Locality of Topographic Ground Truth Data for Salt Marsh Lidar DEM Elevation Bias Mitigation

Stephen C. Medeiros, *Member, IEEE*, Joelle S. Bobinsky, and Khalid Abdelwahab

***Abstract*—Lidar-derived digital elevation models (DEMs) are crucial for modeling salt marsh evolution, forecasting inundation depth, frequency, and duration, and simulating sea level rise. Advances in lidar acquisition and data processing techniques have led to increased accuracy; however, in densely vegetated coastal salt marsh areas, lidar-derived DEMs are generally unreliable without adjustment. In this article, we investigate the need for local topographic ground truth data to train random forest (RF) DEM adjustment models for two similar northern Gulf of Mexico salt marshes. Two RTK-GNSS field surveys were conducted to acquire ground truth topographic elevations near St. Marks, Florida, USA ($n = 377$) and Pascagoula, MS, USA ($n = 610$). These elevations, along with Sentinel-2A MSI reflectance values and lidar DEM elevations, were used to validate and train local and combined RF salt marsh DEM adjustment models. The local RF models achieved mean absolute error values of 0.054 m and 0.045 m in the leave-one-out cross-validation for St. Marks and Pascagoula, respectively. Elevation bias predictions using remote RF models were far worse and those using the combined RF model were marginally worse. Using the local RF predictions to mitigate the bias in the lidar DEMs improved their accuracy by 69.1% for St. Marks and 90.9% for Pascagoula. The DEM elevation was identified as the most important predictor. This evidence suggests that local ground truth data are necessary for mitigating bias in salt marsh lidar DEMs although it remains to be seen if increasing the data set size and incorporating additional hydrologic predictor variables could narrow the accuracy gap.**

***Index Terms*—Coastal, lidar, multispectral, remote sensing, Sentinel.**

I. Introduction

LIDAR digital elevation models (DEMs) are known to be inaccurate in coastal salt marshes because the laser pulses cannot always reflect off of the true marsh platform surface. The presence of standing water and dense vegetation (typically tall, stiff grasses with heights greater than 1 m) is the primary cause of the widespread biases in topographic elevation data products such as point clouds and bare earth DEMs [1]–[3].

The ramifications of this persistent bias are acutely evident when lidar DEMs are used as input data for marsh evolution models in microbial environments. Models such as the marsh equilibrium model (MEM) [4] and HydroMEM [5] rely on the marsh platform topography to serve as a starting point for future projections of marsh migration and evolution [6], [7]. When the initial state of the marsh platform is erroneously biased to a higher elevation than the upper part of the tidal inundation frame (mean high water or MHW), the modeled marsh is not accurately inundated in the simulations and subsequently does not receive a realistic sediment loading. This was the case in the northern Gulf of Mexico (NGOM) when lidar data from 2007 to 2008 were used to model a marsh with a tidal range of approximately 25 cm [7]. Unless this bias is addressed, marsh biomass density, zonation, and migration projections will be inaccurate from the start and all emergent effects from future conditions such as sea level rise (SLR) will be unreliable.

Accurate representation of coastal terrain is also an important factor in hurricane storm surge modeling, regardless of the local tide range [8]–[10]. Coastal salt marshes are often the first or second physical buffer zone standing between an incoming surge and upland property. Therefore, it exerts substantial influence over the propagation (depth, extent, and timing) of storm surge inundation. Since this is a major cause of the destruction and economic disruption associated with tropical cyclones [11], accurate coastal flood predictions for both immediate event-scale decisions such as evacuation orders as well as longer-term risk assessments and resilient infrastructure planning rely on coastal DEMs as their primary input data source.

To mitigate this persistent elevation bias, corrections must be applied to the marsh surface topography. However, it is impractical to adjust multicounty or regional scale DEMs based on field data alone. To address this, techniques have been developed that rely on vegetation characteristics, such as height [2], remotely sensed biomass density [12], local tidal frame elevations [13], and lidar waveform data [14]. In addition to the high elevation bias, lidar DEMs also tend to flatten out the underlying microtopography including small tidal creeks, making the spatial distribution of the bias magnitudes nonlinear [12]. Approaches such as random forest (RF) [7], [15], multiple regression [12], and gradient boosted nonparametric regression [14] have been shown to be effective in the past and the research is trending toward simpler models that require fewer field measured vegetation characteristics for model training. When comparing different machine learning techniques including RF, support vector machine, k-nearest neighbor, and artificial neural network, RF models have emerged as the most useful due to their computational efficiency, resistance to overfitting, ability to handle small datasets, and

Manuscript received 13 May 2022; revised 17 June 2022; accepted 27 June 2022. Date of publication 8 July 2022; date of current version 27 July 2022. This work was supported by the U.S. National Oceanic and Atmospheric Administration's National Centers for Coastal Ocean Science under award NA16NOS4780208. *(Corresponding author: Stephen C. Medeiros.)*

Stephen C. Medeiros and Joelle S. Bobinsky are with the Embry-Riddle Aeronautical University, Daytona Beach, FL 32114 USA (e-mail: medeiros@erau.edu; bobinsj1@my.erau.edu).

Khalid Abdelwahab is with the University of Central Florida, Orlando, FL 32816 USA (e-mail: khalidwaw@gmail.com).

Digital Object Identifier 10.1109/JSTARS.2022.3189226

explainability (trained RF models produce feature importance metrics using out-of-bag testing) [15]. Additionally, Belgiu and Drăguţ discussed how RF classifiers are particularly beneficial for remote sensing purposes, including studies with multisource data [16]. Hu *et al.* used an RF regression model with data from remote sensing and field measurements to estimate AGBM to produce a global mangrove forest AGBM map [17]. This work supported the findings of Fassnacht *et al.* who concluded that combining lidar data with many reference sample units and an RF model produces biomass predictions with the lowest error [18].

In terms of the aforementioned reference samples or ground truth training data, previous work has shown that the results of remote predictions of geospatial characteristics, such as shoreline position or DEM error, are mixed [19]. Here, remote predictions are defined to be the results generated by models trained on nonlocal ground truth data. Local, site-specific ground truth data have been shown to enhance predictions of DEM error [2], [3], [20].

This article presents two RF models for mitigating the bias in the lidar DEMs of two ecologically similar estuarine systems in the NGOM. Sentinel-2A imagery, the lidar DEM elevations, and field-measured topographic elevations were used to train and validate the models and no other *in situ* vegetation measurements were used. The RF models were also used to make remote DEM error predictions on nonlocal ground truth data to investigate the effectiveness of remote predictions in this application.

II. Methodology

To mitigate the effects of DEM bias in this microtidal region, DEM elevations in coastal salt marsh/emergent wetland areas were systematically and variably lowered. The process for lowering the lidar DEM elevations used machine learning to integrate *in situ* topographic measurements with multispectral satellite imagery. The overall objective of the lidar DEM bias mitigation effort was to provide a topographically accurate marsh platform model for use in simulations of salt marsh evolution and migration in response to SLR. The effectiveness of using nonlocal ground truth data to train the machine learning models was also investigated.

