Methods

BERM: a Belowground Ecosystem Resiliency Model for estimating *Spartina alterniflora* belowground biomass

Jessica L. O'Connell[1] (iD), **Deepak R. Mishra**[2] (iD), **Merryl Alber**[3] (iD) and **Kristin B. Byrd**[4] (iD)

[1]Department of Marine Science, University of Texas at Austin, Port Aransas, TX 78373, USA; [2]Department of Geography, University of Georgia, Athens, GA 30602-3636, USA;

[3]Department of Marine Sciences, University of Georgia, Athens, GA 30602-3636, USA; [4]US Geological Survey, Moffett Field, CA 94035, USA

Summary

- Spatiotemporal patterns of *Spartina alterniflora* belowground biomass (BGB) are important for evaluating salt marsh resiliency. To solve this, we created the BERM (Belowground Ecosystem Resiliency Model), which estimates monthly BGB (30-m spatial resolution) from freely available data such as Landsat-8 and Daymet climate summaries.
- Our modeling framework relied on extreme gradient boosting, and used field observations from four Georgia salt marshes as ground-truth data. Model predictors included estimated tidal inundation, elevation, leaf area index, foliar nitrogen, chlorophyll, surface temperature, phenology, and climate data. The final model included 33 variables, and the most important variables were elevation, vapor pressure from the previous four months, Normalized Difference Vegetation Index (NDVI) from the previous five months, and inundation.
- Root mean squared error for BGB from testing data was 313 g m^{-2} (11% of the field data range), explained variance (R^2) was 0.62–0.77. Testing data results were unbiased across BGB values and were positively correlated with ground-truth data across all sites and years ($r = 0.56$–0.82 and 0.45–0.95, respectively).
- BERM can estimate BGB within *Spartina alterniflora* salt marshes where environmental parameters are within the training data range, and can be readily extended through a reproducible workflow. This provides a powerful approach for evaluating spatiotemporal BGB and associated ecosystem function.

Introduction

Spatiotemporal belowground biomass (BGB) patterns from tidal marsh vegetation provide an important window into ecosystem function and resiliency. High BGB in tidal marshes increases soil organic matter and promotes vertical accretion by increasing soil volume and resisting erosion (Kirwan & Guntenspergen, 2012; Cahoon *et al.*, 2021). Where BGB is high, marsh elevation is higher and marshes are more resilient to sea level rise (Nyman *et al.*, 2006; Mudd *et al.*, 2009; Kirwan & Megonigal, 2013). Marshes with greater BGB also tend to sequester more carbon, depositing as much as 1713 g C m^{-2} yr^{-1} in soils, termed as 'blue carbon' (Mcleod *et al.*, 2011). Therefore, BGB assessments can identify vulnerable marshes vs those that may continue to provide valuable services such as carbon (C) sequestration. This is particularly important as sea level rise is accelerating and the Anthropocene has seen global marsh losses of 50% to 65% (Dahl, 1990; Lotze *et al.*, 2006; Gedan & Silliman, 2009). Previous landscape estimates of BGB typically relied either on time consuming field sampling (see Stagg *et al.*, 2017), or landscape

modeling that depends on a constant root : shoot ratio, which misses spatiotemporal variability (Morris *et al.*, 2002; Mudd *et al.*, 2009; Langston *et al.*, 2021). This article presents a new solution for estimating near-real time spatiotemporal trajectories in salt marsh BGB, the Belowground Ecosystem Resiliency Model (BERM). BERM provides landscape-scale estimates of BGB of the salt marsh plant, *Spartina alterniflora*, Loisel (Kartesz, 2015; USDA & NRCS, 2019) (=*Sporobolus alterniflorus*; Peterson *et al.*, 2014a,b), a cosmopolitan herbaceous graminoid that typically grows in monoculture. We focused on *Spartina alterniflora* because it is found on all three United States coasts and on nearly every continent as a native or colonizing species (Pennings & Bertness, 2001; Strong & Ayres, 2016). To estimate *Spartina alterniflora* BGB, BERM uses a combination of readily available climate datasets, digital elevation model (DEM) data, and remote sensing derived aboveground biophysical variables. The procedure employed in BERM can be adapted to evaluate BGB patterns of other plant species and communities in the future. This article aims to introduce BERM and provide examples of the model output.

BERM estimates BGB on a monthly time-step and includes variables related to gross and net primary production (GPP/NPP). GPP variables can constrain and partition estimates of biomass, given that GPP = NPP + autotrophic respiration, and NPP = total annual BGB production + total annual aboveground biomass (AGB) production. GPP/NPP can be readily derived through remote sensing methods (Ruimy *et al.*, 1994; Running *et al.*, 2004; Turner *et al.*, 2006; Gitelson *et al.*, 2014; Tao *et al.*, 2018). For example, GPP is often estimated through canopy photosynthesis models that assume a close relationship between GPP, canopy chlorophyll (CHL) concentration, and leaf area index (LAI), both often measured through remote sensing techniques (Gitelson *et al.*, 2006, 2014). LAI provides an index of the leaf area over which CHL is distributed, together signaling photosynthetic capacity.

Production efficiency models are another method for estimating GPP/NPP. For NPP, the productivity quantity most related to BGB, such models rely on the fraction of absorbed photosynthetically active radiation (fAPAR), downwelling PAR, and production efficiency (ε), e.g. the conversion efficiency of light energy into organic matter (Monteith, 1972; Kumar & Monteith, 1981; Ruimy *et al.*, 1994). This approach can be adapted to a remote sensing framework by substituting vegetation indices for fAPAR, such as the Normalized Difference Vegetation Index (NDVI) (e.g. Ruimy *et al.*, 1994). Furthermore, ε is a complex variable that represents the combined effects of photosynthetic efficiency (often referred to as light use efficiency) and respiration losses, or C use efficiency. The ε value is usually estimated as a function of temperature, soil moisture, vapor pressure (used to estimate vapor pressure deficit/evapotranspiration), or PAR (e.g. Ruimy *et al.*, 1994; Xiao *et al.*, 2011; Barr *et al.*, 2013; Massmann *et al.*, 2019; Hawman *et al.*, 2021), and by including these variables in our model we potentially capture quantities that drive ε.

In marshes, high salinity can cause physiological stress (Barr *et al.*, 2013), for which marsh elevation, temperature, tidal flooding, and precipitation might serve as readily-available proxies. This is because higher elevations receive less soil flushing from tidal flooding, higher temperatures increase evapotranspiration and concentrate soil salts, and precipitation decreases soil salinity through freshwater dilution (Mendelssohn & Morris, 2002). An added complication is that tidal flooding can either stimulate or reduce production. Shallow and abbreviated tidal flooding stimulates production by flushing salt from marsh soils and delivering dissolved nutrients (Mendelssohn & Morris, 2002). Deep and prolonged flooding conversely reduces plant available oxygen, creating metabolic deficits, and, when extreme, can result in plant death and marsh loss (Morris *et al.*, 2002; Kirwan & Megonigal, 2013). Thus, including biophysical variables that affect salinity is likely important for predicting GPP/NPP and therefore BGB.

We also needed to account for the partitioning of biomass between aboveground vs belowground production. For this, we relied on estimates of AGB and foliar nitrogen (N) concentration, as these are linked with BGB in theoretical and empirical studies (Morris *et al.*, 2013; O'Connell *et al.*, 2014, 2015). We previously used foliar N concentration and AGB to explain up to 86% of the variation in root : shoot ratios and 76% of end-of-growing season BGB in freshwater *Scheonoplectus acutus* marshes (O'Connell *et al.*, 2014, 2015). Foliar N is an integrative measure of plant available N (O'Connell *et al.*, 2015), is a recognized driver of root productivity (Morris *et al.*, 2013), and contributes to *Spartina alterniflora* NPP (Morris, 1982). Further, root : shoot ratios generally decrease with increasing canopy foliar N across plant species (Levin *et al.*, 1989; Dewar, 1993; McConnaughay & Coleman, 1999; Deegan *et al.*, 2012; Morris *et al.*, 2013; O'Connell *et al.*, 2014, 2015), especially in coastal wetlands where growth is N-limited (Morris *et al.*, 2013).

BERM estimates BGB at a monthly time-step. To accomplish this, we measured the timing of plant spring green-up and monthly variation in the relationship between BGB and aboveground variables such as AGB. These relationships vary seasonally because *Spartina alterniflora* canopies senesce in the winter and sprout in spring from underground rhizomes (Gallagher *et al.*, 1984). To support spring growth, belowground plant resources are nearly depleted through translocation into aboveground tissues (Gallagher, 1983). Belowground reserves are then gradually returned, but are not fully replenished until late summer (Gallagher, 1983). Thus, AGB, BGB, and root : shoot ratios vary seasonally. Further, spring green-up dates for *Spartina alterniflora* can vary by > one month across short distances, driven by elevation-related differences in marsh soil temperature (O'Connell *et al.*, 2020). Thus, plants meters apart can be in different phenological stages. Our modeling process accounts for spatiotemporal phenology, and aboveground vs belowground variation by estimating the start of the growing season on a pixelwise basis.

To parameterize an easily deployable model, we relied on freely available data and utilized an open-science research process. Vegetation parameters we tested as potential inputs in BERM included those listed in Table 1, which represent metrics related to plant resource availability, photosynthetic and production capacity, phenology, and environmental stress. BERM inputs were combined through a machine learning framework capable of capturing complex relationships. Table 1 parameters can either be estimated through remote sensing methods (Kokaly *et al.*, 2009; Mishra *et al.*, 2012; Byrd *et al.*, 2014; O'Connell *et al.*, 2014, 2015, 2020; Ghosh *et al.*, 2016; Alber & O'Connell, 2019), or through existing datasets such as Daymet or the three-dimensional (3D) digital elevation map (Thornton *et al.*, 2017; US Geological Survey, 2019). Additionally, we used a reproducible research flow to document our modeling approach and include an open-source model code DOI with this manuscript, allowing continued use and development of BERM.

BERM predicts monthly gridded site-wide spatiotemporal variation in *Spartina alterniflora* BGB. We accomplished this by: (1) creating a comprehensive remotely derived dataset of potential aboveground BGB proxies; (2) applying a machine learning approach to model BGB through a reproducible research flow; (3) evaluating model performance across training and testing datasets; and (4) providing example BERM output to demonstrate its utility. We also describe BERM development and discuss its applications and current limitations. This is the first step

Table 1 We tested 131 potential features for inclusion as *Spartina alterniflora* belowground biomass model predictors, listed here along with their data sources.

