Using Carbon Fluxes and Spectral Bands to Model Ecosystem Processes in Wetlands

Qi Jia Sun

ABSTRACT

Wetlands serve as significant carbon sinks in global climate dynamics. However, their sensitivity to climate variability can convert them quickly into carbon sources under adverse climate conditions. This study explores the efficacy of lidar technology in providing high-resolution data to model wetland carbon flux and vegetation structure. By employing a combination of UAV-based LiDAR surveying and multispectral imaging, this research gathered extensive ecological and topographical data from three wetland sites within the Sacramento San Joaquin River Delta: Twitchell Island, Mayberry Slough, and Dutch Slough. This paper focused on three primary objectives: 1) to evaluate the potential of LiDAR data in calculating NDVI, 2) to evaluate the potential of LiDAR data in estimating carbon flux within wetlands and 3) to develop and test a graph neural network designed to predict carbon fluxes through lidar statistics and multispectral imagery data. Data processing and exploratory analysis were conducted using software tools to process, analyze, and visualize LiDAR and multispectral data, revealing significant correlations between LiDAR-derived features such as point density, intensity, and elevation with carbon flux. The findings demonstrate that LiDAR offers a promising tool for ecological modeling and carbon flux estimation in wetland ecosystems. By modeling the data with a graph neural network, the study highlights the challenges associated with the data and emphasizes the potential of remote sensing technologies to elevate ecological monitoring and conservation strategies.

KEYWORDS

LiDAR, wetlands, carbon flux, NDVI, graph neural networks, LiDAR statistics, modeling

INTRODUCTION

Wetlands are pivotal in global climate regulation, acting as carbon sinks through the sequestration of atmospheric carbon in soil and vegetation (Nahlik & Fennessy, 2016). The carbon sequestration process in wetlands is characterized by significant fluctuations between photosynthesis, decomposition, and respiration regulating atmospheric carbon dioxide levels (Valach et. al 2021). However, wetlands are extremely sensitive to climate variability — small changes in climate can drastically alter their hydrology and carbon storage capacity — transforming them from carbon sinks to sources (Zhang et. al 2023). Thus, wetland conservation is essential for biodiversity, water quality, and climate regulation in the global carbon cycle.

LiDAR technology is an important method of remote sensing that can effectively penetrate dense foliage, creating intricate high-resolution models of wetlands, forest canopies, and the ground surface (Lang and McCarty 2014). Wetland ecosystems have complex vegetative structures and carbon dynamics, creating challenges in ecological modeling (Bian et al. 2021). As LiDAR generates high-resolution models, LiDAR-based methods may be key in helping better understand wetland vegetative complexities (Sharma & Naik 2023). The significance of LiDAR in ecological research is highlighted by its role in estimating effective leaf surface area, which influences photosynthesis efficiency and a plant's capacity to intercept light (Wang & Fang 2020, Hu et al. 2020). Within wetland ecosystems, where carbon flux is a key indicator of ecological health, the ability to measure leaf area accurately is indispensable. Current methods for measuring ecosystem carbon flux typically involve setting up eddy covariance towers to collect data over time (Rosentreter 2022). However, in less accessible areas, constructing these towers may be inconvenient and in some cases, disruptive to the ecosystem. Thus, precise LiDAR data, which captures vegetative structures more effectively than any other remote sensing technology, could be instrumental in enhancing our predictive ability for carbon fluxes and deepening our understanding of the intricate process of photosynthesis (Yuan et. al 2018).

This research intends to address a central research question: Can LiDAR data offer accurate estimations of NDVI and carbon footprint within wetlands? This research also seeks to address three pivotal subquestions. First, I want to find which LiDAR features are essential for accurately calculating NDVI within wetland ecosystems. I hypothesize that attributes such as point density, point intensity, and height are critical in this context. Second, the study seeks to

determine the efficacy of predicting carbon flux. I hypothesize that by better representing vertical vegetation structures, I will see similar results to NDVI. Third, the research seeks to use a graph neural network to represent wetland structure and to predict carbon flux. I hypothesize that additional information from the ecosystem will be required to help adjust the model for predictions. By using existing LiDAR data from UAVs and flux data from Ameriflux towers on Twitchell Island, Dutch Slough, and Mayberry Slough, developing features, and fine-tuning a graph neural network, this paper seeks to help better position the potential of using LiDAR data within wetland ecosystem studies.

