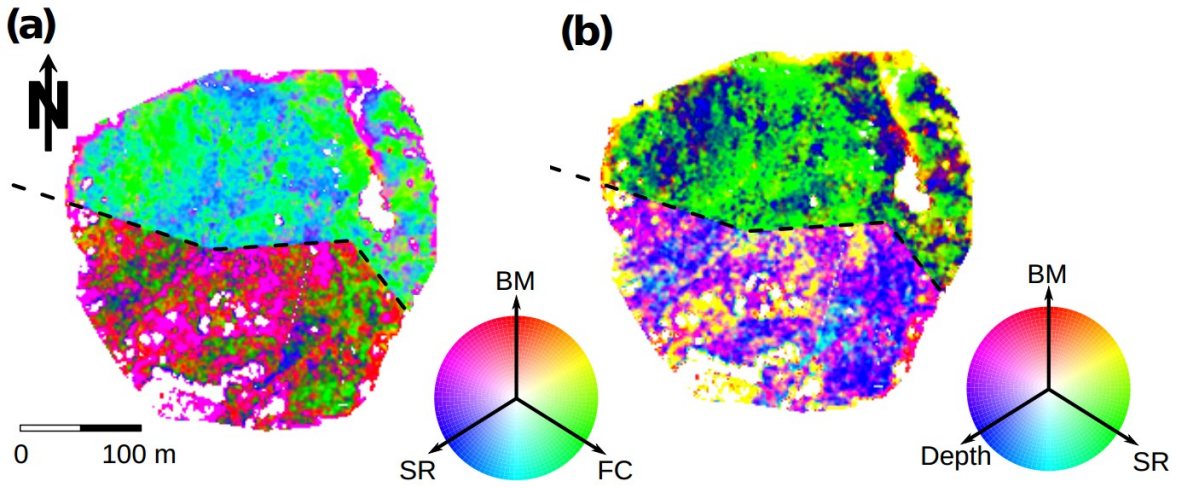

Fig. 7. Spatial distribution of the PLS-PM LV scores produced by RF regressions (Fig. 3 iii). RGB map representations using only aboveground peatland characteristics (a), and a mix of above with belowground variables (b). The variables were normalized (0-255) and plotted with a 20–80% percentiles linear stretching. BM = vascular aboveground biomass, FC = floristic composition, SR = vascular species richness and Depth = soil depth. The dashed line is the border between the conservation (southern) and the managed (northern) areas.

4. Discussion

4.1. Linkages between vegetation attributes and belowground C stocks

The conservation area (to the south) presents higher cover of shrub species (more aboveground biomass) and abundance of native species related to a peatland in the original state (Fig. 7), such as *Sphagnum magellanicum*, *Sticherus cryptocarpus* and *Blechnum cordatum* (low NMDS axis score; Fig. A2). In contrast, the managed area (to the north) has higher species richness and presence of vascular invasive species such as *Plantago lanceolata*, *Agrostis capillaris* and *Trifolium repens* (higher NMDS axis score; Fig. A2), which relates to a more degraded peatland (Dorrepaal et al., 2005; Jonsson & Wardle, 2009). The differences in species composition relate to the colonization of less water-logged areas by vascular plants. The colonization of vascular plants decrease the reservoirs of belowground C stocks in the long term by producing changes in soil nutrient content and pH will eventually hamper further growth of *Sphagnum* species (Fenner & Freeman, 2011). Vascular plants further increase the oxigenization of soils, the microbial activity and decomposition rate of the organic matter, which facilitates liberation of ancient carbon as CO_2 to the atmosphere (Walker et al., 2016) and inhibits further accumulation of C (Chen et al., 2017, 2018; Cong et al., 2014; Gorham, 1991; Lange et al., 2015). This process is likely to be faster and occur more frequently under climate change scenarios, as an increasing temperature promotes drought and accelerates colonization by vascular plants and microbial activity (Fenner & Freeman, 2011).

The negative effects that colonization of vascular plants exert over belowground C stocks is well represented in the PLS-PM model: vegetation height, aboveground biomass, floristic composition (higher scores showing an increase of invasive vascular species) and species richness depicted negative effects (Fig. 4b). From all aboveground variables, only vegetation height showed a significant direct relation to the belowground C stocks. This is clearly an expression of the fact that bog vegetation is low when peat accumulation is high. When considering also indirect effects through the use of moderator variables (especially soil depth), all the above mentioned variables increased their relative importance. This shows the importance of studying variable inter-dependencies when modeling complex systems (see section 4.3 below).

4.2. Model performances

The predictions of random forests (RF) and PLS-PM using plot-based information (Fig. 3a) showed high accuracies in both cases. Nevertheless, PLS-PM outperformed RF significantly using linear relations and fewer variables, indicating that SEM is indeed useful to constrain the models using ecological expert knowledge.

RF outperformed PLS-PM significantly when using remote sensing data to directly predict belowground C stocks (Fig. 3b). This showed that PLS-PM could not handle the high colinearity of the remote sensing data. Although PLS is known to be able to address a certain degree of colinearity, machine learning algorithms like RF behave more flexible by applying complex relationships in higher dimensional feature spaces (Svetnik et al., 2003).

We also tested the use of PLS and support vector machines (SVM) regressions for the predictions. Their performances in terms of r^2, %RMSE and bias were lower than those obtained by RF, hence their results were not presented here.