A. Research Setting

The first setting for this study is the Apalachee Bay/St. Marks region of Florida's NGOM coast consisting of Gulf, Franklin, Wakulla, Jefferson, and Taylor counties (see Fig. 1). This area is microtidal with MHW, mean tide level (MTL), and mean low water (MLW) of 0.228 m, 0.059 m, and −0.110 m NAVD88, respectively. The second area of interest for this study is the Pascagoula River estuary in Mississippi and Alabama lying primarily in Jackson (MS) and Mobile (AL) counties. This area is also microtidal with MHW, MTL, and MLW of 0.239 m, 0.033 m, and −0.174 m NAVD88, respectively.

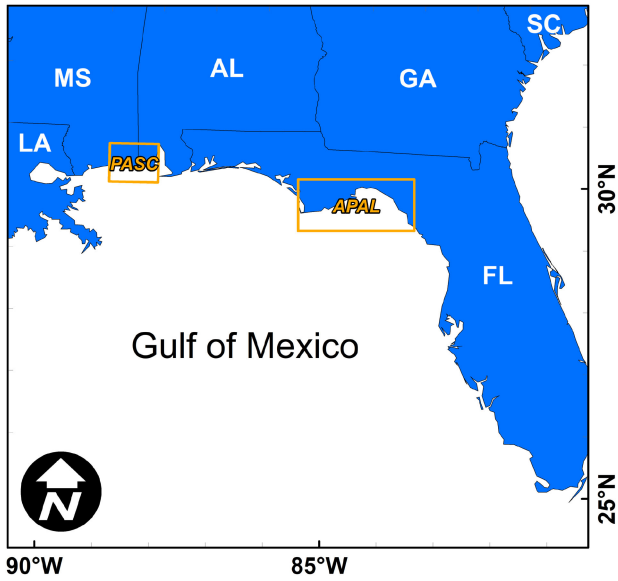


Fig. 1. Research settings in Pascagoula, MS (west) and Apalachee Bay, FL (east).

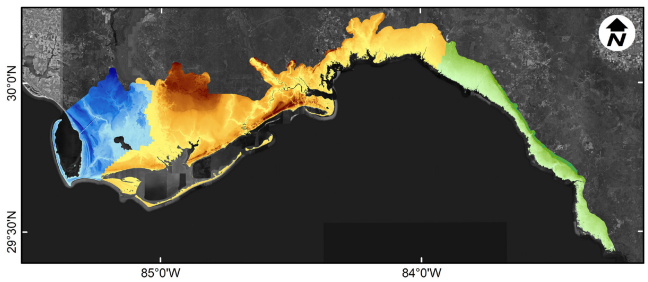


Fig. 2. Lidar DEM products used as the bases for bias mitigation in Apalachee Bay. The western (blue), central (gold), and eastern (green) lidar were acquired in 2017, 2018, and 2007, respectively.

B. Lidar Digital Elevation Model (DEM)—Apalachee Bay

The source lidar DEMs used as the basis for the bias mitigation effort in the Apalachee Bay region were provided by the North West Florida Water Management District and are now available for download from NOAA Digital Coast (https://coast.noaa.gov/digitalcoast/). The lidar DEM used in this study was compiled from three sources: 2-m resolution Lower Choctawhatchee DEM from lidar collected between 9 April and 17 May 2017 with an Aggregate Nominal Point Density (ANPD) of 4.05 points per square meter (ppsm) [21]; 1-m resolution Florida Panhandle DEM from lidar collected between 31 March and 10 May 2018 with an unreported ANPD although the report includes the Nominal Point Density values of 4.3 and 8.7 ppsm for two configurations of the Riegl VQ-1650i sensor used during mission [22]; and, in the far eastern end of the study area, an older 4.7-m resolution Florida Division of Emergency Management 2007 DEM from lidar acquired in the summer of 2007 with a reported ANPD of 4 ppsm [23] was used due to the lack of more recent data in this area (see Fig. 2).

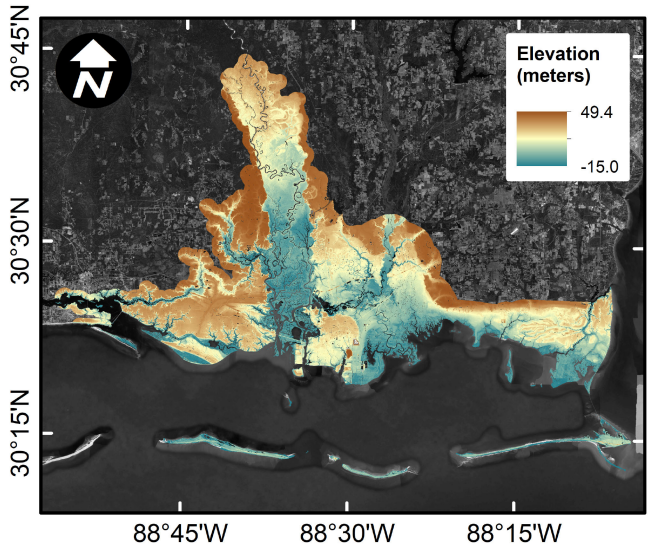


Fig. 3. 2014 USGS Lidar DEM for Pascagoula.

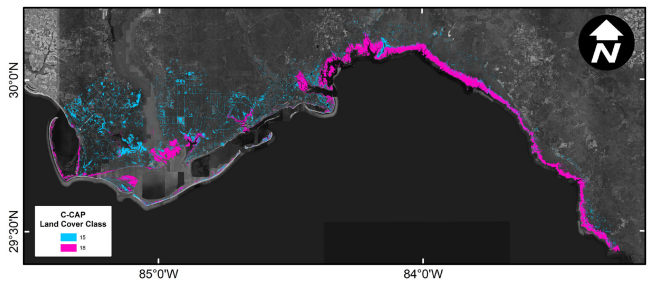


Fig. 4. Apalachee Bay land cover mask: 15—Palustrine Emergent Wetlands (blue) and 18—Estuarine Emergent Wetlands (pink).

These three DEMs were reprojected to NAD83(2011) Florida State Plane North in meters (EPSG: 6440), resampled, and coregistered to 4.7-m resolution using bilinear interpolation, and elevations were converted to meters NAVD88. They were mosaicked together with priority to the more recent data (2018, 2017, and then 2007). The overall DEM was clipped to just the coastal areas (see Fig. 2) and checked for discontinuities using topographic profile transects, especially at the boundary between the 2018 and 2007 DEMs (gold and green areas, respectively). There were no abrupt discontinuities that would indicate an error in the combined DEM.

Since the DEM bias mitigation adjustment is focused on salt marshes and coastal wetlands, the 30-m resolution 2016 Coastal Change Analysis Project (C-CAP) land use data [24] were used to restrict the DEM adjustments to areas classified as estuarine emergent wetlands (C-CAP Class 18) and palustrine emergent wetlands (C-CAP Class 15). These C-CAP regions were used to constrain the DEM bias mitigation adjustments to areas identified as coastal wetlands or salt marshes (see Fig. 4 and 5). Open water areas (C-CAP Class 21) were also masked out of the source DEM to minimize or eliminate any changes to the estuarine and nearshore bathymetry.