Potential features	Data source
Elevation (m)	Digital elevation model (DEM)
Green-up day of year (DOY)	DOY where Σ(mean daily soil temp > 9.9°C) > 202
Days since green-up	Calculated from green-up day of year
Aboveground biomass (AGB) (g m^{-2})[a,b,c,d,e]	Modeled from Landsat-8 data
Percent foliar chlorophyll (CHL) (mg chlorophyll g^{-1} dried leaf tissue)[a,b,c,d,e]	Modeled from Landsat-8 data
Leaf Area Index (LAI)[a,b,c,d,e]	Modeled from Landsat-8 data
Total foliar nitrogen (N)[a,b,c,d,e]	AGB × percent foliar N
Percent foliar N (g N g^{-1} dried leaf tissue)[a,b,c,d,e]	Modeled from Landsat-8 data
Normalized Difference Vegetation Index (NDVI)[a,b,c]	Calculated from Landsat-8 bands
Inundation intensity[a,b]	(Mean Higher High Water (MHHW) (from National Oceanic and Atmospheric Administration (NOAA) station) − elevation (from a DEM))/(MHHW − mean sea level (from NOAA station)) × percent flooded observed
Dry intensity[a,b]	(elevation (from DEM) − Mean Lower Low Water (MLLW) (from NOAA station))/(mean sea level (from NOAA station) − MLLW) × percent dry observed
Percent flooded observed	Modeled from Landsat-8 data
Land surface temperature (LST) (°C)[a,b,c]	Calculated from Landsat-8 bands
Day length (s d^{-1})	Daymet
Total daily photosynthetically active radiation (PAR) (MJ m^{-2} d^{-1})[a,b]	Calc. from Daymet vars: (2.114 × shortwave radiation × day length)/ 1 000 000
Precipitation (mm d^{-1})[a,b]	Daymet
Daily maximum temperature (°C)[a,b]	Daymet
Daily minimum temperature (°C)[a,b]	Daymet
Daily average vapor pressure (Pa)[a,b]	Daymet

[a]Included lags of the previous 1–5 months of the variable.
[b]Included rolling previous 1–5 months mean of the variable.
[c]Included change over the previous 1–5 months in the variable.
[d]Included change in the variable over growing season.
[e]Included the variable value from the end of previous growing season.

towards enabling long-term coast-wide assessments of BGB in *Spartina alterniflora* marshes.

Materials and Methods

Overview

To develop BERM, we conducted a comprehensive field campaign to collect biophysical ground-truth data for model training. We began by creating a series of extreme gradient boosting (EXB) machine learning models to predict aboveground vegetation parameters related to plant performance, resource availability, photosynthetic capacity, and phenology (AGB, LAI, CHL, foliar N) (Table 1), where Landsat-8 data served as the predictors. We also gathered other environmental data, as listed in Table 1. We then calculated derived aboveground vegetation and environmental variables (Table 1). Next, we combined the aboveground predictors in a final EXB model, and included a parameter selection step to remove unimportant variables. Once parameterized, BERM relies only on these finally selected datasets to make a prediction, and can theoretically estimate any *Spartina alterniflora* marsh pixel over the available time series of Landsat-8 and Daymet, provided the input parameters are within the parameter space of our training data.

Study sites

BERM calibration data came from four salt marshes along the Georgia coast, USA, including flux tower marsh A (measured 2013–2019), and three sites sampled during the 2016 growing season, flux tower marsh B, UGAMI marsh, and Skidaway. All sites were dominated by *Spartina alterniflora* monocultures and experienced similar environmental conditions (Fig. 1; Table 2). Flux tower marshes A and B were in relatively undisturbed marshes 1. 5 km from each other, *c.* 620 m north and southeast of an eddy covariance carbon flux tower operated by the Georgia Coastal Ecosystems Long-Term Ecological Research (GCE-LTER) project on Sapelo Island, GA. Flux tower marsh B was in a lower elevation area with greater tidal creek density than flux tower marsh A. UGAMI marsh, adjacent to the University of Georgia Marine Institute also on Sapelo Island, was a formerly impounded marsh that experienced altered hydrological cycles (Craft, 2001). The Skidaway marsh was on Skidaway Island, GA, just south of Savannah, GA, in a more urbanized watershed. Each marsh consisted of vegetated areas traversed by tidal channels. Tall (> 70 cm), medium (30–70 cm) and short (< 30 cm) *Spartina alterniflora* height-forms were found at each site along elevation-driven marsh edge to interior gradients. This characteristic height-form zonation is caused by elevation-related environmental factors such as the degree of tidal flooding, soil salinity, and soil anoxia. The marsh interior represents a higher elevation area with less frequent flooding, creating a higher salinity, stressful environment with shorter vegetation (Mendelssohn & Morris, 2002). Tides were semi-diurnal with a range of ~1.2 m over the vegetation.

Field data

We collected field data to estimate AGB, BGB, LAI, foliar N, and CHL concentration. AGB was estimated allometrically from stem height and counts within permanent 1 m^2 plots. Leaf CHL was estimated via a SPAD chlorophyll meter (Spectrum Technologies, Aurora, IL, USA), calibrated against leaf-level CHL extractions (see Supporting Information Methods S1; Fig. S1). Foliar N concentration was estimated via a CHN Analyzer (see Methods S2). LAI was measured via an AccuPAR LP-80
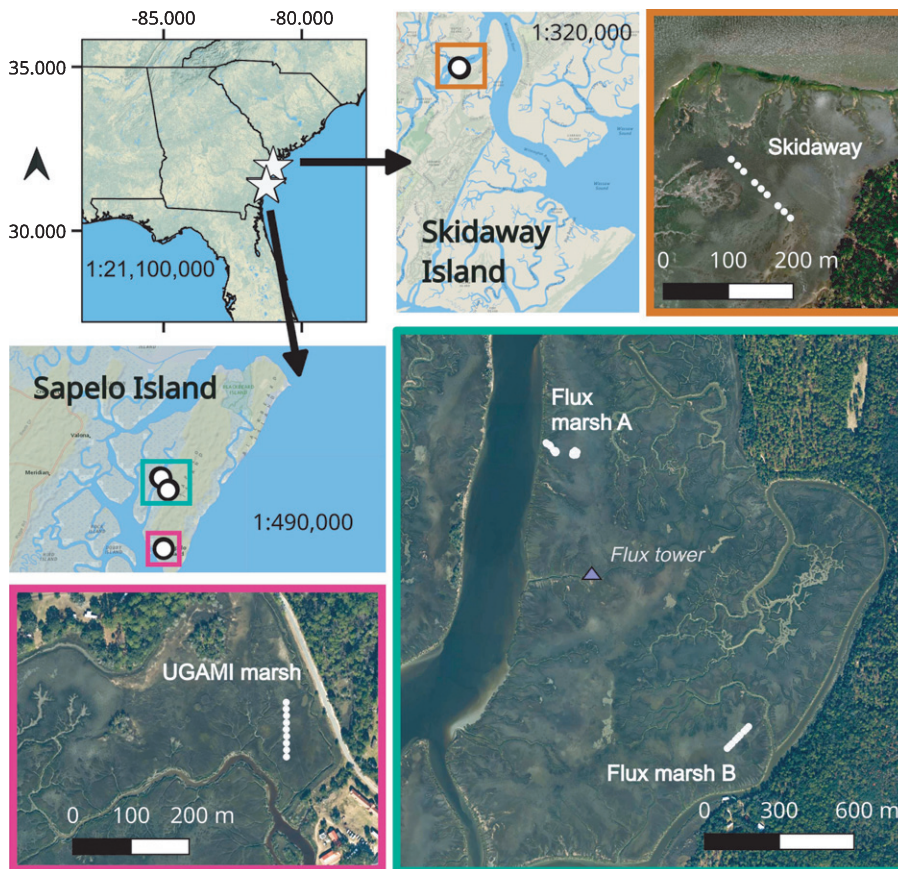
Fig. 1 Field data collections in our study came from four *Spartina alterniflora* salt marshes in Georgia, USA. Flux tower marshes A and B were adjacent to an eddy covariance carbon flux tower on Sapelo Island, GA. UGAMI marsh, also on Sapelo Island, was an impounded marsh that experienced altered hydrological cycles. The Skidaway marsh was on Skidaway Island, GA, just south of Savannah, GA, in a more urbanized watershed. In the overview map (top left), white stars show the locations of Skidaway (north star) and Sapelo (south star) islands. White circles in the close-up maps indicate locations of vegetation plots.

Ceptometer (Meter Environment, Pullman, WA, USA), which measures LAI via canopy light transmittance. For each plot, LAI was calculated from the average of two downwelling (above canopy) and three upwelling (below canopy) radiation measurements. To estimate BGB, we collected standing root cores (7.62 cm × 30 cm) near the permanent plots. Cores were centered on a plant stem clump. Clumps averaged three stems (range: 1–9 stems). Aboveground and belowground core material was washed on 1-mm mesh, sorted into live and dead vegetation, and oven-dried at 60°C to constant mass. Live BGB material was identified by root turgor, color, and texture. We estimated plot-scale BGB by multiplying the stem counts from the permanent vegetation plots by the core live root : shoot ratio, a scaling method that accounts for the stem variation between low-density tall-form plants and high-density short-form plants. The number of field samples differed among sites and variables (see Methods S3, Table S1).

We also estimated the start of the *Spartina alterniflora* growing season. To help with this, we collected soil temperature data, which predicts the annual day of spring green-up (O'Connell *et al.*, 2020). We collected soil temperature with temperature probes (Hobo UA-002-08; Onset Computer Corp., Bourne, MA, USA), buried 10 cm into the soil at 18 points along a marsh edge to interior transect. These probes collected measurements at 15-min intervals during a month-long field campaign from January to February 2018 (see Alber & O'Connell, 2019). We used this information to calibrate a soil temperature model and

estimate the start of the growing season for each pixel (see Section 'Estimating the day of spring green-up').