When using UAV-based independent variables, the aboveground vegetation variables were predicted with higher accuracies than the belowground variables (Table 4). Nevertheless,

predictions of aboveground vegetation variables still showed high uncertainties (i.e. average r^2 of ~0.61 and %RMSE of ~22.68). In this study, the accuracies for the vascular aboveground biomass, species richness and floristic composition are likely to be affected by three factors: 1) the destructive sampling performed in a 0.5 m × 0.5 m area located at the center of the 4 m^2 plots (Fig. 1) prior to UAV acquisition; 2) the time lag of two years between the field (2014) and the UAV (2016) campaigns. In this time period, *Sphagnum* and shrub biomass and abundance were expected to remain stable due to their slow growth but herbaceous properties may have experienced shifts; 3) the inaccuracies during UAV data georectification may have added errors in the five field plots where markers were not visible from the hyperspectral orthomosaic. These observations could have been skipped during the analysis, but we decided to use them anyway as the number of observations was already low.

The prediction of the aboveground variables showed relatively similar results as other studies: Castillo-Riffart et al. (2017) estimated species richness in the same study are using satellite-based sensors, yielding model r^2 between 0.54 and 0.6 and %RMSE between 18% and 20%. Moreover, Harris et al. (2015) mapped the floristic composition of a temperate peatland in the west coast of Wales using a full-range airborne-based hyperspectral data and PLS regressions. They obtained a r^2 in validation of 0.72 and RMSE 0.18 (in feature space units). Estimations of aboveground biomass in acidic peatlands have been obtained mainly by allometric equations and land cover maps obtained by high resolution imageries (e.g. Beilman et al., 2008).

Compared to the RF models based on UAV-based predictors alone (Fig. 3b), the hybrid model coupling PLS-PM plot-based information and RF UAV predictions (Fig. 3c) resulted in significant improvements . This suggests that the use of vegetation attributes as predictors improved the estimation of belowground C stocks by incorporating known ecological relations constraining the model. SEM theoretical constructs are generalizable, allowing their comparison across sites (Grace et al., 2007) and scales (Grace et al., 2016), while empirical models based on remote sensing data alone, despite precision and ease of implementation, lack portability as they are largely affected by sensor and site-specific conditions (Kattenborn et al., 2017; Schmidtlein et al., 2012; Vuolo et al., 2013). The proposed approach showed similar accuracies as other studies using a more complex set of predictors (such as peat thickness, dry bulk density and carbon concentration) and/or allometric functions (Akumu & McLaughlin, 2014; Beilman et al., 2008; Dargie et al., 2017; Draper et al., 2014; Gumbricht et al., 2017; Jaenicke et al., 2008). The advantage of using aboveground vegetation characteristics as proxies are: 1) field efforts and costs are reduced by allowing non-destructive sampling. This may decrease soil sampling and laboratory analyses, and 2) variables can be mapped by remote sensing with higher accuracies than belowground variables (e.g. Castillo-Riffart et al., 2017; Harris et al., 2015). The aboveground vegetation variables used in the present study relate to changes in C fluxes (e.g. gas-exchange to the atmosphere) that slowly alter the C stocks (Fenner & Freeman, 2011). Therefore, the use of such proxies will always require validation as relations between vegetation proxies and belowground C stocks may depend on site specific conditions, such as floristic composition, management treatments and successional stages. Nevertheless, improving the understanding of the interactions between below and aboveground plant and community properties could advance the operationalization of such mapping approaches by decreasing the amount of data required for training and validation.

The accuracies obtained in this investigation are still not sufficient for an operational monitoring of peatland belowground C stocks. To achieve this goal, higher accuracies are needed (e.g. > 80% of

variance explained) to decrease the amount of field data without compromising model performances. Likewise, the approach portability needs to be assessed, as site specific approaches are not ideal for management tasks. To improve the estimation of belowground C stocks with the proposed approach, an increase of the prediction accuracy of the indirect proxies is needed (i.e. vascular species richness and aboveground biomass). For example, the development of an accurate allometric equation to assess aboveground biomass in a non-destructive manner would facilitate the proposed procedure. The vascular species richness estimation could be enhanced e.g. by using detailed multi-scale canopy 3D and textural information based on the point cloud and the orthomosaic (Lopatin et al., 2019, Kattenborn et al. 2019)

4.3. Variable importance

The plot-based PLS-PM model showed significant links (direct + indirect) between soil depth, species richness, vegetation height and belowground C stocks (Fig. 4). The model showed that for every centimeter increase of soil depth, centimeter decrease of vegetation height and reduction of species numbers by one species, there is a belowground C stocks increase of ~0.88 kg m^{-2}, ~0.53 kg m^{-2} and ~0.68 kg m^{-2}, respectively. In comparison to these results, the RF model was rather uninformative: the RF Gini purity index (Table 3) showed that soil depth was the only meaningful predictor of belowground C stocks in the RF model. This suggest that considering only direct relations (i.e. regression coefficients or model variable importance) may lead to underestimations of the ecological importance of several variables in the model. For example, in the PLS-PM plot-based model, influences of species richness and floristic composition on determining belowground C stocks using direct links (path coefficients) yielded non-significant results ($\beta < 0.1$ each), whereas presenting strong importance if indirect links were considered (species richness = -0.68; floristic gradient = -0.40); the assessment of only direct influences may lead to an erroneous evaluation of the ecosystem dependencies (Irwin, 2006).