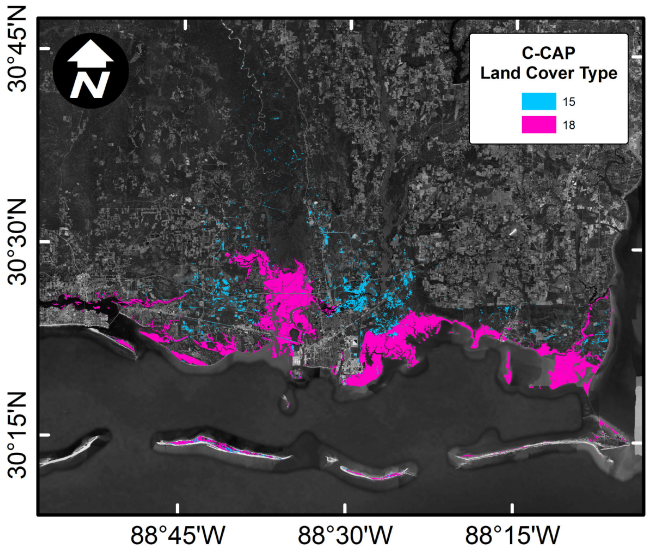


Fig. 5. Pascagoula land cover mask: 15—Palustrine Emergent Wetlands (blue) and 18—Estuarine Emergent Wetlands (pink).

C. *Lidar Digital Elevation Model (DEM)—Pascagoula River*

The source lidar DEM used as the basis for the bias mitigation effort for the Pascagoula River region was developed by the United States Geological Survey and downloaded from NOAA Digital Coast. The 3-m resolution CoNED Topobathymetric DEM is based on lidar data collected in 2014 [25]. This DEM was reprojected to NAD83(2011) UTM Zone 16N (EPSG: 6345) and the elevations were converted to meters NAVD88. The DEM was then clipped to include the lower Pascagoula River estuary as well as the coastal areas at the mouth of the river. The same C-CAP coastal wetlands masking strategy used in Apalachee Bay was also used in Pascagoula. In addition, because the Pascagoula source DEM contained both topographic and bathymetric data, areas classified by C-CAP as open water including rivers, creeks, nearshore, and offshore were masked out and not included in the DEM adjustment (see Fig. 3).

D. In Situ *Topographic Data Acquisition*

The *in situ* topographic data for the Apalachee Bay region were collected in the salt marsh and adjacent upland areas around the St. Marks lighthouse (see Fig. 4) using Real Time Kinematic Global Navigation Satellite System (RTK-GNSS) survey equipment in March of 2018. Virtual Reference Station corrections were obtained from the Florida Permanent Reference Network maintained by the Florida Department of Transportation in Real Time Correction Message (RTCM) version 3.1 format. A total of 377 spot elevations were collected across various transects located with an emphasis on capturing the topographic gradient from the water surface to the high marsh and uplands.

The *in situ* topographic data for the Pascagoula region were also collected using RTK-GNSS. Single base corrections were obtained from the Gulf Coast Geospatial Center Real Time Network operated by the University of Southern Mississippi in RTCM 3.1 MAX format. A total of 610 spot elevations were

collected along transects in the lower river estuary designed to capture topographic gradient like the protocol in St. Marks (see Fig. 5). Topographic data at both sites were converted to orthometric heights in meters NAVD88 in real time by using a Geoid separation file of the Continental US based on GEOID12B.

E. Satellite Imagery Acquisition and Processing

Sentinel-2A MSI Level 1C data were downloaded from the European Space Agency (ESA) Copernicus website. The following are the scene identifiers acquisition dates, times (GMT), and the verified tide elevation (NAVD88) from the Apalachicola tide station (National Ocean Service Station ID: 8728690) used in Apalachee Bay:

1) T16RFT, March 8, 2018, 16:11, 0.008 m;
2) T16RFU, March 31, 2018, 16:19, −0.072 m;
3) T16RGT, March 8, 2018, 16:11, 0.008;
4) T16RGU, March 8, 2018, 16:11, 0.008;
5) T17RKN, March 28, 2018, 16:10, −0.100 m;
6) T17RKP, March 8, 2018, 16:11, 0.008.

Similarly, the following are the scene identifiers acquisition dates, times (GMT), and the tide elevation (NAVD88) from the Pascagoula NOAA Lab tide station (National Ocean Service Station ID: 8741533) used in Pascagoula:

1) T16RCU, March 19, 2019, 16:30, 0.072 m;
2) T16RCV, March 19, 2019, 16:30, 0.072 m.

The images were chosen because they were the most cloud-free and were captured at the same time of year as the ground truth topographic data. Each scene was processed to Level 2A bottom of atmosphere (BOA) reflectance products using the sen2cor software provided by ESA. For each scene, bands 2 (blue, central wavelength 492.4 nm/bandwidth 66 nm), 3 (green, 559.8 nm/36 nm), 4 (red, 664.6 nm/31 nm), and 8 (near infrared, 832.8 nm/106 nm) were extracted [26], mosaicked, and reprojected to NAD83(2011) Florida State Plane North for Apalachee Bay, and NAD83(2011) UTM zone 16N for Pascagoula, using the ESA SNAP software.

F. Geospatial Data Fusion

All data were integrated using ArcGIS to produce the training and validation data sets for the bias mitigation machine learning models as well as the application points across the study areas.

First, the topographic spot elevations were converted to a point feature class containing *xyz* coordinates as measured in the field. The lidar DEM elevations along with the reflectance values from the four multispectral satellite imagery spectral bands were interpolated onto the *in situ* spot elevations. The difference between the measured marsh platform elevation and lidar DEM elevation was calculated and added as a field; this represents the DEM error and serves as the label in the machine learning model explained below. The following is a list of the fields in the point data corpus (∗ indicates predictor fields and ∗∗ indicates the target value or label):

1) Northing (meters), field measured *y* coordinate;
2) Easting (meters), field measured *x* coordinate;
3) Elevation (meters), field measured *z* coordinate;
4) ∗DEM_Elevation (meters), elevation interpolated from lidar DEM;
5) ∗B2_Blue, representing the 492.4-nm band reflectance;
6) ∗B3_Green, representing the 559.8-nm band reflectance;
7) ∗B4_Red, representing the 664.6-nm band reflectance;
8) ∗B8_NIR, representing the 832.8-nm band reflectance;
9) ∗∗Elev_DIFF, elevation difference between the lidar DEM and field measured elevation, calculated from other fields as DEM_Elevation – Elevation.

The reflectance values in the processed S2A MSI reflectance data are provided in digital number (integer) format. The S2A MSI specifications indicate that these integer reflectance values are computed by multiplying the raw floating-point reflectance values by a quantification value. In the metadata for all scenes used in this article, the quantification value is 10 000 [26]. Therefore, this was the value used to convert BOA reflectances to floating point values, which are more easily handled as input to a machine learning model.

The projected, clipped, and C-CAP-masked lidar DEM, which represents the collective set of points that need to be adjusted, was converted to a point feature class in ArcGIS. Similar to the field data, the reflectance values from the four-satellite imagery spectral bands were interpolated onto the points to form the application data corpus. All the predictor fields listed earlier, along with Northing and Easting, are present in the application data corpus.