Deriving comprehensive remotely-sensed datasets of potential aboveground proxies for predicting BGB

Preprocessing Landsat-8 data We used the field data to train remote sensing models developed from the Tier 1 Landsat-8 surface reflectance product, freely available from the US Geological Survey (USGS). Landsat-8 provides moderate resolution (30 m × 30 m) multispectral data with global coverage and 16-d return intervals, and collects observations close to solar noon. Image acquisition time is provided in the image metadata. At our location, multiple Landsat-8 scenes overlap, resulting in up to four observations/month/pixel. We associated our field data with their spatiotemporally corresponding Landsat-8 pixel footprints for the period of observation (June 2013–August 2019, Flux tower marsh A: five pixels/scene; Other sites: three pixels/scene). We averaged field observations from the same date and pixel before model training (3–5 observations/pixel/date). We preprocessed the Landsat-8 data by removing pixels representing clouds and cloud shadows, identified via the Landsat-8 'pixel_qa' mask, a quality control flag (Foga *et al.*, 2017). We additionally removed pixels representing flooded marsh conditions, identified via an in-house flood model. This random forest model predicts binary flood status (dry/flooded) from Landsat-8 surface reflectance, where images from the 'GCESapelo' PhenoCam served as

Table 2 Mean values for ground-truth data representing the key characteristics used to estimate *Spartina alterniflora* belowground biomass for each site.

Variable	Flux marsh A	Flux marsh B	Skidaway	UGAMI marsh
Belowground biomass (BGB) (g m^{-2})	1256 (481–2495)	640 (191–1062)	461 (217–918)	1503 (799–2221)
Above ground biomass (AGB) (g m^{-2})	336 (137–788)	161 (98–221)	237 (124–440)	264 (152–479)
Leaf area index (LAI)	1.05 (0.41–2.28)	0.60 (0.30–1.19)	0.75 (0.39–1.01)	1.21 (0.86–1.78)
Percent foliar nitrogen (N)	0.98 (0.80–1.32)	0.89 (0.68–1.15)	0.67 (0.52–0.89)	0.98 (0.72–1.45)
Percent foliar chlorophyll (CHL)	0.66 (0.50–0.85)	0.57 (0.41–0.72)	0.58 (0.48–0.68)	0.75 (0.60–0.99)
Green-up day of year	34 (6–51)	46 (46–46)	47 (46–49)	47 (46–49)
Elevation (m)	0.73 (0.65–0.80)	0.72 (0.69–0.75)	0.79 (0.72–0.87)	0.80 (0.75–0.85)
Inundation intensity	0.13 (0.02–0.48)	0.10 (0.05–0.17)	0.08 (0.02–0.23)	0.06 (0.03–0.11)
Percent flooded observed	0.27 (0.05–0.70)	0.22 (0.13–0.29)	0.20 (0.10–0.38)	0.15 (0.08–0.21)
Daily maximum temperature (°C)	31 (27–35)	31 (27–35)	31 (27–35)	31 (27–35)
Daily minimum temperature (°C)	22 (17–25)	22 (18–25)	21 (16–24)	22 (18–25)
Daily mean precipitation (mm d^{-1})	3.45 (1.29–9.10)	3.44 (1.23–9.33)	4.90 (1.87–8.06)	3.44 (1.26–9.23)

Numbers in parenthesis indicate the range observed (minimum–maxium) across the dataset.

ground-truth data. The 'GCESapelo' PhenoCam is part of the National PhenoCam network and auto-collects digital imagery of the *Spartina alterniflora* marsh area near the flux tower every 30 min year-round, and clearly depicts flooding events (see O'Connell & Alber, 2016). We also used upland and permanent creek observations in the model as end-points of homogeneous pixels (100% cover of dry upland or deep water). This model used 15 579 pixel observations (1195 were flooded observations), 7848 of which were withheld from model training as testing data (including 359 flooded observations). Within the testing data, the model was 99% accurate for identifying flooded pixels and misclassified nine observations as flooded that were actually dry.

After removing clouds and flooding from the Landsat-8 time series, lower elevations with more flooding had more gaps and thus lower observation temporal density. For example, for Landsat pixels corresponding with our field data, pixels with elevations of 0.6, 0.7, 0.8 and 0.9 m had 40%, 22%, 9% and 11% of pixels classified as flooded, respectively. Note that most of the pixels were at elevations of 0.7 and 0.8 m (e.g. 32% and 51% of data, respectively), which is similar to the elevation distribution within vegetated *Spartina alterniflora* marsh in the broader study area.

Before joining the field and satellite data, we used linear interpolation to fill gaps in the field data time series, providing a field estimate that was as close as possible to the Landsat-8 overpass date. For field data, we did not extrapolate to gap-fill data outside of our field season, and there were no gaps > 1 month. We could gap-fill the field data because we assumed plant production was progressive through time, which may cause us to miss small disturbances, but allows us to track annual cycles in vegetation, as the marsh did not experience severe disturbances that would change plant trajectories. The result of these steps was a regularly spaced field dataset that could be joined to the remote sensing data despite remote sensing data gaps caused by clouds and tides.

EXB modeling After pre-processing, the next step was to predict aboveground parameters from the remote sensing data. We began with physical quantities related to aboveground vegetation: AGB, CHL, LAI, and foliar N (Table 1). Our process was to create a

series of EXB models for these variables, which could then be used to estimate BGB through a final EXB model.

EXB is a machine learning tool developed as a scalable implementation of the gradient boosting framework (Friedman *et al.*, 2000; Friedman & Meulman, 2003; Elith *et al.*, 2008; Chen & Guestrin, 2016), available in the R 'XGBOOST' package. EXB is an ensemble method that combines many weak learners, e.g. machine learning predictive models, into a stronger overall predictor that is both computationally efficient and minimizes overfitting (Friedman, 2001; Chen & Guestrin, 2016). EXB does not have model assumptions, trees are not independent of one another, and is robust to the inclusion of correlated and non-informative predictors (Friedman & Meulman, 2003; Elith *et al.*, 2008; Chen *et al.*, 2018).

EXB modeling workflow Our EXB modeling workflow employed nested resampling (Bischl *et al.*, 2012) (Fig. S2), which divides the data through inner and outer cross-validated data folds. This fits the model multiple times, which is computationally expensive but avoids model bias and overfitting (Bischl *et al.*, 2012) (see Methods S4). During the creation of training and testing sets, we also kept observations from the same site and date together, a procedure known as spatial cross-validation, which helps account for spatial auto-correlation within a site to avoid overfitting (Schratz *et al.*, 2019) (Fig. S2). We additionally stratified by site to ensure that training and testing sets contained observations from all sites, and were thus more representative of the larger data.

The outer resampling for intermediate aboveground models used only a single cross-validation, consisting of a single training and testing set (65% and 35% of data, respectively). To create the inner resampling in all models, we divided the training data from the outer fold only once into an inner training and testing set (85% and 15% of the outer training data, respectively). As mentioned earlier, this inner resampling was used to tune the model hyperparameters on the inner training data, and validated against the inner testing data before model fitting. To select hyperparameters, we used a grid search with a resolution of 20,

and selected the hyperparameters that best fit the data as indicated by a minimized root mean squared error (RMSE) between the observed and predicted inner testing data outcome, where RMSE was: $RMSE = ((\Sigma(obs_i - pred_i)^2)/n)^{(1/2)}$ for each variable, obs_i and $pred_i$ were the observed and predicted values for each sample, and n was the sample size (Hyndman & Koehler, 2006). For the final BGB model, we used five cross-validated folds in the outer resampling, resulting in five inner training and testing sets for model tuning and five outer training and testing sets for final model fitting (Fig. S2).

We evaluated the goodness of model fit for the testing data in the outer resampling only, based on mean absolute error (MAE), RMSE, normalized RMSE (nRMSE), coefficient of variation of RMSE (COV RMSE) and the Pearson's correlation coefficient, r, between the field collected ground-truth data and the predicted BGB outcome. For these calculations, we used $MAE = \Sigma(obs_i - pred_i)/n$ (Willmott & Matsuura, 2005), $nRMSE = RMSE/(obs_{max} - obs_{min}) \times 100$, where obs_{max} and obs_{min} were the maximum and minimum observed values, respectively, and COV $RMSE = RMSE/(obs_{mean}) \times 100$.

EXB models of aboveground vegetation We employed the model workflow (see Section 'EXB modeling workflow') to create EXB models for AGB, LAI, CHL, and foliar N, where the field data, including gap-filled estimates, were the response variables. As predictors, we used Landsat-8 spectral reflectance Bands (B) 1–8, vegetation indices (NDVI, Visible Atmospherically Resistant Index (VARI)), and a phenology index (Rouse *et al.*, 1974; Gitelson *et al.*, 2002; O'Connell *et al.*, 2017).

$$NDVI = (B5 - B4)/(B5 + B4)$$

$$VARI = (B3 - B4)/(B3 + B4 + B2)$$

$$Phenology\,Index = (B5 - B6)/(B6 + B5)$$

We chose these indices because they track vegetation but represent different aspects of growth, as they were not correlated (Pearson's r of 0.30, −0.03, and 0.38, respectively for correlations among NDVI vs VARI, NDVI vs Phenology Index, and VARI vs Phenology Index).

Once we had working biophysical models, we predicted these parameters for any Landsat-8 pixel representing *Spartina alterniflora* marsh across the available time series of Landsat-8, which at the time of analysis was May 2013 to December 2019. In this way, we estimated aboveground biophysical proxies even for months when we did not have corresponding field data (e.g. at Flux marsh B, UGAMI marsh, and Skidaway).

Gathering additional aboveground physical variables needed to estimate belowground biomass (BGB) We also acquired gridded climate data to inform our BGB model, including land surface temperature (LST) as estimated from Landsat-8 thermal bands, derived following methods described in Alber & O'Connell (2019). Additionally, we used Daymet 1-km products, which are daily surface weather and climatological summaries including

daylight (in seconds), total daily precipitation (in millimeters), incident shortwave radiation flux density (average over the daylight period in W m^{-2}), maximum and minimum air temperature (daily 2-m air temperature in °C), and vapor pressure (daily average partial pressure of water vapor in pascals).