METHODS

Study site

The data collection and data pre-processing methods were paid for by the Berkeley Biometeorology Lab and conducted by a third-party entity. The data was collected across three distinct environments within the Sacramento San Joaquin River Delta: Dutch Slough, Mayberry Slough, and Twitchell Island. These locations are wetland ecosystems with differing features that provide a diverse setting for geospatial analysis. The areas chosen for the study were within a 15-kilometer radius of the eTrac Continuously Operating Reference Station "VORTEX," located in Antioch, CA (NOAA). This selection of sites enabled a comprehensive examination of ecological and topographical characteristics within the delta region. These areas were chosen due to the availability of the data as well as the existence of the eddy-covariance towers that have continuously collected data for long periods (Baldocchi 2014). The carbon flux data is available publicly on the Ameriflux websites.

LiDAR data collection methods

The data collection employed a combination of UAV LiDAR surveying and multispectral imaging to capture high-resolution data of the study area. A total of 34 UAV flights were conducted using a DJI Matrice 600 Pro, equipped with a Riegl MiniVUX-1 Laser Scanner and an Agrowing Sony Alpha a7RII 10-band Digital Camera (Bauer & Nagy 2024). The UAV flights

were meticulously planned to optimize LiDAR data density, camera overlap, and overall flight efficiency, maintaining a consistent altitude of 90 meters above ground level. Ground Control Points were strategically placed across the survey area to prepare for the UAV survey (USGS). Marked with black and white aerial panels, these GCPs were essential for ensuring the accuracy of the LiDAR data, with each point undergoing a three-minute observation period via RTK GNSS (Fan et. al 2019).

LiDAR data processing

Trajectory processing was carried out using Novatel Inertial Explorer Xpress, which integrated GPS and GLONASS data from both the mobile platform and the base station to refine the UAV's trajectory (NovAtel). LiDAR data processing was conducted using Phoenix Spatial Explorer software, where raw scan data was time-matched with the trajectory data to produce geo-referenced scan data. Further classification and analysis of the LiDAR data were performed using LiDAR360 software.

The multispectral images captured by the UAV were batch-processed for radiometric calibration, using a 24-square Color-Checker board placed on the launch pad (Tocci et. al 2022). These images were then orthorectified to generate high-resolution mosaics (Huang et. al 2021). Additionally, the research involved the generation of Digital Terrain Models and Digital Elevation Models using Autodesk Civil3D 2023. This process used classified LiDAR data to create ASCII files that served as the basis for terrain modeling. Finally, the data included rigorous accuracy and density analyses of the LiDAR data. The LiDAR data's vertical accuracy was assessed by comparing the elevation of GCPs with the point cloud data using the LiDAR QC tool in Global Mapper (Zayed et. al 2023). Additionally, the Density Quality Analysis tool in LiDAR360 was utilized to evaluate the point cloud density for each site (Petras et. al 2023).

Carbon data collection and data processing

Eddy covariance is a method primarily capturing pointwise measurements of gasses like carbon dioxide at specific locations using sensors on flux towers (Sanchez et. al 2022). However, due to its localized nature, it does not directly offer a direct comprehensive view of gas

exchanges over larger areas. Unlike variables such as temperature or precipitation, flux data cannot be easily extrapolated or obtained at specific points (Shao et. al 2024). Thus, to extrapolate carbon flux tower data to broader spatial scales, data from the eddy covariance sensors is used alongside models to create probabilistic maps indicating the origins of air parcels interacting with the sensor. The Kljun footprint model estimates the source of atmospheric measurements taken at a specific location and stands out due to its flexibility and applicability to a wide range of heights, atmospheric stabilities, and surface roughness, making it highly versatile for ecological and meteorological studies (Kljun et. al 2015). The Kljun footprint model operates by calculating the statistical distribution of upwind surface areas that influence the measurements at the sensor location based on wind direction, wind speed, and atmospheric stability. The model outputs a footprint function that represents the probability distribution of source areas, indicating the likelihood that different locations contributed to the detected fluxes at the tower.