G. Elevation Adjustment Model

The spatially variable elevation adjustments used to mitigate the bias in the lidar DEM were determined using an RF model. The RF model was implemented in Python using the scikit-learn module [27]. The RF model hyperparameters were left at their default values with the exception of the number of trees being set to 300 (n_estimators parameter). Preliminary tests indicated that more than 300 trees offered no increase in prediction accuracy. The random state parameter was also set to an arbitrary number (59) for reproducibility purposes.

Due to the relatively small ($n = 377$ for St. Marks and $n = 610$ for Pascagoula) size of the field data corpus, RF model validation was executed using a leave-one-out cross-validation protocol. In this study, one record in the training dataset was held out, the model was trained on the remaining data, and the trained model was used to predict the held-out value. This was repeated for all records in the training data. The statistical metrics used included a 1-to-1 plot with a coefficient of determination (R^2), mean absolute error (MAE), and mean absolute percentage error (MAPE). After the validation phase was complete, the entire data sets were used to train the production RF models used to adjust the DEMs for publication, presentation, and marsh modeling. For comparison, a linear regression (LR) model was also constructed and validated in the same manner on the same data to justify the use of the more complex RF model.

To examine the necessity of local ground truth topographic data, the production RF models were each used to predict the DEM errors in the other study region (i.e., the APAL model was

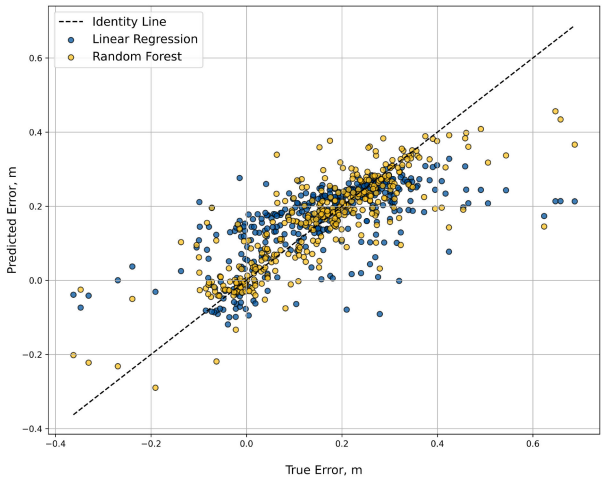


Fig. 6. Prediction results from the validation of the RF and LR models for the Apalachee Bay ($n = 377$) region.

TABLE I
DEM ERROR PREDICTION RESULTS FOR APALACHEE BAY VALIDATION ($N = 377$)

Metric	*Random Forest*	*Linear Regression*
MAE (m)	0.054	0.080
MAPE (%)	0.87	2.02
R^2	0.701	0.435

used to predict errors in the PASC DEM and vice versa). A combined model (ALL) was also validated on the entire dataset from both regions using the leave-one-out cross-validation procedure explained earlier.

Lastly, kernel density estimation plots were used to depict the error distributions of the source and adjusted (RF and LR) DEMs for the APAL, PASC, and ALL scenarios.

III. RESULTS AND DISCUSSION

A. *Lidar Bias Mitigation—Apalachee Bay*

The results of the leave-one-out cross-validation procedure for Apalachee Bay are shown in Fig. 6. Recall that this reflects the results of using 376 data points to train the model and testing it on the single held-out value for a total of 377 training/test cycles. The RF model achieved $R^2 = 0.701$, which indicates more scatter around the identity line than expected. However, it was superior to the LR model that achieved $R^2 = 0.435$.

As shown in Table I, the RF model's MAE = 0.054 m is superior to that of LR and is a 69.1% improvement over the MAE = 0.177 m for the unadjusted DEM.

As stated in Section II-G, all data were used to train the production RF model used to produce a bias-mitigated or adjusted DEM for the Apalachee Bay region. Fig. 7 shows the entire adjusted DEM and there appear to be no marked discontinuities that would indicate errors in the source data or bias mitigation

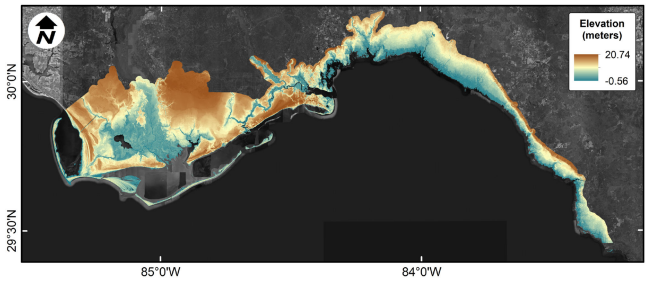


Fig. 7. Lidar DEM of the Apalachee Bay FL region adjusted using the RF model.

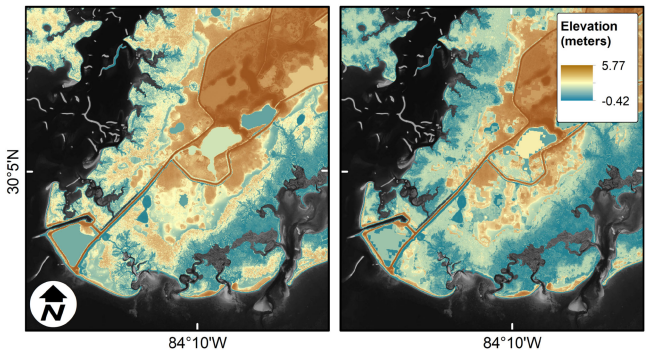


Fig. 8. Comparison of original (left) and RF adjusted (right) DEMs in the area around St. Marks Lighthouse.

process. The zoomed-in area in Fig. 8 shows a manmade impoundment in the vicinity of a major regional landmark, St. Marks Lighthouse. Visual observation indicates that the adjusted DEM is lower than the original and the variability in the terrain is preserved or in some cases enhanced. Examples of this are visible in the northeastern and southwestern areas of the site. In the original DEM, the elevations in this area were relatively uniform as shown by the brown-orange block in the image. This may be attributable to the common practice of hydro-flattening implemented in the source DEM generation process. In the adjusted DEM, there is clearly more topographic variability in this section. This is evidence of the dense marsh vegetation not only causing the documented bias in lidar DEM, but also the artificial smoothing of the underlying microtopography.

B. *Lidar Bias Mitigation—Pascagoula*

The results of the leave-one-out cross-validation procedure for the Pascagoula region are shown in Fig. 9. The data were more clustered around the 1-to-1 line than the Apalachee Bay data with the RF model achieving $R^2 = 0.817$ and the LR model achieving an $R^2 = 0.757$. Like the Apalachee Bay region, the RF was the better performing model however the gap between the two methods was narrowed.

As shown in Table II, the RF model's MAE = 0.045 m is superior to that of LR and is a 90.9% improvement over the MAE = 0.493 m for the unadjusted DEM. However, the LR model also performed well with MAE = 0.054 m representing an 89.1% improvement over the unadjusted DEM.