We also included pixel elevation. For the flux tower marshes this was estimated from 1-m DEM, corrected via a Trimble R6 real time kinematic (RTK) global positioning system (GPS) receiver, with sub-meter vertical and horizontal accuracy (see Hladik *et al.*, 2013). This DEM was resampled to match the 30-m spatial resolution of Landsat-8. Elsewhere, we measured plot elevation in the field via RTK, which we used too created a pixel average elevation. Elevation served as a proxy for the environmental variables that vary along coastal marsh elevation gradients.

Finally, we included metrics related to tidal inundation intensity (based on local tidal flood heights and the frequency of flooding), as well as dryness intensity (based on a lack of tidal flushing, soil saturation, and frequency of dry observations). We used the random forest flood model described previously to estimate the percent of cloud-free Landsat pixels that were flooded or dry. Landsat-8 collects observations at solar noon every 16 d during cloud-free conditions. Thus, percent flooded or dry pixel observations do not capture all flooding, but, given a sufficiently long time series, should provide a proxy for the frequency and duration of flooding.

We combined these percent flooded and dry variables with the normalized local height below Mean Higher High Water (MHHW) and above Mean Lower Low Water (MLLW) for each pixel, related to the depth of tidal flooding during high tides and soil saturation during low tides, respectively. We normalized these values by dividing by the range between MHHW or MLLW and Mean Sea Level (MSL). These were monthly water level observations in the vertical datum NAVD88 from the nearest benchmarked NOAA sampling station, Fort Pulaski, GA (Station ID: 8670870). Final inundation and dryness metrics were:

Inundation intensity = (MHHW − elevation)/
        (MHHW − MSL) × percent flooded.
Dryness intensity = (elevation − MLLW)/
        (MSL − MLLW) × percent dry.

Estimating the day of spring green-up For each pixel, we estimated the day of green-up and the growing season day (number of days since green-up). This required that we first estimate soil temperature from cloud and tide-free Landsat-8 LST observations. For this, we estimated LST for each pixel from Landsat-8 B10 as described in Alber & O'Connell (2019). Then we filled LST gaps of > 1 month. For these gaps, we estimated LST as the average difference between mean air temperature (T_a, estimated from Daymet as (maximum T_a − minimum T_a)/2) and pixel LST for the given week in the year. We then interpolated LST for each pixel to create a daily LST estimate.

Once we had a daily LST estimate, we predicted pixel mean daily soil temperature. To predict soil temperature, we used the tidbit field soil temperature data as ground-truth information and adapted the method of Plauborg (2002), who estimated soil temperature from lags of previous air temperature combined with Fourier terms. We adapted this method to use LST and lags of

mean air temperature. We explored the inclusion of lags of 1 to 11 d because Zheng *et al.* (1993) noted that soil temperature was influenced by past temperatures of up to 11 d. We started by adding lags incrementally in a stepwise fashion and removed terms that were non-significant and did not reduce Akaike information criterion (AIC) by more than 2. We also included pixel elevation as a predictor, as we previously identified elevation as a significant factor influencing spatial variation in soil temperature (Alber & O'Connell, 2019; O'Connell *et al.*, 2020). We also included Fourier terms (sine and cosine) for estimating the influence of time of year on soil temperature. Models we tried took the form:

$$T_s = \alpha_0 + \alpha_1 \text{Elevation} + \alpha_2 \text{LST} + \alpha_3 T_{a(D-1)} + \ldots$$
$$+ \alpha_n T_{a(D-n)} + \beta_1 \sin\left(2\pi D^{(1/h)}\right)$$
$$+ \delta_1 \cos\left(2\pi D^{(1/h)}\right) + \beta_2 \sin\left(4\pi D^{(1/k)}\right)$$
$$+ \delta_1 \cos\left(4\pi D^{(1/k)}\right)$$

where T_s was the mean temperature at 10 cm soil depth, D was day of year, LST was the daily interpolated LST estimate for that pixel, T_a is the air temperature at lags of $D-1$ through $D-n$, and α_n, β_n, and δ_n were estimated constants. The Fourier terms had constants k and h, which we estimated by fitting all possible k and h from 1 to 60 and selecting as best those that minimized AIC. Values of zero for α_n, β_n, and δ_n allow the term to drop from the model.

Once we had soil temperature, we estimated the day of spring green-up for each pixel based on the total growing degree days experienced by roots by calculating total soil growing degree days (TDD):

$$\text{TDD} = \Sigma \text{Daily mean } T_s > 9.9°C$$

where 9.9°C is the physiological base temperature identified by O'Connell *et al.* (2020), and temperatures above the physiological base temperature contribute to phenological development. Plants are assumed to initiate spring green-up when TDD sums above the base temperature exceeds 202, the threshold identified by O'Connell *et al.* (2020). We validated this green-up model with PhenoCam-derived estimates of green-up date for marsh areas within the PhenoCam field of view (see O'Connell *et al.*, 2020). We also calculated growing season days (days since spring green-up date). We also included LST itself as a BGB predictor because it represents temperatures experienced by shoots, and aboveground physiology and growth should vary with temperature. Thus, we included temperatures experienced by both roots and shoots in the modeling process.

Further processing of aboveground predictors BGB at any time point might depend on antecedent conditions. To account for this, we considered combinations of lagged predictor variables. Before we could evaluate antecedent conditions, we first needed to transform the remote sensing predicted aboveground vegetation from an irregular time series with gaps from clouds and tides into a

regularly spaced time series. For this, we used linear approximation to create a single monthly estimate for each predicted variable that corresponded to the middle of each month. Then we created variables representing past vegetation dynamics. These variables included the rolling monthly means of all potential predictors for periods of 1–5 months, the change in each variable over one to five months, and changes in each variable from the start of the growing season. Up to five months was considered so that even end of season (spring, summer, fall, winter) growth included 2 months from the previous season, allowing previous conditions to influence the results. For nongrowing season months (Nov-green-up), we included peak growing season means from the previous growing season, as evidence from field data suggested past growing season data might determine over-wintering BGB. We could create these derived variables for all sites by using the aboveground models to predict observations where we lacked field data. Landsat-8 data were first available in 2013. Because some of the included variables required previous year estimates, we began our belowground model predictions in June 2014, providing an adequate time series of data to inform the model.

Final extreme gradient boosting (EXB) model to estimate belowground biomass (BGB) from aboveground predictors

To estimate monthly pixel-wise BGB, we used the EXB modeling workflow described in Section 'EXB modeling workflow' to create a final set of BGB EXB models, where field-estimated BGB was the response variable. To select model predictors, we began with those listed in Table 1, used the inner cross-validation to tune model hyperparameters as described in Section 'EXB modeling workflow', and then refit the inner training data with these hyperparameters. We then extracted the variable importance from the results, where variable importance was the gain or fractional contribution of each predictor to the model and ranges from 0 to 1 (low-high importance). We retained predictors with variable importance > 0.005 across at least three of the five inner training sets. We selected 0.005 as the cut-off because variables with less importance contribute little to the prediction. We then used only these selected parameters and re-tuned the hyperparameters before fitting the model to the outer training set. This resulted in five outer training models for BGB. We predicted BGB from each of these final five outer cross-validated models and used the mean as the final model prediction. For the final BGB model, we also calculated additional goodness-of-fit metrics obtained by regressing the observed values (obs$_i$) (response variable) against the predicted values (pre$_i$) (see Methods S5).

Example model output

BERM provides monthly predictions of BGB and the aboveground inputs that inform BGB at a 30 m × 30 m spatial grid. To illustrate BERM outputs, we applied it to a sample area (Fig. 2, methods and intermediate results in Methods S6). A landscape evaluation of spatiotemporal vegetation trends is outside the scope of this article, as our main goal was to describe the BERM algorithm.
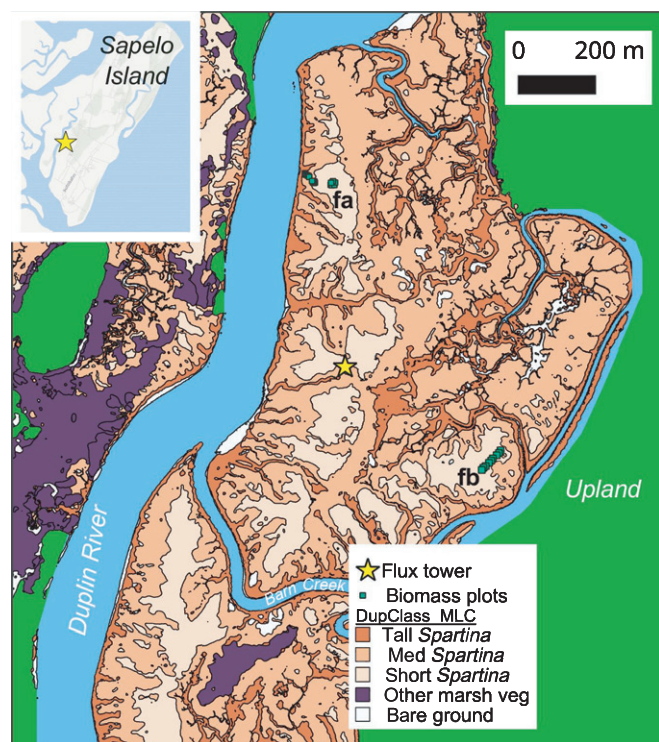
Fig. 2 An overview of the land cover classifications in the flux tower salt marsh area on Sapelo Island, where extensive *Spartina alterniflora* marshes are found (modified from Hladik *et al.*, 2013). Locations of coincident vegetation biomass plots are also shown (fa, Flux marsh A; fb, Flux marsh B). We applied the Belowground Ecosystem Resiliency Model (BERM) workflow to the areas of *Spartina alterniflora* cover in this region, in order to provide examples of BERM outputs.

Results

Comprehensive remotely derived dataset of potential aboveground proxies for predicting belowground biomass (BGB)

We used the EXB workflow to create remote sensing models of all of the aboveground vegetation characteristics considered for inclusion in BERM. When compared with testing data, models for estimating AGB, CHL, foliar N, and LAI, had nRMSEs of < 11%, and correlations ranged from 0.20 to 0.57 (Table S2; Figs S3–S6). We also estimated model environmental variables (see Table 1 for a list and Figs S7–S9), such as inundation intensity, which was generally highest in the fall and lowest in the spring (Fig. S8).