These datasets are derived from these models and measurements are compiled into TIFF files containing 48 layers each, representing 30-minute intervals from 00:30 to 24:00 UTC. One of these layers is displayed in Figure 1. This carbon data was averaged throughout the 48 layers for a daily average. Each pixel in these high-resolution Geotiffs represents a percentage of the total footprint, encoded logarithmically for data efficiency. Each file uses a placeholder value of 1e-28 for pixels not part of the footprint to not add additional value to the total sum. Thus, the sum of all pixel values in a single footprint should be less than or equal to 1.0. Additionally, the dataset includes daily heat maps and contour lines representing 80% of the confidence interval of the footprint during both daytime and nighttime, displayed in blue and green respectively. A yellow-to-red gradient in the maps represents a 2D histogram, indicating how frequently each pixel appears within the 80% threshold of each footprint throughout the day.

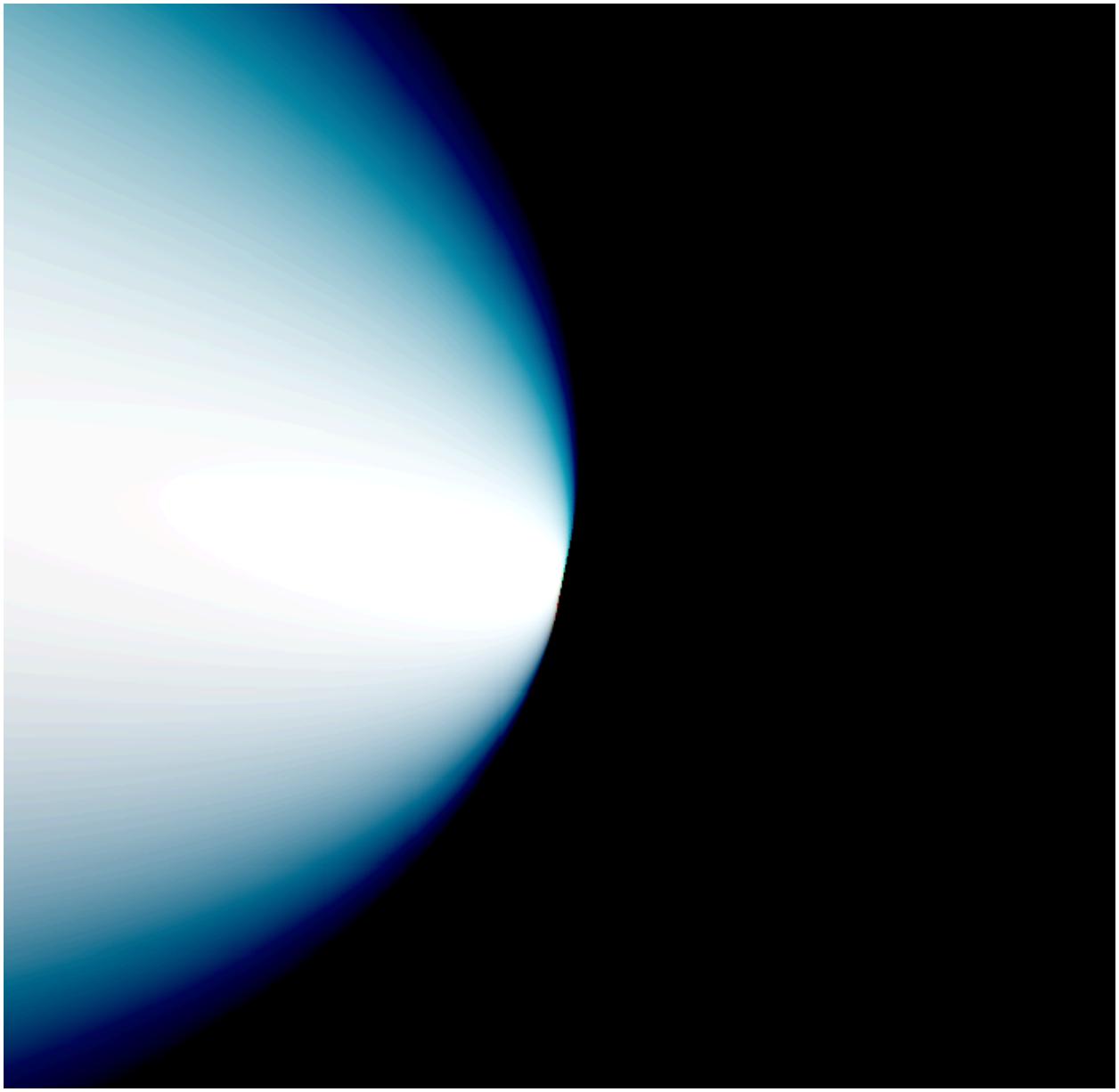

Figure 1. Kljun calculated footprint values for Twitchell Island at time 00:00.