Although ecological interdependencies and indirect linkages can be estimated as well by applying separate linear or multiple regressions (e.g. Gough et al., 1994) or analyses of variance (e.g. Dorrepaal et al., 2005), SEM tracks these relationships while allowing the prediction of its variables within a composite model (e.g. Grace et al., 2016).

The variable importance of the random forests models using the UAV-based predictors showed that the canopy 3D information outperformed the spectral information in all cases except in the direct prediction of belowground C stocks (see Supplementary data), where the first MNF component yielded the highest importance. In addition, variables such as the mode and the coefficient of variation of the point cloud also depicted a high importance. This indicates that for predicting belowground C stocks the combination of spectral and canopy 3D information has a high potential.

4.4. Belowground C stocks extrapolation

The prediction maps (Fig. 6) agreed with our knowledge about the study site, where the belowground C stocks has a smoother transition between the conservation and managed areas than the aboveground vegetation variables (Cabezas et al., 2015).

The total amount of C stocks estimated with the hybrid model was ~233,790 kg m^{-2}. This is a rather low value compared to other natural peatlands in Chilean Patagonia (Cabezas et al., 2015), which can be attributed to its young stage (Díaz et al., 2008). Chiloé Island presents many similar isolated

small peatlands, which, despite their relatively low individual C pools, sum up to considerable amounts of total C stocks (McClellan et al., 2017). Given the fact that these wetlands are numerous and isolated, the measurement bias and field sampling costs could be high (Jobe & White, 2009). The lower field operational costs of assessing aboveground vegetation attributes compared to soil samples makes the presented mapping approach suitable to estimate the overall C contribution of the landscape.

Lower CV pixel values were obtained in areas were vegetation variables presented larger variability, whereas higher CV were caused mainly by extremes in aboveground biomass and species richness (Fig. 7). The models resulted in higher uncertainties for the conservation area due to a higher abundance of shrubs, while the managed area presented high uncertainties in eroded grasses with presence of exotic species. This difference is an expression of the low representativeness of pure aboveground biomass and exotic rich plots compared to plots with intermedium gradient values (i.e. only one pure shrub plot in the conservation area and two plots with high number of invasive species in the managed area were included in the analysis).

Maps of the vegetation attributes evaluated in this study can be obtained at different spatial scales. For example, vegetation height can be obtained by airborne LiDAR, photogrammetry based on UAV, airborne or high resolution optical satellites (e.g. WorldView or Pleiades stereo-imagery; Maack et al., 2015) or SAR interferometry (e.g. TanDEM-X; Kattenborn et al., 2015). More research is needed to link such products with ecological studies to assess if the hypothesis developed from local ecological studies hold across spatial scales (Pettorelli et al., 2014). Further work is needed to find stable and mechanistic links between belowground C stocks and vegetation characteristics, which would facilitate a continuous monitoring of soil pools while reducing the amount of destructive soil samples. We believe that SEM is suitable to find ecological meaningful interactions and to link ecosystem properties with remote sensing. Nonetheless, the relationships presented in this investigation needs to be tested against new data to see if their assumptions hold across landscapes.

5. Conclusions

This study evaluated the use of remotely sensed vegetation attributes as proxies to predict peatland belowground C stocks. By using a hybrid model that couples plot-based ecological knowledge of the ecosystem functioning (structural equation modeling) with remote sensing estimates (machine learning), we conclude the following:

1. The use of vegetation characteristics such as vascular vegetation height, aboveground biomass and species richness as proxies resulted in more accurate belowground C stock estimations, compared to the use of raw remote sensing data. The use of vegetation characteristics is advantageous because of their relatively easy and non-destructive assessment. The relationships found for vegetation proxies and belowground C stock is likely to be site specific, and therefore change with geographic position, species composition or state of the peatland. Hence, more investigation is needed to check whether the relationships found here hold across landscapes.
2. The assessment of indirect relationships between the vegetation attributes and belowground C stocks improved the interpretation of variable importance. Considering only direct linkages (coefficients) resulted in underestimation of the vascular species richness and floristic composition contribution to the model.

3. Structural equation models are suitable to track indirect and bi-directional links among components in a defined model. Nonetheless, they cannot handle correctly the number and dimensionality of the remote sensing data, so their integration with machine learning algorithms is appropriate for mapping purposes.
4. The structural equation model showed flexibility to fit the large gradients of peatland vegetation attributes caused by land use differences.
5. Peatland belowground C stocks is difficult to estimate directly using optical and 3D canopy height remote sensing information. Algorithms that take into account more complexity in the predictors, such as random forests and support vector machines, are able to model the relations between canopy reflectance/canopy and the C gradient, obtaining higher accuracies than linear models. Nevertheless, these accuracies were not sufficient for a reliable estimation of belowground C stocks.

The integration of ecological expertise into remote sensing applications has great potential to improve not only the final mapping accuracy, but to contribute to the knowledge of ecosystem functioning and processes. Nevertheless, more investigation is needed to find generalized aboveground proxies of belowground C stocks to effectively decrease the amount of soil samples needed for model calibration.