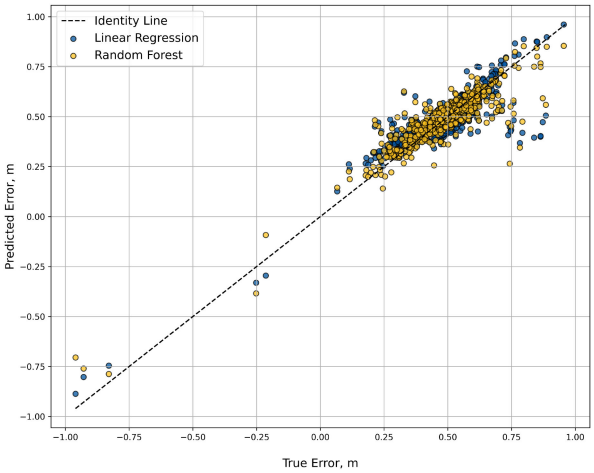


Fig. 9. Prediction results from the validation of the RF and LR models for the Pascagoula ($n = 610$) region.

TABLE II
DEM ERROR PREDICTION RESULTS FOR THE PASCAGOULA VALIDATION ($N = 610$)

Metric	*Random Forest*	*Linear Regression*
MAE (m)	0.045	0.054
MAPE (%)	0.10	0.12
R^2	0.817	0.757

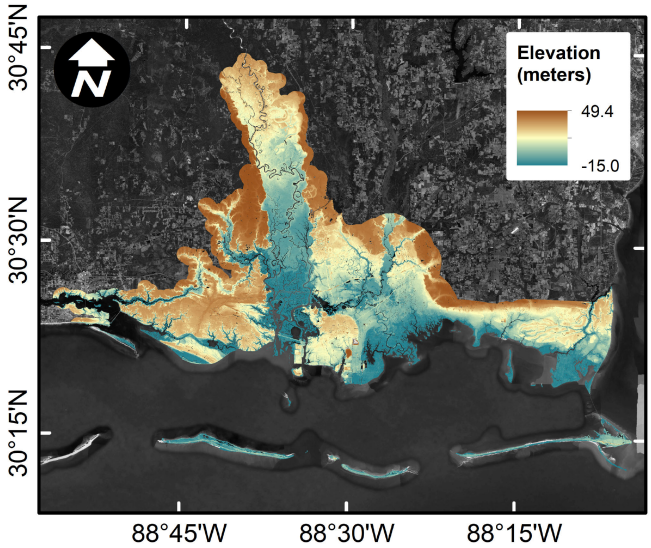


Fig. 10. Lidar DEM of Pascagoula adjusted using the RF model.

The production RF model was also applied to the regional lidar DEM for the Pascagoula region. The final adjusted DEM is shown in Fig. 10 and a side-by-side comparison of the adjusted and source DEMs in a zoomed-in area is shown in Fig. 11. In the side-by-side comparison, the right image showing the adjusted

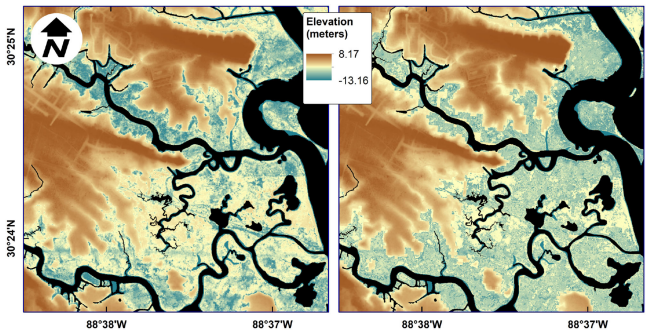


Fig. 11. Comparison of original (left) and RF adjusted (right) DEMs.

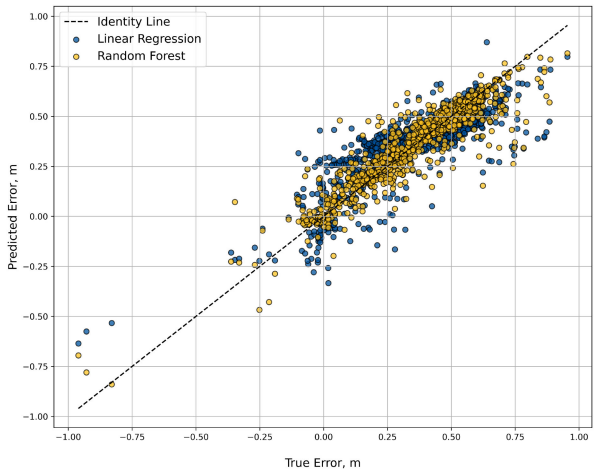


Fig. 12. Prediction results from bootstrap validation of the RF and LR models for the combined Apalachee Bay and Pascagoula datasets ($n = 987$).

TABLE III
DEM ERROR PREDICTION RESULTS FOR THE COMBINED APALACHEE BAY AND PASCAGOULA VALIDATION ($N = 987$)

Metric	*Random Forest*	*Linear Regression*
MAE (m)	0.053	0.097
MAPE (%)	0.45	1.20
R^2	0.861	0.686

DEM displays more consistent blue coloring indicating lower elevations adjacent to the creek network.

C. Cross-Validation Using Other Location

To determine the value and necessity of *in situ* data for each location requiring a lidar DEM adjustment, we employed a regionally mixed model application procedure. First, a leave-one-out bootstrap validation procedure was performed on RF and LR models using the entire $n = 987$ dataset. The results of the bootstrap validation procedure for the combined dataset (ALL) are shown in Fig. 12 and Table III.

As shown in Table III, the ALL RF model's MAE = 0.053 m falls between those of the APAL and PASC. The ALL LR model

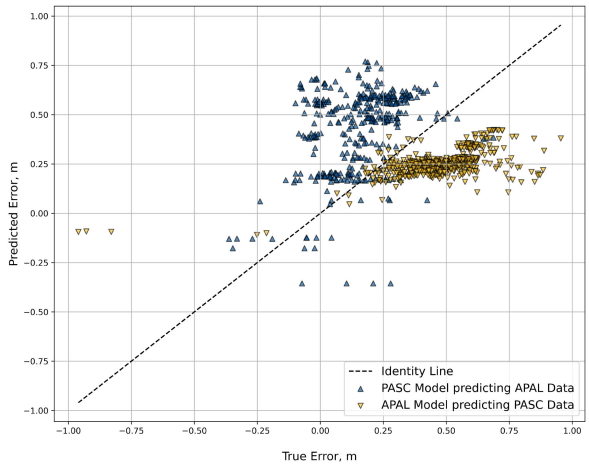


Fig. 13. Remote prediction validation results.

was worse than the LR models of both the APAL and PASC cases. The performance of the ALL RF model was superior in terms of its $R^2 = 0.861$.

Subsequently, each production RF model was used to predict DEM errors for the other location in a remote prediction test. Fig. 13 shows that neither production model was able to predict the DEM errors at the other location. Using the Pascagoula model to predict the errors in the Apalachee Bay DEM resulted in a wide scatter and a high elevation bias. The predictions made using the Apalachee Bay model to predict the errors in the Pascagoula DEM were tightly clustered between elevations 0.00 and 0.50 m and biased low.