LiDAR statistics

LiDAR data consists of XYZ coordinates (latitude, longitude, and elevation) and intensity values that correspond to the reflectivity of the surfaces from which the laser pulses are reflected (Neon). From this data, LiDAR can represent the structure and height of the data in a way that other remote sensing surveys cannot. Based on a predefined pixel, I can compute pixel-level statistics on the data. I have chosen four key statistics.

1. Mean elevation is a pivotal metric in wetland environments, as it dictates water saturation levels, vegetation types, and the area's overall hydrology (Buffington et. al 2021). I focus on mean elevation because it directly influences local climate conditions. For instance, higher mean elevations within a pixel may suggest the presence of densely stacked vegetation or tall trees, indicative of distinct ecological characteristics. Conversely, lower elevations might correspond to roads, open water, or bare ground, elements with very little carbon capacity (Zhang et. al 2021).
2. Mean density reflects the structural and density attributes of vegetation (Cekada et al. 2010). This metric is essential for estimating biomass and, consequently, the carbon storage capacity of wetlands. Denser point measurements suggest robust vegetation, crucial for effective carbon sequestration and ecosystem productivity. Areas with sparse vegetation may indicate less biomass and potentially different ecological functions (Li et al. 2013).
3. 95 percentile height can identify some of the taller areas where there could be more influence from trees on carbon flux (Lang 2014). This metric also correlates with biomass levels, pinpointing areas of higher carbon storage. Typically, taller vegetation, especially trees, will exhibit a higher 95% confidence interval of height, signifying significant biomass accumulation.
4. Mean intensity offers clues about the ground surface and vegetation types through reflectivity data, aiding in the classification of wetland areas and assessing vegetation health (Dai et. al 2022). This measure can help distinguish between wet and dry areas, which is crucial since these conditions affect carbon exchange rates differently. Variable reflectance properties can indicate how surfaces with similar characteristics might exhibit comparable carbon flux outputs.

These LiDAR statistics are generated in R (lidR), which processes LiDAR point cloud data to analyze its spatial and reflective attributes. I set the pixel size to the coarsest resolution available for LiDAR imagery at 30 cm. This selection helps mitigate the challenges posed by the sparsity and non-uniform distribution of LiDAR data across different areas. Larger pixel sizes are beneficial as they help in averaging the data, providing a more consistent output, and reducing

the effects of data clustering. After computing these statistics, the data is exported into a .tif file, making it suitable for further analysis and processing in ArcGIS.