Acknowledgements

This investigation was funded by the Graduate School for Climate and Environment of the Karlsruhe Institute of Technology [VH-GS-304], the Chilean National Commission for Science and Technology [FONDECYT 1130935] and CR² [CONICYT/FONDAP/15110009]. We further thank Gamaya for their collaboration with the UAV hyperspectral sensor, Julián Cabezas, Ariel Valdés and Jose Ignacio Calderón for their crucial participation in the field survey and to Rocío Araya-Lopez for her valuable comments on the manuscript.

Appendix A. PLS path modeling extra information

Table A1. PLS-PM outer model specifications. The weights correspond to the outer model coefficients while the loading are similar to the correlations in regression analysis. Non-significant predictors (n.s.) are highlighted in gray.

Latent variables	Predictor	Weight (ω)	Loading (*l*)
Vascular vegetation height	Vegetation height	1.00	1.00
Floristic composition	NMDS 1	1.00	1.00
	NMDS 2	n.s.	n.s.
	NMDS 3	n.s.	n.s.
Vascular aboveground biomass	Bryophytes biomass	n.s.	n.s.
	Herbaceous biomass	0.65	0.94
	Shrubs biomass	0.45	0.87
Vascular species richness	Graminoid richness	0.50	0.90
	Forbs richness	0.59	0.93
	Shrub richness	n.s.	n.s.
SOM	Soil organic matter depth	1.00	1.00

C stocks	Belowground C stocks	1.00	1.00

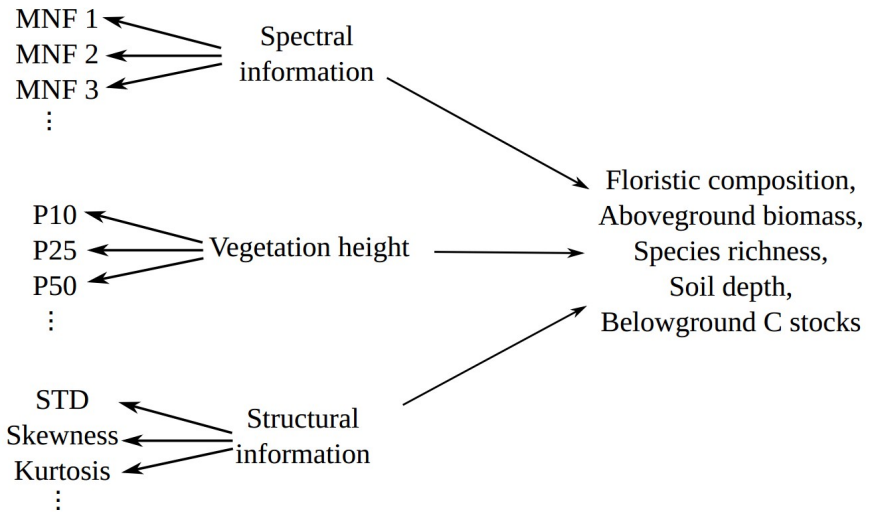

Fig. A1. PLS-PM structural model for the estimation of vegetation proxies, soil depth and belowground C stocks using UAV-based predictors (Fig. 3b).