Based on these results, we recommend that any DEM adjustment model should contain ground truth training data within the area of interest. However, the performance of the combined model indicates that a dataset containing a large range of errors may be able to train an effective yet more generalized and less region-specific DEM bias mitigation model.

D. Error Distributions

Kernel density estimation plots for the errors in the lidar DEM both before and after adjustment are shown in Fig. 1. The error distributions shifted closer to zero when both the RF- and LR-based corrections were applied. The RF model has the highest density and tightest distribution around zero.

E. Feature Importance From Random Forest (RF) Models

RF models use an internal out-of-bag testing procedure during training to produce the feature importance for each predictor feature. These values are an indication of the relative contribution of each predictor variable to the overall prediction of the DEM error (see Table III).

As shown in Table IV, the lidar DEM elevation is the most important predictor in all validation scenarios except ALL. In that case, the blue band from the S2A imagery has slightly higher importance.

TABLE IV
FEATURE IMPORTANCE VALUES FROM THE RANDOM FOREST MODELS

Predictor	*APAL model*	*PASC model*	*ALL model*
DEM_Elevation	0.373	0.628	0.367
B2_Blue	0.124	0.089	0.391
B3_Green	0.231	0.094	0.101
B4_Red	0.138	0.089	0.064
B8_NIR	0.135	0.101	0.077

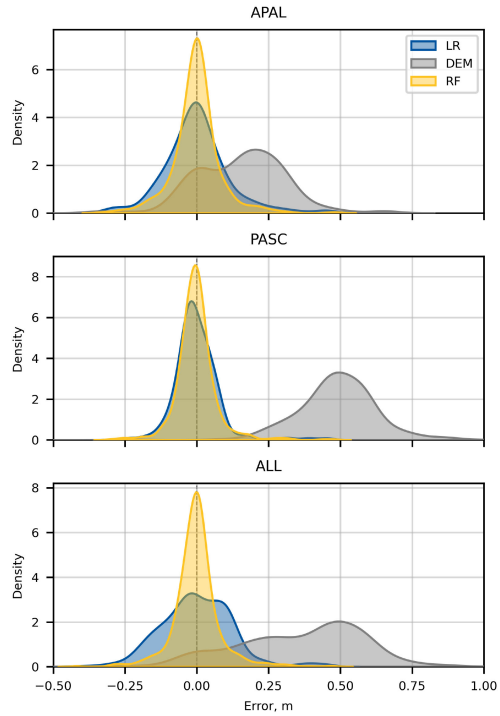


Fig. 14. Kernel density estimation plots for elevation errors in the source DEMs (gray) and after adjustment using the random forest (RF, gold) and linear regression (LR, blue) models for the Apalachee Bay (APAL, $n = 377$), Pascagoula (PASC, $n = 610$), and combined (ALL, $n = 987$).

F. Discussion

The RF models developed in this study achieved a sub-6-cm MAE when used to predict the bias in the lidar DEMs for the lower Pascagoula River and Apalachee Bay marshes. This level of DEM correction is effective in many modeling applications in the study regions including tidal, storm surge, and marsh evolution. The kernel density estimation plots (see Fig. 14) also show that the RF model has the highest density and tightest distribution around zero. This indicates that the RF model was able to handle the wide distribution of error magnitudes and effectively mitigate the high elevation bias in the source DEM.

The superior performance of the PASC model in relation to the APAL model could be the result of the larger number of training data records or the increased resolution of the lidar DEM (3 m in PASC versus 4.7 m in APAL). The larger training dataset is known to improve the performance of machine learning models

so, in this respect, we expected the PASC RF model to perform better. From the resolution perspective, a higher resolution DEM has fewer lidar elevations averaged into the DEM pixel elevation (assuming ANPD values are about equal) so a ground truth elevation lying in a particular pixel has a higher probability of being representative. Therefore, higher resolution DEMs are less sensitive to local slope and microtopography and models trained using local ground truth elevations and high-resolution DEMs can reasonably be expected to perform better.

Other machine learning algorithms have been also used including support vector machine, *k*-nearest neighbor, or artificial neural network, but past work has demonstrated that RF can make effective predictions in marsh environments [15]. One vital issue for discussion is the inclusion of the lidar DEM elevation itself in the predictor feature set. At first, this may appear to be circular logic but there is a crucial reason for doing so. Currently, the presence of the persistent high elevation bias in a salt marsh lidar DEM is determined or confirmed by local ground-truth topographic data in the salt marsh. This is a labor-intensive process that can only be done effectively by capable and experienced personnel to mitigate damage to the marsh environment and produce reliable data. As machine learning models for this application become more ubiquitous and broadly trained on a wider range of data, we would like them to be able to detect the presence of bias as well as correct for it. By including the DEM elevation as a predictor feature, we are taking a step towards training a machine learning model that may, based on the DEM elevation in combination with other predictor variables, determine that there is no bias present in a particular pixel or set of pixels. In our mind, this is the most important future work direction in this field.

With that in mind, we must also consider other sources of error in these methods and estimate their impact to the results. Potential sources of error in the development of the DEM adjustment models presented here came from the timing, precision, and accuracy of the ground truth and remotely sensed observations.

From the timing perspective, we synchronized the satellite imagery sensing dates with the ground truth data acquisition dates to associate the spectral characteristics of the vegetation with the true elevation of marsh surface. However, an argument can be made for synchronizing the satellite imagery sensing dates with the lidar acquisition dates to associate the spectral characteristics of the vegetation to the bias, which is subsequently estimated and removed using a model like the RF presented here. Further testing could prove that the latter is a more robust approach. As stated in the methodology section, all data were acquired in the early spring, and in the case of ground truth and satellite imagery, all data were acquired in March. This eliminated many of the seasonal effects associated with the growth and recession cycle of the marsh grasses. In Apalachee Bay, the ground truth data were acquired approximately 3 weeks before the lidar acquisition period so it is unlikely that any significant accretion occurred in the marsh during that period. However, applying the trained model to data collected in 2017 or worse, 2007, could result in erroneous predictions of DEM bias. In Pascagoula, the lidar data were collected in 2014 and the ground truth data in 2019, representing a 5-year difference. According to typical marsh accretion rates for similar systems in this region, we estimate that the accretion was less than 2 cm [28]. Therefore, the seasonal consistency between the lidar acquisition and remote sensing data was determined to be the governing factor.

The water level during the lidar flights and the satellite imagery acquisition is also known to influence both lidar returns and the spectral characteristics of the satellite imagery. The water levels at the time of the lidar data acquisition are unknown and vary over the course of the flight missions. All S2A scenes for Apalachee Bay were acquired when the water was below MTL, therefore it is likely that most of the marsh was above water. For Pascagoula, the two scenes were acquired when the water was $\sim$4 cm above MTL. In this case, some of the marshes may have been inundated, especially on the fringes of tidal creeks, as most of the area in *Juncus*-dominated marshes tend to lie between MHW and just below MTL [29]. The presence of standing water would impact the spectral response of the marsh in the satellite imagery.