Green-up day of year was another input variable we considered, based on mean daily soil temperature, T_s, estimated as:

$$T_s = 4.74 - 2.22\,\text{Elevation} + 0.15\,\text{LST} - 0.71\,T_{a(D-1)}$$
$$+ 1.47\,T_{a(D-2)} - 0.14\cos\left(2\pi D^{(1/4)}\right) + 0.14\cos\left(4\pi D^{(1/7)}\right)$$

where R^2 for the T_s model was 0.99 and *P*-values for all parameters were < 0.004. The subsequent green-up estimate, derived from TDD, predicted green-up day of year with a RMSE of 9.3

d, whereas spatial variation in green-up date across the marsh can be > one month (O'Connell *et al.*, 2020) (Fig. S9).

Belowground Ecosystem Resiliency Model (BERM) for estimating belowground biomass (BGB)

The final five cross-validated outer EXB models for predicting BGB included only 33 out of the 131 potential predictors (Fig. 3). These five models were averaged together to create our final result and relied heavily on flooded observations, elevation, as well as longer-term averages of LST, AGB, and inundation. In general, longer-term metrics were more important than shorter-term ones. Vegetation-related metrics, such as AGB, LAI, foliar N, CHL, and phenology variables, 'green-up day of year' and 'days since green-up' were also included in the model. Of the physical variables, those related to vapor pressure, precipitation, maximum and minimum air temperature, LST, inundation, dryness, and flooded observations were included.

BERM predicted BGB across testing data with RMSE of 313 g m^{-2} (RMSE range across all cross-validations: 282–347 g m^{-2} Table 3; Fig. 4), whereas field data ranged from 191 to 2495 g m^{-2}. Maximum and minimum BGB and AGB were both highest during September and lowest during January/February (Figs S3, 5), which corresponded with the field data. Regression of observed vs predicted testing data from each cross-validation (Methods S5; Table 4) had explained variance (R^2) of 0.62 to 0.77. Further, in four out of the five cases, cross-validated results were unbiased (regression intercept β_0 equal to zero) (e.g. the regression slope β_1 did not differ from the 1 : 1 line). The decomposition of the sums of squares prediction error (SSPE) (Methods S5) revealed that > 89% of the prediction error was from unexplained variance rather than model bias.

Evaluation of model performance across sites and years

We also compared testing data model performance across sites and years (Table 3). Mean Pearson's *r* was 0.83, with the lowest *r* at flux tower marsh B and highest at flux tower mash A (Table 3). nRMSE averaged 11% across all the data, but was only < 25% at flux marsh A. MAEs from the five cross validation datasets spanned both negative and positive values across all sites except UGAMI marsh. Across years at flux tower marsh A, the only site with multiple years of data available, nRMSE was generally less than 25%, suggesting BERM could capture inter-annual cycles when appropriate training data were provided. However, for the year with the fewest testing data observations (2014), nRMSE averaged 39%, highlighting the importance of observation density for improving the model.

Discussion

BERM estimates near real time spatiotemporal variation in BGB of *Spartina alterniflora* salt marshes from remote sensing data. BGB predictions corresponded with field estimates (average *r* = 0.83; nRMSE = 11%). The model provided unbiased estimates across the range of observed BGB, suggesting it reliably tracks
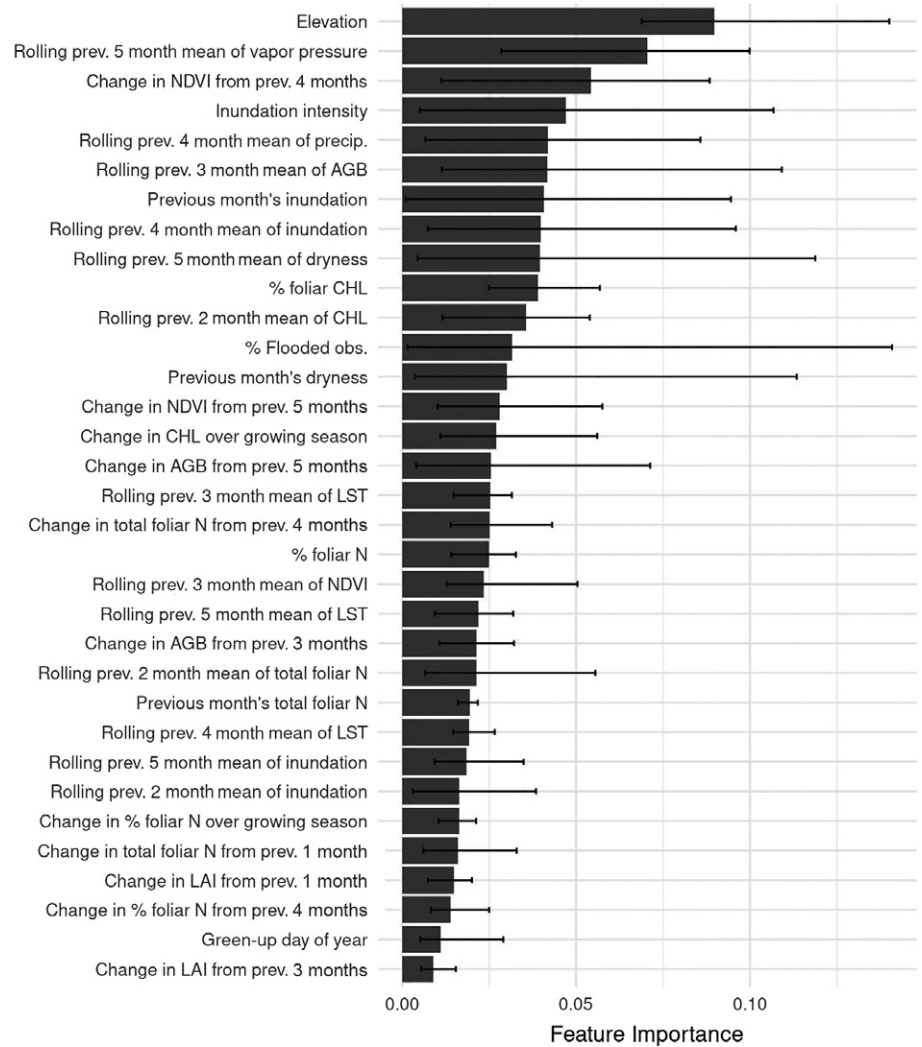
Fig. 3 Features selected for inclusion in the final Belowground Ecosystem Resiliency Model (BERM) output and their average importance across the cross-validated models, where error bars span the range of importance values across these models (minimum–maximum). In total 33 features were selected for inclusion in the model.

BGB patterns. Most of the model error was attributable to unexplained variance, and likely resulted from causes that are challenging to measure through remote sensing methods, such as genetic diversity, herbivory pressure, soil metabolites, or other factors. However, BERM relies on a broad suite of biophysical predictors, and can track belowground plant growth related to change in aboveground vegetation and environmental gradients. The accuracy is sufficient to evaluate spatial patterns, identify areas experiencing either very high or low biomass production, and track relative patterns and trajectories. BERM can be directly applied to marshes where the range of environmental variables is similar to those here (Table 2), and can be extended to other systems where training data are available. Because BERM provides site-wide estimates, this is a promising approach for estimating BGB patterns and trends at broad-scales.

Some of the most important BERM variables were related to wetland hydrology, such as inundation intensity and elevation (Fig. 3). Inundation intensity was a combination of inundation frequency/duration and water depth. Thus, lower elevation marsh areas with higher water depths and poorly drained areas with longer flood duration had increased inundation intensity.

This hydrology relationship was not surprising, as both flooding duration and depth drive plant productivity (Kirwan & Megonigal, 2013; Voss *et al.*, 2013), through relationships with tidal flushing of soil metabolites, nutrient delivery, soil salinity, and soil anoxia. Tidal flooding brings in oxygenated water and nutrients, as well as flushes soil salinity, which can have a positive impact on plant growth following tidal draining. However, tidal flooding also slows plant gas exchange (Kathilankal *et al.*, 2008), and deep and prolonged flooding causes respiratory deficits (Mendelssohn *et al.*, 1981).

Variables related to ε (production efficiency) were also important in BERM, including temperature, precipitation, and vapor pressure. BERM included LST, a proxy for plant canopy temperature. Early in the year, increasing LST likely increases plant metabolism and production, whereas later in the summer, high canopy temperatures during drought can cause higher soil salinity, water limitation, and plant stress (Schalles *et al.*, 2013; Miklesh & Meile, 2018). For example, Schalles *et al.* (2013) suggested drought on Sapelo Island negatively impacted *Spartina alterniflora* AGB by increasing soil salinity and evapotranspiration, resulting in long-term AGB declines of as much as 39%

Table 3 Goodness-of-fit metrics for the final *Spartina alterniflora* belowground biomass output, including mean absolute error (MAE) (g m^{-2}), root mean squared error (RMSE) (g m^{-2}), normalized RMSE (nRMSE)(%), coefficient of variation of RMSE (COV RMSE) (%), and correlation (*r*) for all the testing data combined ('all'), for testing data from each site ('fluxa', 'fluxb', 'skida', and 'ugami') (for site comparisons, data are from 2016), and for each year (2014–2019, where 2014 predictions were only available after June)

Type	MAE	RMSE	nRMSE	COV RMSE	Correlation	n-train	n-test
All	8 (−91–78)	312 (282–347)	11 (10–12)	28 (25–31)	0.83 (0.79–0.88)	223 (219–234)	116 (105–120)
Site							
fluxa	8 (−80–92)	383 (180–622)	11 (10–12)	27 (25–29)	0.82 (0.80–0.87)	191 (189–199)	101 (93–103)
fluxb	60 (−77–153)	225 (122–265)	26 (14–30)	35 (19–41)	0.56 (0.03–0.99)	10 (9–12)	5 (3–6)
skida	163 (−2–313)	307 (206–382)	44 (29–55)	67 (45–83)	0.67 (0.45–0.98)	10 (9–12)	5 (3–6)
ugami	−175 (−388–−58)	383 (180–622)	27 (13–44)	25 (12–41)	0.72 (0.51–0.95)	11 (11–12)	6 (5–6)
Year							
2014	138 (113–164)	366 (223–570)	39 (31–47)	29 (23–35)	0.47 (0.01–0.92)	3 (0–5)	2 (0–5)
2015	−41 (−215–141)	316 (246–374)	24 (19–29)	27 (21–32)	0.68 (0.53–0.88)	36 (20–55)	24 (5–40)
2016	90 (33–144)	339 (266–409)	16 (13–20)	31 (24–37)	0.74 (0.61–0.85)	39 (35–40)	21 (20–25)
2017	82 (−28–221)	247 (188–356)	15 (11–21)	24 (18–34)	0.85 (0.79–0.96)	45 (40–50)	21 (16–26)
2018	−70 (−145–4)	278 (217–344)	10 (8–13)	20 (16–25)	0.95 (0.92–0.97)	48 (36–54)	24 (18–36)
2019	150 (−138–382)	366 (223–570)	21 (13–33)	39 (24–61)	0.45 (−1.00–1.00)	21 (16–27)	8 (2–13)