Appendix B. Floristic composition information

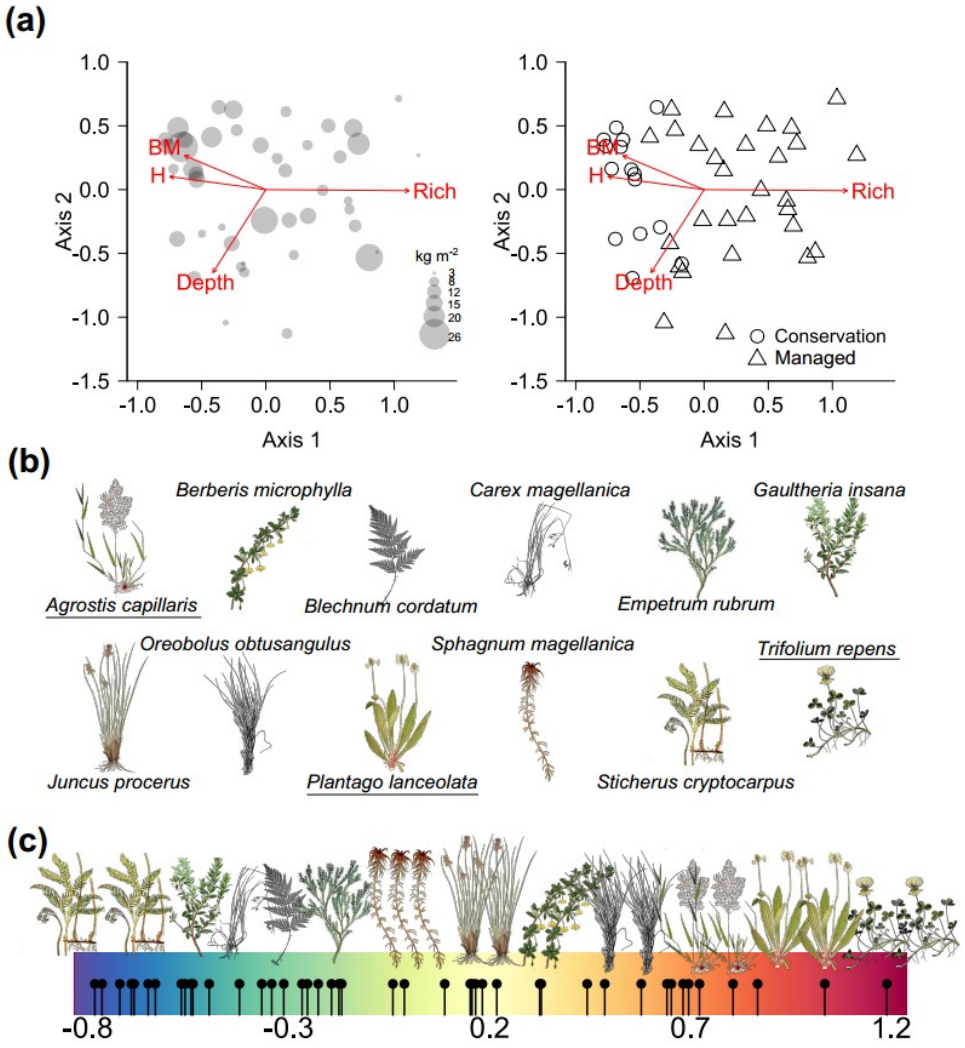

Fig. A2. Results of the floristic gradient. (a) Distribution of plots (sample points) in the two-dimensional ordination space of NMDS, showing the distribution of the belowground C stocks (left) and the management types (right) along the gradients. Close plots feature a similar species composition, while remote plots are more dissimilar. Vectors illustrate the correlations of the axes with the PLS-PM latent variables (H = vascular vegetation height, BM = vascular aboveground biomass, Rich = vascular species richness and Depth = soil depth). (b) Typical species identified with the isopam clustering algorithms in the NMDS first axis. Underlined species names indicate exotic species. (c) Floristic gradient obtained in the NMDS first axis with the plot locations (black dots) and the placement of the commons species on it.

Appendix B. Applied scripts

The R and Python scripts applied in this investigation can be found in the following repository:

https://github.com/JavierLopatin/Peatland-carbon-stock

References

Akumu, C.E., & McLaughlin, J.W. (2014). Modeling peatland carbon stock in a delineated portion of the Nayshkootayaow river watershed in Far North, Ontario using an integrated GIS and remote sensing approach. CATENA 121, 297–306. doi:10.1016/j.catena.2014.05.025

Araya-López, R.A., Lopatin, J., Fassnacht, F.E. & Hernández, H.J. (2018). Monitoring Andean high altitude wetlands in central Chile with seasonal optical data: A comparison between Worldview-2 and Sentinel-2 imagery. ISPRS J. Photogramm. Remote Sens. doi:10.1016/j.isprsjprs.2018.04.001

Beilman, D.W., Vitt, D.H., Bhatti, J.S. & Forest, S. (2008). Peat carbon stocks in the southern Mackenzie River Basin: Uncertainties revealed in a high-resolution case study. Glob. Chang. Biol. 14, 1221–1232. doi:10.1111/j.1365-2486.2008.01565.x

Breiman, L. (2001). Random Forests. Mach. Learn. 45, 5–32.

Cabezas, J., Galleguillos, M. & Perez-Quezada, J.F. (2016). Predicting Vascular Plant Richness in a Heterogeneous Wetland Using Spectral and Textural Features and a Random Forest Algorithm. IEEE Geosci. Remote Sens. Lett. 13, 646–650. doi:10.1109/LGRS.2016.2532743

Cabezas, J., Galleguillos, M., Valdés, A., Fuentes, J.P., Pérez, C. & Perez-Quezada, J. (2015). Evaluation of impacts of management in an anthropogenic peatland using field and remote sensing data. Ecosphere 6, 1–24. doi:http://dx.doi.org/10.1890/ES15-00232.1

Castillo-Riffart, I., Galleguillos, M., Lopatin, J. & Perez-Quezada, J.F. (2017). Predicting Vascular Plant Diversity in Anthropogenic Peatlands: Comparison of Modeling Methods with Free Satellite Data. Remote Sens. 9, 681. doi:10.3390/rs9070681

Chen, H., Mommer, L., van Ruijven, J., de Kroon, H., Fischer, C., Gessler, A., Hildebrandt, A., Scherer-Lorenzen, M., Wirth, C. & Weigelt, A. (2017). Plant species richness negatively affects root decomposition in grasslands. J. Ecol. 105, 209–218. doi:10.1111/1365-2745.12650

Chen, S., Wang, W., Xu, W., Wang, Y., Wan, H., Chen, D., Tang, Z., Tang, X., Zhou, G., Xie, Z., Zhou, D., Shangguan, Z., Huang, J., He, J.-S., Wang, Y., Sheng, J., Tang, L., Li, X., Dong, M., Wu, Y., Wang, Q., Wang, Z., Wu, J., Chapin, F.S. & Bai, Y. (2018). Plant diversity enhances productivity and soil carbon storage. Proc. Natl. Acad. Sci. 115, 4027–4032. doi:10.1073/pnas.1700298114

Cong, W.-F., van Ruijven, J., Mommer, L., De Deyn, G.B., Berendse, F. & Hoffland, E. (2014). Plant species richness promotes soil carbon and nitrogen stocks in grasslands without legumes. J. Ecol. 102, 1163–1170. doi:10.1111/1365-2745.12280

Dargie, G.C., Lewis, S.L., Lawson, I.T., Mitchard, E.T.A., Page, S.E., Bocko, Y.E. & Ifo, S.A. (2017). Age, extent and carbon storage of the central Congo Basin peatland complex. Nature 542, 86–90. doi:10.1038/nature21048

Díaz, M.F., Bigelow, S. & Armesto, J.J. (2007). Alteration of the hydrologic cycle due to forest clearing and its consequences for rainforest succession. For. Ecol. Manage. 244, 32–40. doi:10.1016/j.foreco.2007.03.030

Díaz, M.F., Larraín, J., Zegers, G. & Tapia, C. (2008). Floristic and hydrological characterization of Chiloé Island peatlands, Chile. Rev. Chil. Hist. Nat. 81, 455–468.