The horizontal georeferencing accuracy for S2A imagery is estimated to be less than one 10-m pixel [30] so it is fair to say the reflectances used here can reasonably be assumed to represent the reflectances at the ground truth locations. The blue, green, red, and NIR bands from the S2A MSI imagery were selected for multiple reasons. Red and NIR are known to indicate vegetation health as they are the two bands used in the normalized difference vegetation index (NDVI) and its derivatives. However, the presence of the Red and NIR as predictor variables does not capture all the information inherent in spectral ratio indices like NDVI and should not be construed as drop-in replacements for them. The relatively low feature importances of the Red and NIR predictors shown in Table IV support this. These four bands are also frequently combined in instruments on satellite, aircraft, and small unmanned aerial system platforms so an investigation into their use in this context is intended to provide a basis for future data collection efforts aimed at machine learning applications. For example, the PlanetScope product produces multichannel imagery using these four bands at approximately 4-m resolution with a daily revisit time [31]. Due to differences in instrument characteristics such as band center wavelengths and bandwidths, transferring the findings presented here should be undertaken with caution and should consider retraining the models with reflectances generated by the new sensor. Furthermore, our granular approach using the reflectances from each band separately as predictor values should be tested against or even combined with those that use indices like NDVI as this may reveal some of the additional signals contained within the ratio-based indices, for example, photosynthetic activity, and proxies for vegetation height and density [3], [32].

IV. Conclusion

While the ability to accurately capture the topography of salt marsh platforms using lidar has improved, there is still a persistent high bias that is significant enough to negatively impact marsh modeling results in microtidal environments. This bias is largely attributed to the dense grasses over the salt marsh surface as well as tidal inundation at the time of the lidar data acquisition.

Local ground-truth elevation data in coastal salt marshes are necessary to detect and quantify the surface elevation bias and to train machine learning approaches to mitigate it.

The method presented here used Sentinel-2A multispectral imagery coupled with ground-truth elevations to construct an RF bias mitigation model. The RF model reduced the bias, as measured by MAE of 69% in Apalachee Bay, Florida, USA, and 90% in Pascagoula, Mississippi, USA, to MAE of approximately 5 cm. This was superior to the MAE achieved by an LR model fit using the same source data thereby justifying the decision to use the more computationally intensive RF approach. Additionally, the source lidar DEM elevation was the most important predictor variable in the RF. This suggests that inundation potential and bare ground visibility, known to vary with elevation, matter more than vegetation vigor in predicting lidar elevation bias in salt marshes. This also suggests that future models trained on more diverse ground-truth topographic data may be able to distinguish between DEM pixels with the high elevation bias and those without.

Future work will focus on increasing the ground truth data corpus size, incorporating other relevant environmental variables such as local mean sea level elevation, using higher resolution remote sensing data including ratio-based indices to develop a more universally applicable bias mitigation model.

Acknowledgment

The authors would like to thank Karim Alizad, Madeline Foster-Martinez, Ryan Lauve, Katherine Renken, Gwen Miller, Mike Archer, and Jonathan Pitchford for their assistance in the field; Terry Peacock, Refuge Manager, St. Marks National Wildlife Refuge, for facilitating access to the St. Marks field site; and Willa Brantley, Director, Bureau of Wetlands Permitting, Mississippi Department of Marine Resources for facilitating access to the Pascagoula Site.

References

[1] K. Alizad, S. C. Medeiros, M. R. Foster-Martinez, and S. C. Hagen, "Model sensitivity to topographic uncertainty in meso- and microtidal marshes," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 807–814, 2020, doi: 10.1109/jstars.2020.2973490.

[2] C. Hladik and M. Alber, "Accuracy assessment and correction of a LIDAR-derived salt marsh digital elevation model," *Remote Sens. Environ.*, vol. 121, pp. 224–235, 2012, doi: 10.1016/j.rse.2012.01.018.

[3] K. J. Buffington, B. D. Dugger, K. M. Thorne, and J. Y. Takekawa, "Statistical correction of lidar-derived digital elevation models with multispectral airborne imagery in tidal marshes," *Remote Sens. Environ.*, vol. 186, pp. 616–625, 2016, doi: 10.1016/j.rse.2016.09.020.

[4] J. T. Morris, P. V. Sundareshwar, C. T. Nietch, B. Kjerfve, and D. R. Cahoon, "Responses of coastal wetlands to rising sea level," *Ecology*, vol. 83, no. 10, pp. 2869–2877, 2002, doi: 10.1890/0012-9658083[2869:ROCWTR]2.0.CO;2.

[5] K. Alizad *et al.*, "A coupled, two-dimensional hydrodynamic-marsh model with biological feedback," *Ecological Model.*, vol. 327, pp. 29–43, 2016, doi: 10.1016/j.ecolmodel.2016.01.013.

[6] J. T. Morris, "Ecological engineering in intertidal salt marshes," *Hydrobiologia*, vol. 577, no. 1, pp. 161–168, 2007, doi: 10.1007/s10750-006-0425-4.

[7] K. Alizad, S. C. Hagen, J. T. Morris, S. C. Medeiros, M. V. Bilskie, and J. F. Weishampel, "Coastal wetland response to sea-level rise in a fluvial estuarine system," *Earth's Future*, vol. 4, no. 11, pp. 483–497, 2016, doi: 10.1002/2016EF000385.

[8] M. V. Bilskie, D. Coggin, S. C. Hagen, and S. C. Medeiros, "Terrain-driven unstructured mesh development through semi-automatic vertical feature extraction," *Adv. Water Resour.*, vol. 86, pp. 102–118, Dec. 2015, doi: 10.1016/j.advwatres.2015.09.020.

[9] J. C. Dietrich *et al.*, "Hurricane Gustav (2008) waves and storm surge: Hindcast, synoptic analysis and validation in Southern Louisiana," *Monthly Weather Rev.*, vol. 139, pp. 2488–2522, 2011, doi: 10.1175/2011MWR3611.1.

[10] S. Bunya *et al.*, "A high-resolution coupled riverine flow, tide, wind, wind wave and storm surge model for southern Louisiana and Mississippi. Part I: Model development and validation," *Monthly Weather Rev.*, vol. 138, pp. 345–377, 2010.

[11] M. Baradaranshoraka, J.-P. Pinelli, K. Gurley, X. Peng, and M. Zhao, "Hurricane wind versus storm surge damage in the context of a risk prediction model," *J. Struct. Eng.*, vol. 143, no. 9, 2017, Art. no. 04017103, doi: 10.1061/(asce)st.1943-541x.0001824.

[12] S. C. Medeiros, S. C. Hagen, J. F. Weishampel, and J. J. Angelo, "Adjusting lidar-derived digital terrain models in coastal marshes based on estimated aboveground biomass density," *Remote Sens.*, vol. 7, no. 4, pp. 3507–3525, 2015. [Online]. Available: http://www.mdpi.com/2072-4292/7/4/3507

[13] K. Alizad *et al.*, "Dynamic responses and implications to coastal wetlands and the surrounding regions under sea level rise," *PLoS One*, vol. 13, no. 10, 2018, Art. no. e0205176, doi: 10.1371/journal.pone.0205176.