Sample size for training and testing data are provided in the columns n-train and n-test, respectively. Numbers in parenthesis indicate the range observed (minium–maximum) in the testing data across all the cross-validated models.
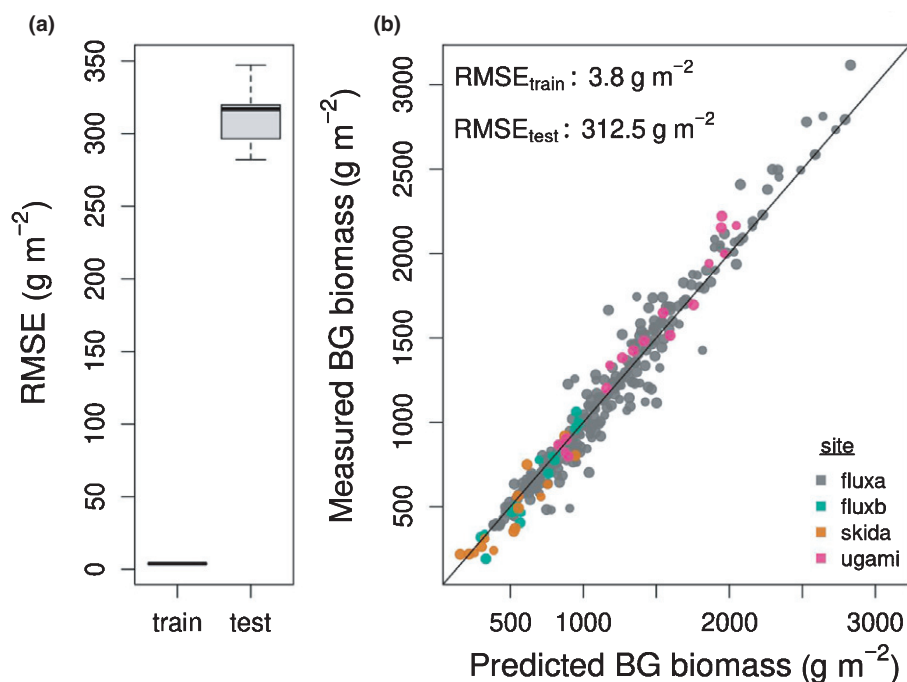

Fig. 4 Range in root mean squared error (RMSE) of fit between field measured and model predicted *Spartina alterniflora* belowground biomass (g m^{-2}) for the training and testing sets across all five cross-validations (a); Average measured vs predicted belowground biomass across all five cross-validations for the training and testing sets combined (b).

over a 28-yr period. Such droughts can be exacerbated by high temperatures and porewater salinity. In Sapelo *Spartina alterniflora* marshes, the latter varies from 20 to 60 pss-78, depending on year, season, and marsh elevation, as a consequence of precipitation, temperature, evapotranspiration, tidal flooding frequency, and river discharge (Miklesh & Meile, 2018). LST is also related to soil temperature, which varies spatially and likely similarly influences root metabolism (Alber & O'Connell, 2019). Vapor pressure, a variable used to calculate vapor pressure deficit and evapotranspiration, also is related to plant water stress (Massmann *et al.*, 2019). Thus, it was not surprising that LST, precipitation, and vapor pressure were important model predictors. Altogether, these environmental variables influence rates of

photosynthesis and respiration, and were likely important for constraining model productivity estimates.

BERM also relied on aboveground variables that drive plant productivity. For example, NDVI, a proxy for fAPAR, and canopy CHL were both important and have a close relationship with GPP/NPP (Monteith, 1977; Ruimy *et al.*, 1994; Gitelson *et al.*, 2006). CHL and fAPAR represent the amount of CHL available to conduct photosynthesis and the light available to simulate photosynthetic activity, respectively. Additionally, root : shoot ratios vary predictably with plant nutrient status (Morris, 1982; Morris *et al.*, 2013; O'Connell *et al.*, 2014), and foliar N provides a proxy for nutrient status. Thus, the combination of foliar N and AGB can help estimate BGB (O'Connell *et al.*,
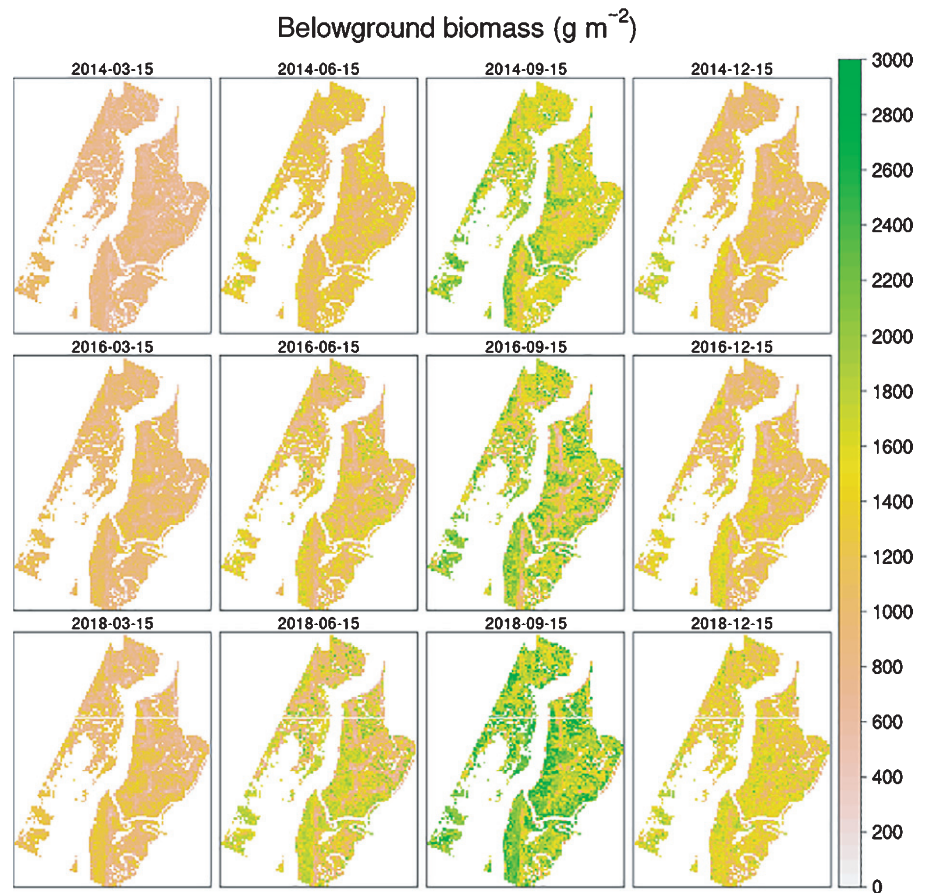
Fig. 5 An example of a Belowground Ecosystem Resiliency Model (BERM) predicted time series for *Spartina alterniflora* belowground biomass (g m⁻²) for the flux tower marsh area. See Fig. 2: for land cover classifications for this area.

Table 4 Additional goodness-of-fit metrics for testing data from each cross-validated (CV) model for the final *Spartina alterniflora* belowground biomass output.

CV	R^2	β_0	β_1	SSPE	U_{bias}	$U_{\beta1=1}$	U_{error}
1	0.69	70.2 ± 71.7	0.92 ± 0.06	10 288 907	0.01	0.02	0.97
2	0.66	45.9 ± 75.4	0.97 ± 0.06	12 044 544	0.00	0.00	1.00
3	0.77	−75.4 ± 65.9	1.03 ± 0.05	9 389 720	0.02	0.00	0.98
4	0.67	93.8 ± 75.3	0.85 ± 0.06**	10 737 935	0.06	0.05	0.89
5	0.62	185.3 ± 73.4*	0.91 ± 0.07	14 469 802	0.07	0.02	0.92

These metrics are derived from the regression of the observed values (response variable) against the Belowground Ecosystem Resiliency Model (BERM) predicted values. The metrics include explained variance (R^2) for the regression, the regression intercept (β_0) and slope (β_1) (mean ± SE), where intercepts that significantly differ from zero and slopes that significantly differ from one are indicated. Intercepts differing from zero reflect a consistent model bias across all observed values, whereas slopes differing from one reflect model inconsistency from the 1 : 1 line over the range of observed values. We also present the sums of squares prediction error (SSPE), which we further decomposed through Theil's partial inequality coefficients into different sources of error, including the proportion of error resulting from differences between observed vs predicted values (U_{bias}), the proportion of error associated with deviations in slope from the 1 : 1 line ($U_{\beta1=1}$), and the proportion of error from unexplained model variance (U_{error}). Units are in g m⁻² for the slope, intercept, and SSPE.
*, $P < 0.05$ for H_0: $\beta_0 = 0$; **, $P < 0.05$ for H_0: $\beta_1 = 1$.

2015). We additionally included the day of spring green-up for each pixel, allowing us to account for spatiotemporal heterogeneity in growing season timing (O'Connell *et al.*, 2020).

Evaluation of model performance across sites and years

While BERM performed well overall, there was variability in model performance across sites and years. Performance was best in Flux tower marsh A in terms of MAE, nRMSE, and correlation, likely because multiple years were available for training. Similarly, goodness-of-fit metrics were better for years where we had more testing data for model evaluation. For this site, as well as for other sites where nRMSE averaged < 28% (Flux tower marsh B, UGAMI marsh), we are confident BERM can produce quantitative estimates of BGB and evaluate spatiotemporal patterns, though the model could be improved by additional ground-truth data.