Dorrepaal, E. (2007). Are plant growth-form-based classifications useful in predicting northern ecosystem carbon cycling feedbacks to climate change? J. Ecol. 95, 1167–1180.

doi:10.1111/j.1365-2745.2007.01294.x

Dorrepaal, E., Cornelissen, J.H.C., Aerts, R., Wallén, B. & Van Logtestijn, R.S.P. (2005). Are growth forms consistent predictors of leaf litter quality and decomposability across peatlands along a latitudinal gradient? J. Ecol. 93, 817–828. doi:10.1111/j.1365-2745.2005.01024.x

Draper, F.C., Roucoux, K.H., Lawson, I.T., Mitchard, E.T.A., Honorio Coronado, E.N., Lähteenoja, O., Torres Montenegro, L., Valderrama Sandoval, E., Zaráte, R. & Baker, T.R. (2014b). The distribution and amount of carbon in the largest peatland complex in Amazonia. Environ. Res. Lett. 9, 124017. doi:10.1088/1748-9326/9/12/124017

Feilhauer, H., Asner, G.P., Martin, R.E. & Schmidtlein, S. (2010). Brightness-normalized Partial Least Squares Regression for hyperspectral data. J. Quant. Spectrosc. Radiat. Transf. 111, 1947–1957. doi:10.1016/j.jqsrt.2010.03.007

Fenner, N., & Freeman, C. (2011). Drought-induced carbon loss in peatlands. Nat. Geosci 4, 895–900.

Ferner, J., Schmidtlein, S., Guuroh, R.T., Lopatin, J. & Linstädter, A. (2018). Disentangling effects of climate and land-use change on West African drylands' forage supply. Glob. Environ. Chang. 53, 24–38. doi: 10.1016/j.gloenvcha.2018.08.007

Gorham, E. (1991). Northern peatlands: role in the carbon cycle and probable responses to climatic warming. Ecol. Appl. 1. doi:10.2307/1941811

Gough, L., Grace, J.B. & Taylor, K.L. (1994). The Relationship between Species Richness and Community Biomass: The Importance of Environmental Variables. Oikos 70, 271–279.

Grace, J.B., Anderson, T.M., Olff, H. & Scheiner, S.M. (2010). On the specification of structural equation models for ecological systems. Ecol. Monogr. 80, 67–87. doi:10.1890/09-0464.1

Grace, J.B., Anderson, T.M., Seabloom, E.W., Borer, E.T., Adler, P.B., Harpole, W.S., Hautier, Y., Hillebrand, H., Lind, E.M., Pärtel, M., Bakker, J.D., Buckley, Y.M., Crawley, M.J., Damschen, E.I., Davies, K.F., Fay, P.A., Firn, J., Gruner, D.S., Prober, S.M. & Smith, M.D. (2016). Integrative modelling reveals mechanisms linking productivity and plant species richness. Nature 529, 1–10. doi:10.1038/nature16524

Grace, J.B., & Bollen, K.A. (2005). Interpreting the results from multiple regression and structural equation models. Bull. Ecol. Soc. Am. 9623, 296–300. doi:10.1890/0012-9623(2005)86

Grace, J.B., Michael Anderson, T., Smith, M.D., Seabloom, E., Andelman, S.J., Meche, G., Weiher, E., Allain, L.K., Jutila, H., Sankaran, M., Knops, J., Ritchie, M. & Willig, M.R. (2007). Does species diversity limit productivity in natural grassland communities? Ecol. Lett. 10, 680–689. doi:10.1111/j.1461-0248.2007.01058.x

Green, A.A., Berman, M., Switzer, P. & Craig, M.D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Trans. Geosci. Remote Sens. 26, 65–74.

Gumbricht, T., Roman-Cuesta, R.M., Verchot, L., Herold, M., Wittmann, F., Householder, E., Herold, N. & Murdiyarso, D. (2017). An expert system model for mapping tropical wetlands and peatlands reveals South America as the largest contributor. Glob. Chang. Biol. doi:10.1111/gcb.13689

Harris, A., Charnock, R. & Lucas, R.M. (2015). Hyperspectral remote sensing of peatland floristic gradients. Remote Sens. Environ. 162, 99–111. doi:10.1016/j.rse.2015.01.029

Hribljan, J.A., Suarez, E., Bourgeau-Chavez, L., Endres, S., Lilleskov, E.A., Chimbolema, S., Wayson, C., Serocki, E. & Chimner, R.A. (2017). Multidate, multisensor remote sensing reveals high density of carbon-rich mountain peatlands in the páramo of Ecuador. Glob. Chang. Biol. 23, 5412–5425. doi:10.1111/gcb.13807