[14] J. N. Rogers, C. E. Parrish, L. G. Ward, and D. M. Burdick, "Improving salt marsh digital elevation model accuracy with full-waveform lidar and nonparametric predictive modeling," *Estuarine, Coastal Shelf Sci.*, vol. 202, pp. 193–211, 2018, doi: 10.1016/j.ecss.2017.11.034.

[15] H. M. Cooper, C. Zhang, S. E. Davis, and T. G. Troxler, "Object-based correction of LiDAR DEMs using RTK-GPS data and machine learning modeling in the coastal EverGlades," *Environ. Model. Softw.*, vol. 112, pp. 179–191, Feb. 2019, doi: 10.1016/j.envsoft.2018.11.003.

[16] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogrammetry Remote Sens.*, vol. 114, pp. 24–31, Apr. 2016, doi: 10.1016/j.isprsjprs.2016.01.011.

[17] T. Hu, Y. Zhang, Y. Su, Y. Zheng, G. Lin, and Q. Guo, "Mapping the global Mangrove forest aboveground biomass using multisource remote sensing data," vol. 12, no. 10, 2020, Art. no. 1690. [Online]. Available: https://www.mdpi.com/2072-4292/12/10/1690

[18] F. E. Fassnacht, H. Latifi, and F. Hartig, "Using synthetic data to evaluate the benefits of large field plots for forest biomass estimation with LiDAR," *Remote Sens. Environ.*, vol. 213, pp. 115–128, Aug. 2018, doi: https://doi.org/10.1016/j.rse.2018.05.007.

[19] S. Banks, K. Millard, J. Pasher, M. Richardson, H. Wang, and J. Duffe, "Assessing the potential to operationalize shoreline sensitivity mapping: Classifying multiple wide fine quadrature polarized RADARSAT-2 and Landsat 5 scenes with a single random forest model," *Remote Sens.*, vol. 7, no. 10, pp. 13528–13563, 2015. [Online]. Available: https://www.mdpi.com/2072-4292/7/10/13528

[20] A. McClure, X. Liu, E. Hines, and M. C. Ferner, "Evaluation of error reduction techniques on a LIDAR-derived salt marsh digital elevation model," *J. Coastal Res.*, vol. 32, no. 2, pp. 424–433, 2016. [Online]. Available: https://doi.org/10.2112/JCOASTRES-D-14-00185.1

[21] Dewberry, "FL lower choctawhatchee NWFWMD lidar 2017 B17," Dewberry, Tampa, FL, USA, Nov. 2018. [Online]. Available: https://coast.noaa.gov/htdata/lidar2_z/geoid12b/data/8681/supplemental/NWFWMD_FL_Lower_Choctawhatchee_Topo_Lidar_Project_Report.pdf

[22] I. Airborne Imaging, "LiDAR acquisition and calibration report – Florida Panhandle," Airborne Imaging, Inc., Calgary, AB, Canada, 2018. [Online]. Available: https://coast.noaa.gov/htdata/lidar3_z/geoid18/data/8945/supplemental/3143_NWFWMD_FL_PanHandle_Aquisition_Report.pdf

[23] Dewberry, "Final report of specific purpose lidar survey - LiDAR, breaklines and contours for Taylor County, Florida," Dewberry, Fairfax, VA, USA, Oct. 2009. Accessed: Jun. 10, 2022. [Online]. Available: https://noaa-nos-coastal-lidar-pds.s3.amazonaws.com/laz/geoid18/554/supplemental/Taylor_County_FL_LiDAR_Survey_Report.pdf

[24] Office for Coastal Management, National Oceanic and Atmospheric Administration, "C-CAP regional land cover and change," 2016. [Online]. Available: https://coast.noaa.gov/digitalcoast/data/ccapregional.html

[25] OCM Partners, "Topobathymetric model of the Northern Gulf of Mexico, 1888 to 2013," 2014. [Online]. Available: https://www.fisheries.noaa.gov/inport/item/49465

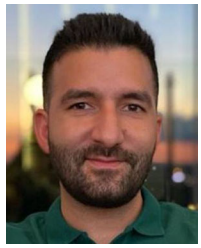
[26] Thales Alenia Space, "Sentinel-2 products specification document," European Space Agency, Paris, France, Mar. 2021, Accessed: Aug. 20, 2021. [Online]. Available: https://sentinel.esa.int/documents/247904/685211/Sentinel-2-Products-Specification-Document.pdf/fb1fc4dc-12ca-4674-8f78-b06efa871ab9?t=1616068001033
[27] F. Pedregosa, G. Varoquaux, A. Gramfort, and V. Michel, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011, doi: 10.5555/1953048.2078195.
[28] V. D. Hansen and J. A. Nestlerode, "Carbon sequestration in wetland soils of the northern Gulf of Mexico coastal region," *Wetlands Ecol. Manage.*, vol. 22, no. 3, pp. 289–303, 2013, doi: 10.1007/s11273-013-9330-6.
[29] J. T. Morris *et al.*, "Integrating LIDAR elevation data, multispectral imagery and neural network modelling for marsh characterization," *Int. J. Remote Sens.*, vol. 26, no. 23, pp. 5221–5234, 2007, doi: 10.1080/01431160500219018.
[30] B. Vajsová and P. J. Åstrand, "New sensors benchmark report on Sentinal-2A," Joint Research Centre, European Union, Strasbourg, France, Rep. EUR 27674 EN, 2015. Accessed: 24 Mar. 2022. [Online]. Available: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC99517/lb-na-27674-en-n%20.pdf
[31] Planet Labs PBC, "Planet imagery product specifications," Planet Labs PBC, San Francisco, CA, USA, Jun. 2020. Accessed: 23 Mar. 2022. [Online]. Available: https://earth.esa.int/eogateway/documents/20142/37627/Planet-combined-imagery-product-specs-2020.pdf
[32] J. R. Holmquist *et al.*, "Scalability and performance tradeoffs in quantifying relationships between elevation and tidal wetland plant communities," *Mar. Ecol. Prog. Ser.*, vol. 666, pp. 57–72, 2021, doi: 10.3354/meps13683.

Stephen C. Medeiros (Member, IEEE) was born in Massachusetts, MA, USA, in 1977. He received the B.S. and M.S. degrees in civil engineering from the Florida Institute of Technology, Melbourne, FL, USA, in 1999 and 2002, respectively, and the Ph.D. degree in civil engineering from the University of Central Florida, Orlando, FL, USA, in 2012.

He is currently an Assistant Professor with the Department of Civil Engineering, Embry-Riddle Aeronautical University, Daytona Beach, FL, USA. His research focuses on the fusion of remote sensing, lidar, field measurements, and computer modeling to study the resilience of civil infrastructure and ecosystems to natural hazards.

Dr. Medeiros is a Registered Professional Engineer in the State of Florida.

Joelle S. Bobinsky was born in New Mexico, USA, in 1998. She received the B.S. and M.S. degrees in civil engineering from Embry-Riddle Aeronautical University, Daytona Beach, FL, USA, in 2021 and 2022, respectively.



Khalid Abdelwahab was born in Victoria, BC, Canada, in 1991. He received the B.S. and M.S. degrees in civil engineering from the University of Central Florida, Orlando, FL, USA, in 2016 and 2019, respectively.