BERM performed poorest at Skidaway, the only site not on Sapelo Island. Both biotic and abiotic conditions may differ there, for example, in terms of flooding and weather dynamics, or *Spartina alterniflora* population structure. It would be useful to have multi-year data to allow the model to learn from these differences. When adequate ground-truth data are lacking, we caution against using BERM for quantitative BGB estimates, particularly for sites where conditions are very different from those modeled here. However, BERM should still be useful for evaluating spatiotemporal patterns.

Future applications

BERM-estimated BGB has multiple applications within coastal marshes. BERM can map plant productivity trends across broad spatial scales and over time, and could form the basis of long-term monitoring. Areas with low BGB are potentially at risk from sea level rise, as BGB contributes to vertical marsh accretion (Nyman *et al.*, 1993; DeLaune & Pezeshki, 2003; Mudd *et al.*, 2009) and might indicate priority sites for marsh restoration, as BGB may be a more sensitive indicator of ecosystem function than AGB (Stagg *et al.*, 2017). BGB is also important for evaluating wetland carbon sequestration (Morris *et al.*, 2002; DeLaune & Pezeshki, 2003; Mudd *et al.*, 2009). BERM can also be linked with coastal marsh blue carbon models. In the future, BERM could be updated to use other remote sensing data, such as Sentinel-2, or future satellite platforms such as Landsat-9. Further, while we trained our model for use in *Spartina alterniflora* marshes, a similar process should be possible for other plant species. Overall, BERM estimates help map marsh function and provide a critical tool for decision-making in coastal areas. This is the first model of this kind, and can be used for long-term and broad-scale monitoring of ecosystem trajectories.

Author contributions

JLO, DMR, MA and KBB planned and designed the research. JLO collected the field data, conducted the analysis, and wrote the manuscript. DMR, MA and KBB contributed to writing the manuscript and interpreting the results.

ORCID

Merryl Alber https://orcid.org/0000-0002-9467-4449
Kristin B. Byrd https://orcid.org/0000-0002-5725-7486
Deepak R. Mishra https://orcid.org/0000-0001-8192-7681
Jessica L. O'Connell https://orcid.org/0000-0002-2814-0668

Data availability

Model code to apply BERM estimates at broad-scales as described in this article are available at: https://zenodo.org/badge/latestdoi/382914221

The most recent version of BERM is available at: https://github.com/jloconnell/BERM

The data that support the findings of this study are openly available in the GCE-LTER Data Catalog:

Repeat measurements of Flux tower marsh A vegetation data from 2013 to 2019 at: https://doi.org/10.6073/pasta/ed6fb03c81a39babfdbb7be3c076cb65

Repeat measurements of vegetation plots for the 2016 growing season from all sites at: https://doi.org/10.6073/pasta/03f4f78c6498aecca34faf4339591129

Extra LAI measurements from other years, as published in Hawman *et al.* (2021): https://doi.org/10.6073/pasta/53a3e61158aa058aee50be72a1156c83

Soil temperature measurements, as published in Alber & O'Connell (2019): https://doi.org/10.6073/pasta/db52896274df333805f2043efb7e5710

Corrected DEM data for GCE-LTER, as published in Hladik *et al.* (2013): https://doi.org/10.6073/pasta/4c5187ef603f70cd0a77ece24ef0fed9

Vegetation classification map for GCE-LTER, as published in Hladik *et al.* (2013):
https://doi.org/10.6073/pasta/575c75149bdaa30a68507e46607ce784

References

Alber M, O'Connell JL. 2019. Elevation drives gradients in surface soil temperature within salt marshes. *Geophysical Research Letters* **46**: 5313–5322.

Barr JG, Engel V, Fuentes JD, Fuller DO, Kwon H. 2013. Modeling light use efficiency in a subtropical mangrove forest equipped with CO₂ eddy covariance. *Biogeosciences* **10**: 2145–2158.

Bischl B, Mersmann O, Trautmann H, Weihs C. 2012. Resampling methods for meta-model validation with recommendations for evolutionary computation. *Evolutionary Computation* **20**: 249–275.

Byrd KB, O'Connell JL, Di Tommaso S, Kelly M. 2014. Evaluation of sensor types and environmental controls on mapping biomass of coastal marsh emergent vegetation. *Remote Sensing of Environment* **149**: 166–180.

Cahoon DR, McKee KL, Morris JT. 2021. How plants influence resilience of salt marsh and mangrove wetlands to sea-level rise. *Estuaries and Coasts* **44**: 883–898.

Chen T, Guestrin C. 2016. XGBoost: A Scalable Tree Boosting System. In: *KDD '16: proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.* New York, NY, USA: KDD Knowledge Discovery and Data Mining, 785–794.

Chen T, He T, Benesty M, Tang Y. 2018. *Understand your dataset with Xgboost.* R package: Xgboost, vignette. [WWW document] URL https://cran.r-project.org/web/packages/xgboost/vignettes/discoverYourData.html#numeric-v.s.-categorical-variables [accessed 24 May 2021].

Craft CB. 2001. Soil organic carbon, nitrogen, and phosphorus as indicators of recovery in restored *Spartina* marshes. *Ecological Restoration* 19: 87–91.

Dahl TE. 1990. *Wetlands losses in the United States 1780's to 1980's.* Washington, DC, USA: US Department of the Interior, US Fish and Wildlife Service.

Deegan LA, Johnson DS, Warren RS, Peterson BJ, Fleeger JW, Fagherazzi S, Wollheim WM. 2012. Coastal eutrophication as a driver of salt marsh loss. *Nature* 490: 388–392.

DeLaune RD, Pezeshki SR. 2003. The role of soil organic carbon in maintaining surface elevation in rapidly subsiding U.S. Gulf of Mexico coastal marshes. *Water, Air and Soil Pollution: Focus* 3: 167–179.

Dewar RC. 1993. A root-shoot partitioning model based on carbon–nitrogen–water interactions and Munch phloem flow. *Functional Ecology* 7: 356–368.

Elith J, Leathwick JR, Hastie T. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology* 77: 802–813.

Foga S, Scaramuzza PL, Guo S, Zhu Z, Dilley RD, Beckmann T, Schmidt GL, Dwyer JL, Joseph Hughes M, Laue B. 2017. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sensing of Environment* 194: 379–390.

Friedman JH. 2001. Greedy function approximation: a gradient boosting machine. *Annals of Statistics* 29: 1189–1232.

Friedman J, Hastie T, Tibshirani R. 2000. Special invited paper. Additive logistic regression: a statistical view of boosting. *Annals of Statistics* 28: 337–374.

Friedman JH, Meulman JJ. 2003. Multiple additive regression trees with application in epidemiology. *Statistics in Medicine* 22: 1365–1381.

Gallagher JL. 1983. Seasonal patterns in recoverable underground reserves in *Spartina alterniflora* Loisel. *American Journal of Botany* 70: 212–215.

Gallagher JL, Wolf PL, Pfeiffer WJ. 1984. Rhizome and root growth rates and cycles in protein and carbohydrate concentrations in Georgia *Spartina alterniflora* Loisel. Plants. *American Journal of Botany* 71: 165–169.

Gedan KB, Silliman BR. 2009. Patterns of salt marsh loss within coastal regions of North America. In: Silliman BR, Grosholz ED, Bertness MD, eds. *Human impacts on salt marshes: a global perspective*. Berkeley, CA, USA: University of California Press, 253–265.

Ghosh S, Mishra DR, Gitelson AA. 2016. Long-term monitoring of biophysical characteristics of tidal wetlands in the northern Gulf of Mexico – a methodological approach using MODIS. *Remote Sensing of Environment* 173: 39–58.

Gitelson AA, Peng Y, Arkebauer TJ, Schepers J. 2014. Relationships between gross primary production, green LAI, and canopy chlorophyll content in maize: implications for remote sensing of primary production. *Remote Sensing of Environment* 144: 65–72.

Gitelson AA, Stark R, Grits U, Rundquist D, Kaufman Y, Derry D. 2002. Vegetation and soil lines in visible spectral space: a concept and technique for remote estimation of vegetation fraction. *International Journal of Remote Sensing* 23: 2537–2562.

Gitelson AA, Viña A, Verma SB, Rundquist DC, Arkebauer TJ, Keydan G, Leavitt B, Ciganda V, Burba GG, Suyker AE. 2006. Relationship between gross primary production and chlorophyll content in crops: implications for the synoptic monitoring of vegetation productivity. *Journal of Geophysical Research: Atmospheres* 111: D08S11.

Hawman PA, Mishra DR, O'Connell JL, Cotten DL, Narron CR, Mao L. 2021. Salt marsh light use efficiency is driven by environmental gradients and species-specific physiology and morphology. *Journal of Geophysical Research: Biogeosciences* 126: e2020JG006213.

Hladik C, Schalles J, Alber M. 2013. Salt marsh elevation and habitat mapping using hyperspectral and LIDAR data. *Remote Sensing of Environment* 139: 318–330.

Hyndman RJ, Koehler AB. 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting* 22: 679–688.

Kartesz JT. 2015. *The Biota of North America Program (BONAP)*. Chapel Hill, NC, USA: North American Plant Atlas. [WWW document] URL http://bonap.net/napa.

Kathilankal JC, Mozdzer TJ, Fuentes JD, D'Odorico P, McGlathery KJ, Zieman JC. 2008. Tidal influences on carbon assimilation by a salt marsh. *Environmental Research Letters* 3: 44010.

Kirwan ML, Guntenspergen GR. 2012. Feedbacks between inundation, root production, and shoot growth in a rapidly submerging brackish marsh. *Journal of Ecology* 100: 764–770.

Kirwan ML, Megonigal JP. 2013. Tidal wetland stability in the face of human impacts and sea-level rise. *Nature* 504: 53–60.

Kokaly RF, Asner GP, Ollinger SV, Martin ME, Wessman CA. 2009. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment* 113 (Suppl. 1): S78–S91.

Kumar M, Monteith J. 1981. Plants and the daylight spectrum. In *Remote sensing of crop growth*. New York, NY, USA: Academic Press, 133–144.

Langston AK, Alexander CR, Alber M, Kirwan ML. 2021. Beyond 2100: elevation capital disguises salt marsh vulnerability to sea-level rise in Georgia, USA. *Estuarine, Coastal and Shelf Science* 249: 107093.