Irwin, R.E. (2006). The consequences of direct versus indirect species interactions to selection on traits: pollination and nectar robbing in Ipomopsis aggregata. Am. Nat. 167, 315–328. doi:10.1086/499377

Jaenicke, J., Rieley, J.O., Mott, C., Kimman, P. & Siegert, F. (2008). Determination of the amount of carbon stored in Indonesian peatlands. Geoderma 147, 151–158. doi:10.1016/j.geoderma.2008.08.008

Jobe, T.R., & White, P.S. (2009). A new cost distance model for human accessibility and an evaluation of accessibility bias in permanent vegetation plots in Great Smoky Mountains National Park, USA. J. Veg. Sci. 20, 1099–1109. doi:10.1111/j.1654-1103.2009.01108.x

Jonsson, M., & Wardle, D.A. (2009). Structural equation modelling reveals plant community drivers of carbon storage in boreal forest ecosystems. Biol. Lett. 116–119.

Kattenborn, T., Lopatin, J., Förster, M., Braun, A.C., Fassnacht, F.E. (2019). UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data. Remote Sens. Environ. 227, 61–73. doi: 10.1016/j.rse.2019.03.025

Kattenborn, T., Lopatin, J., Kattenborn, G. & Fassnacht, F.E. (2018). Pilot study on the retrieval of DBH and diameter distribution of deciduous forest stands using cast shadows in UAV-based ortomosaics. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-1, 10–12. doi: 10.5194/isprs-annals-IV-1-93-2018

Kattenborn, T., Fassnacht, F., Pierce, S., Lopatin, J., Grime, J.P. & Schmidtlein, S. (2017). Linking plant strategies and plant traits derived by radiative transfer modelling. J. Veg. Sci. 38, 42–49. doi:10.1111/jvs.12525

Kattenborn, T., Maack, J., Faßnacht, F., Enßle, F., Ermert, J. & Koch, B. (2015). Mapping forest biomass from space – Fusion of hyperspectral EO1-hyperion data and Tandem-X and WorldView-2 canopy height models. Int. J. Appl. Earth Obs. Geoinf. 35, 359–367. doi:10.1016/j.jag.2014.10.008

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. Proc. Int. Jt. Conf. Artif. Intell. 2, 1137–1143.

Lange, M., Eisenhauer, N., Sierra, C.A., Bessler, H., Engels, C., Griffiths, R.I., Mellado-Vázquez, P.G., Malik, A.A., Roy, J., Scheu, S., Steinbeiss, S., Thomson, B.C., Trumbore, S.E. & Gleixner, G. (2015). Plant diversity increases soil microbial activity and soil carbon storage. Nat. Commun. 6. doi:10.1038/ncomms7707

Lawson, I.T., Kelly, T.J., Aplin, P., Boom, A., Dargie, G., Draper, F.C.H., Hassan, P.N.Z.B.P., Hoyos-Santillan, J., Kaduk, J., Large, D., Murphy, W., Page, S.E., Roucoux, K.H., Sjögersten, S., Tansey, K., Waldram, M., Wedeux, B.M.M. & Wheeler, J. (2014). Improving estimates of tropical peatland area, carbon storage, and greenhouse gas fluxes. Wetl. Ecol. Manag. 1–20. doi:10.1007/s11273-014-9402-2

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest 2, 18–22.

Lopatin, J., Dolos, K., Hernández, H.J., Galleguillos, M. & Fassnacht, F.E. (2016). Comparing

Generalized Linear Models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. Remote Sens. Environ. 173, 200–210. doi:10.1016/j.rse.2015.11.029

Lopatin, J., Fassnacht, F.E., Kattenborn, T. & Schmidtlein, S. (2017). Mapping plant species in mixed grassland communities using close range imaging spectroscopy. Remote Sens. Environ. 201, 12–23. doi:10.1016/j.rse.2017.08.031

Lopatin, J., Galleguillos, M., Fassnacht, F.E., Ceballos, A. & Hernández, J. (2015). Using a Multistructural Object-Based LiDAR Approach to Estimate Vascular Plant Richness in Mediterranean Forests With Complex Structure. IEEE Geosci. Remote Sens. Lett. 12, 1008–1012. doi: 10.1109/LGRS.2014.2372875

Lopatin, J., Dolos, K., Kattenborn, T, & Fassnach, F. (2019). How canopy shadow affects invasive plant species classification in high spatial resolution remote sensing. Remote Sens. Ecol. Conserv. doi: 10.1002/rse2.109

Ma, Q., Cui, L., Song, H., Gao, C., Hao, Y., Luan, J., Wang, Y. & Li, W. (2017). Aboveground and Belowground Biomass Relationships in the Zoige Peatland, Eastern Qinghai–Tibetan Plateau. Wetlands 37, 461–469. doi:10.1007/s13157-017-0882-8

Maack, J., Kattenborn, T., Fassnacht, F.E., Hernández, J., Corvalán, P. & Koch, B. (2015). Modeling forest biomass using Very-High-Resolution data—Combining textural, spectral and photogrammetric predictors derived from spaceborne stereo images. Eur. J. Remote Sens. 48, 245-261.