Levin SA, Mooney HA, Field C. 1989. The dependence of plant root:shoot ratios on internal nitrogen concentration. *Annals of Botany* 64: 71–75.

Lotze HK, Lenihan HS, Bourque BJ, Bradbury RH, Cooke RG, Kay MC, Kidwell SM, Kirby MX, Peterson CH, Jackson JBC. 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science* 312: 1806–1809.

Massmann A, Gentine P, Lin C. 2019. When does vapor pressure deficit drive or reduce evapotranspiration? *Journal of Advances in Modeling Earth Systems* 11: 3305–3320.

McConnaughay KDM, Coleman JS. 1999. Biomass allocation in plants: ontogeny or optimality? A test along three resource gradients. *Ecology* 80: 2581–2593.

Mcleod E, Chmura GL, Bouillon S, Salm R, Björk M, Duarte CM, Lovelock CE, Schlesinger WH, Silliman BR. 2011. A blueprint for blue carbon: toward an improved understanding of the role of vegetated coastal habitats in sequestering CO_2. *Frontiers in Ecology and the Environment* 9: 552–560.

Mendelssohn IA, McKee KL, Patrick WH. 1981. Oxygen deficiency in *Spartina alterniflora* roots: metabolic adaptation to anoxia. *Science* 214: 439–441.

Mendelssohn IA, Morris JT. 2002. Eco-physiological controls on the productivity of Spartina alterniflora Loisel. In: Weinstein MP, Kreeger DA, eds. *Concepts and controversies in tidal marsh ecology*. Fort Hancock, NJ, USA: Springer Netherlands, 59–80.

Miklesh D, Meile C. 2018. Porewater salinity in a southeastern United States salt marsh: controls and interannual variation. *PeerJ* 6: e5911.

Mishra DR, Cho HJ, Ghosh S, Fox A, Downs C, Merani PBT, Kirui P, Jackson N, Mishra S. 2012. Post-spill state of the marsh: remote estimation of the ecological impact of the Gulf of Mexico oil spill on Louisiana Salt Marshes. *Remote Sensing of Environment* 118: 176–185.

Monteith JL. 1972. Solar radiation and productivity in tropical ecosystems. *The Journal of Applied Ecology* 9: 747.

Monteith JL. 1977. Climate and efficiency of crop production in Britain. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences* 281: 277–294.

Morris JT. 1982. A model of growth responses by *Spartina alterniflora* to nitrogen limitation. *Journal of Ecology* 70: 25–42.

Morris JT, Shaffer GP, Nyman JA. 2013. Brinson review: perspectives on the Influence of nutrients on the sustainability of coastal wetlands. *Wetlands* 33: 975–988.

Morris JT, Sundareshwar PV, Nietch CT, Kjerfve B, Cahoon DR. 2002. Responses of coastal wetlands to rising sea level. *Ecology* 83: 2869–2877.

Mudd SM, Howell SM, Morris JT. 2009. Impact of dynamic feedbacks between sedimentation, sea-level rise, and biomass production on near-surface marsh stratigraphy and carbon accumulation. *Estuarine, Coastal and Shelf Science* 82: 377–389.

Nyman JA, Delaune RD, Roberts HH, Patrick WH. 1993. Relationship between vegetation and soil formation in a rapidly submerging coastal marsh. *Marine Ecology Progress Series* 96: 269–279.

Nyman JA, Walters RJ, Delaune RD, Patrick JWH. 2006. Marsh vertical accretion via vegetative growth. *Estuarine, Coastal and Shelf Science* 69: 370–380.

O'Connell JL, Alber M. 2016. A smart classifier for extracting environmental data from digital image time-series: applications for PhenoCam data in a tidal salt marsh. *Environmental Modelling & Software* 84: 134–139.

O'Connell JL, Alber M, Pennings SC. 2020. Microspatial differences in soil temperature cause phenology change on par with long-term climate warming in salt marshes. *Ecosystems* 23: 498–510.

O'Connell JL, Byrd KB, Kelly M. 2014. Remotely-sensed indicators of N-related biomass allocation in *Schoenoplectus acutus*. *PLoS ONE* 9: e90870.

O'Connell JL, Byrd KB, Kelly M. 2015. A hybrid model for mapping relative differences in belowground biomass and root: shoot ratios using spectral reflectance, foliar N and plant biophysical data within coastal marsh. *Remote Sensing* 7: 16480–16503.

O'Connell JL, Mishra DR, Cotten DL, Wang L, Alber M. 2017. The Tidal Marsh Inundation Index (TMII): an inundation filter to flag flooded pixels and improve MODIS tidal marsh vegetation time-series analysis. *Remote Sensing of Environment* 201: 34–46.

Pennings SC, Bertness MD. 2001. Salt marsh communities. In: Hay ME, Gaines SD, Bertness MD, eds. *Marine community ecology*. Sunderland, MA, USA: Sinauer Associates, 550.

Peterson PM, Romaschenko K, Arrieta YH, Saarela JM. 2014a. A molecular phylogeny and new subgeneric classification of *Sporobolus* (Poaceae: Chloridoideae: Sporobolinae). *Taxon* 63: 1212–1243.

Peterson PM, Romaschenko K, Arrieta YH, Saarela JM. 2014b. Proposal to conserve the name *Sporobolus* against *Spartina*, *Crypsis*, *Ponceletia*, and *Heleochloa* (Poaceae: Chloridoideae: Sporobolinae). *Taxon* 63: 1373–1374.

Plauborg F. 2002. Simple model for 10 cm soil temperature in different soils with short grass. *European Journal of Agronomy* 17: 173–179.

Rouse J, Haas RH, Schell J, Deering D. 1974. Monitoring vegetation systems in the Great Plains with ERTS. In *Third ERTS Symposium, NASA SP-351*. Washington, DC, USA: NASA, 309–317.

Ruimy A, Saugier B, Dedieu G. 1994. Methodology for the estimation of terrestrial net primary production from remotely sensed data. *Journal of Geophysical Research: Atmospheres* 99: 5263–5283.

Running SW, Nemani RR, Heinsch FA, Zhao M, Reeves M, Hashimoto H. 2004. A continuous satellite-derived measure of global terrestrial primary production. *BioScience* 54: 547–560.

Schalles J, Hladik C, Lynes A, Pennings S. 2013. Landscape estimates of habitat types, plant biomass, and invertebrate densities in a Georgia salt marsh. *Oceanography* 26: 88–97.

Schratz P, Muenchow J, Iturritxa E, Richter J, Brenning A. 2019. Performance evaluation and hyperparameter tuning of statistical and machine-learning models using spatial data. *Ecological Modelling* 406: 109–120.

Stagg CL, Schoolmaster DR Jr, Piazza SC, Snedden G, Steyer GD, Fischenich CJ, McComas RW. 2017. A landscape-scale assessment of above- and belowground primary production in coastal wetlands: implications for climate change-induced community shifts. *Estuaries and Coasts* 40: 856879.

Strong DR, Ayres DA. 2016. Control and consequences of *Spartina* spp. Invasions with focus upon San Francisco Bay. *Biological Invasions* 18: 2237–2246.

Tao J, Mishra DR, Cotten DL, O'Connell J, Leclerc M, Nahrawi HB, Zhang G, Pahari R. 2018. A comparison between the MODIS product (MOD17A2) and a tide-robust empirical GPP model evaluated in a Georgia wetland. *Remote Sensing* 10: 1831.

Thornton M, Thornton P, Wei Y, Vose R, Boyer A. 2017. *Daymet: Station-level inputs and model predicted values for North America*, Version 3. Oak Ridge, TN, USA: ORNL DAAC.

Turner DP, Ritts WD, Cohen WB, Gower ST, Running SW, Zhao M, Costa MH, Kirschbaum AA, Ham JM, Saleska SR *et al.* 2006. Evaluation of MODIS NPP and GPP products across multiple biomes. *Remote Sensing of Environment* 102: 282–292.

US Geological Survey. 2019. *The National Map New data delivery homepage, advanced viewer, lidar visualization*. Reston, VA, USA; Report no. 2019-3032. US Geological Survey [WWW document] URL http://pubs.er.usgs.gov/publication/fs20193032 [accessed 19 October 2020].

USDA, NRCS. 2019. *The PLANTS database*. Greensboro, NC, USA: National Plant Data Team. [WWW document] URL http://plants.usda.gov [accessed 3 May 2021].

Voss CM, Christian RR, Morris JT. 2013. Marsh macrophyte responses to inundation anticipate impacts of sea-level rise and indicate ongoing drowning of North Carolina marshes. *Marine Biology* 160: 181–194.

Willmott C, Matsuura K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research* 30: 79–82.

Xiao J, Zhuang Q, Law BE, Baldocchi DD, Chen J, Richardson AD, Melillo JM, Davis KJ, Hollinger DY, Wharton S *et al.* 2011. Assessing net ecosystem carbon exchange of U.S. Terrestrial ecosystems by integrating eddy covariance flux measurements and satellite observations. *Agricultural and Forest Meteorology* 151: 60–69.

Zheng D, Hunt ER Jr, Running SW. 1993. A daily soil temperature model based on air temperature and precipitation for continental applications. *Climate Research* 2: 183–191.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Leaf CHL concentration vs SPAD measurements.

Fig. S2 Theoretical diagram of the spatial cross-validation workflow.

Fig. S3 BERM predicted time series for aboveground biomass.

Fig. S4 BERM predicted time series for LAI.

Fig. S5 BERM predicted time series for percent foliar N.

Fig. S6 BERM predicted time series for CHL.

Fig. S7 BERM predicted time series for LST.

Fig. S8 BERM predicted time series for inundation intensity.

Fig. S9 BERM predicted time series for green-up.

Methods S1 Chlorophyll extractions and chlorophyll-SPAD model calibration.

Methods S2 Foliar N sample collection, preparation and analysis.

Methods S3 Field sample frequency and numbers.

Methods S4 Extreme gradient boosting modeling workflow explanation.

Methods S5 Goodness-of-fit calculations for the final BGB models.

Methods S6 Example BERM output.

Table S1 Field campaign sampling frequency and dates.

Table S2 Testing data mean goodness-of-fit metrics.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.