McClellan, M., Comas, X., Benscoter, B., Hinkle, R. & Sumner, D. (2017). Estimating Belowground Carbon Stocks in Isolated Wetlands of the Northern Everglades Watershed, Central Florida, Using Ground Penetrating Radar and Aerial Imagery. J. Geophys. Res. Biogeosciences 122, 2804–2816. doi:10.1002/2016JG003573

McGaughey, R.J. (2018). FUSION/LDV: Software for LIDAR Data Analysis and Visualization.

Mercer, J.J. & Westbrook, C.J. (2016). Ultrahigh-resolution mapping of peatland microform using ground-based structure from motion with multiview stereo. J. Geophys. Res. Biogeosciences 121, 2901–2916. doi: 10.1002/2016JG003478

Parish, F., Sirin, A., Charman, D., Joosten, H., Minayeva, T., Silvius, M. & Stringer, L. (2008). Assessment on peatlands, biodiversity and climate change: main report.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Pettorelli, N., Safi, K. & Turner, W. (2014). Satellite remote sensing, biodiversity research and conservation of the future. Philos. Trans. R. Soc. London B Biol. Sci. 369, 20130190.

Phillips, R., & Beeri, O. (2008). The role of hydropedologic vegetation zones in greenhouse gas emissions for agricultural wetland landscapes. Catena 72, 386–394. doi:10.1016/j.catena.2007.07.007

Rocchini, D., Bacaro, G., Chirici, G., Re, D. Da, Feilhauer, H., Foody, G.M., Galluzzi, M., Garzon-Lopez, C.X., Gillespie, T.W., He, K.S., Lenoir, J., Marcantonio, M., Nagendra, H., Ricotta, C., Rommel, E., Schmidtlein, S., Skidmore, A.K., Kerchove, R. Van De, Wegmann, M. & Rugani,

B. (2018). Remotely sensed spatial heterogeneity as an exploratory tool for taxonomic and functional diversity study. Ecol. Indic. 85, 983–990. doi:https://doi.org/10.1016/j.ecolind.2017.09.055

Rudiyanto, Minasny, B., Setiawan, B.I., Saptomo, S.K. & McBratney, A.B. (2018). Open digital mapping as a cost-effective method for mapping peat thickness and assessing the carbon stock of tropical peatlands. Geoderma 313, 25–40. doi:10.1016/j.geoderma.2017.10.018

Sanchez, G., Trinchera, L. & Russolillo, G. (2017). plspm: Tools for Partial Least Squares Path Modeling (PLS-PM).

Schaepman-Strub, G., Limpens, J., Menken, M., Bartholomeus, H.M. & Schaepman, M.E. (2008). Towards spatial assessment of carbon sequestration in peatlands: spectroscopy based estimation of fractional cover of three plant functional types. Biogeosciences Discuss. 5, 1293–1317. doi:10.5194/bgd-5-1293-2008

Schmidtlein, S., Feilhauer, H. & Bruelheide, H. (2012). Mapping plant strategy types using remote sensing. J. Veg. Sci. 23, 395–405. doi:10.1111/j.1654-1103.2011.01370.x

Schmidtlein, S., & Sassin, J. (2004). Mapping of continuous floristic gradients in grasslands using hyperspectral imagery. Remote Sens. Environ. 92, 126–138. doi:10.1016/j.rse.2004.05.004

Smith, L.C., MacDonald, G.M., Velichko, A.A., Beilman, D.W., Borisova, O.K., Frey, K.E., Kremenetski, K. V. & Sheng, Y. (2004). Siberian Peatlands a Net Carbon Sink and Global Methane Source since the Early Holocene. Science (80-.). 303, 353–356. doi:10.1126/science.1090553

Svetnik, V., Liaw, A., Tong, C., Culberson, J., Sheridan & R. Feuston, B. (2003). Random forest: A classification and regression tool for compound classification and QSAR modeling. J. Chem. Inf. Comput. Sci. 43, 1947–1958.

Tenenhaus, M., Esposito, V., Chatelin, T.-M. & Lauro, C. (2005). PLS path modeling. Comput. Stat. Data Anal. 48, 159–205.

Turetsky, M.R., Kane, E.S., Harden, J.W., Ottmar, R.D., Manies, K.L., Hoy, E., Kasischke, E.S. & 5 (2011). Recent acceleration of biomass burning and carbon losses in Alaskan forests and peatlands. Nat. Geosci. 4, 27–31.

van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E. & Yu, T. (2014). Scikit-image: image processing in Python. PeerJ 2, e453. doi:10.7717/peerj.453

Vuolo, F., Neugebauer, N., Bolognesi, S.F., Atzberger, C. & D'Urso, G. (2013). Estimation of Leaf Area Index Using DEIMOS-1 Data: Application and Transferability of a Semi-Empirical Relationship between two Agricultural Areas. Remote Sens. 5, 1274–1291. doi:10.3390/rs5031274

Walker, T.N., Garnett, M.H., Ward, S.E., Oakley, S., Bardgett, R.D. & Ostle, N.J. (2016). Vascular plants promote ancient peatland carbon loss with climate warming. Glob. Chang. Biol. 22, 1880–1889. doi:10.1111/gcb.13213

Weinacker, H., Koch, B. & Weinacker, R. (2004). TREESVIS—a software system for simultaneous 3D-realtime visualization of DTM, DSM, laser raw data, multispectral data, simple tree and building models. Int. Arch. Photogrametry, Remote Sens. Spat. Inf. Sci. XXXVI, 90–95.