Data manipulation

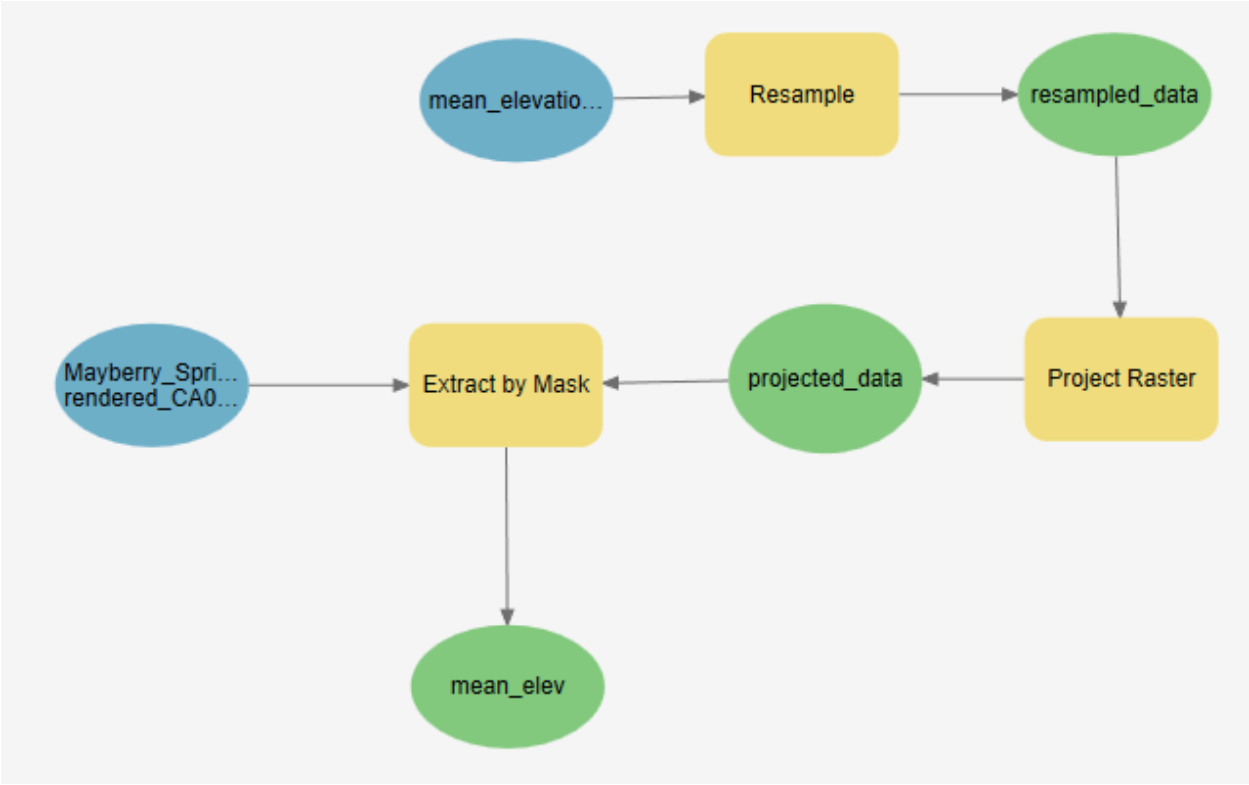

Figure 1. ArcGIS workflow to standardize images

The workflow depicted in this diagram outlines a structured sequence of geospatial data processing steps in ArcGIS Pro, specifically applied to LiDAR statistics Geotiff files for each island. Here, I use the *"mean_elevation.tif"* data as an example. This workflow was applied to all lidar statistics data for every site. The data is first processed with the Resample tool, which adjusts the spatial resolution of the elevation data to match that of the *"Mayberry_Spring NDVI.tif" file*. The file "*mean_elev.tif"* will be used later to compute point-wise correlations. Following resampling, the data undergoes a coordinate system transformation through the Project Raster tool and ensures the elevation data is aligned correctly with the Mayberry Spring NDVI dataset (ESRI).

The adjusted dataset is then passed into the Extract by Mask tool using the Mayberry Spring NDVI file as the mask. This tool selectively extracts areas of interest from the elevation dataset by applying the geographical boundaries of the Mayberry Springs NDVI as a mask. The result is a refined dataset that is resampled, reprojected, and clipped to the areas defined by the NDVI mask. This newly created dataset is ready for further analysis or integration with other geospatial data layers (ESRI).

Exploratory data analysis and feature engineering

I generated basic descriptive statistics in Python, such as means, medians, ranges, and standard deviations, to understand the lidar data's central tendencies and variability. I then visualized these statistics in the form of histograms, box plots, and scatter plots, to discern the distributions and relationships between variables. I also utilized heat maps and correlation matrices to examine the relationships between LiDAR features, such as point density, intensity, and height characteristics, and how they relate to carbon flux. From this information, I then created new variables by transforming and combining existing variables. I employed regression analysis to quantify the relationship between each LiDAR feature and the effective leaf area, and correlation studies helped identify which features are most strongly linked to the effective leaf area. The selection process prioritized features that are not only statistically significant but also provide meaningful ecological insights (Miao et al. 2016).

Dimensionality reduction

Before developing a graph neural network (GNN), I applied dimensionality reduction of the spectral data to reduce the large dimensionality of the data (Morehead et. al 2022). We chose to aggregate data from all three wetlands to increase the diversity of the data and the number of points for training. I clipped the regions to include only those covered by carbon flux measurements. I also applied a mask to roads and water bodies to eliminate potential confounding variables (Mateo-Garcia et. al 2018). These variables can add noise to carbon flux predictions as they are not classified as vegetation. This process was essential in highlighting the most pertinent features of the ecosystem processes, allowing for a streamlined and focused input

into the GNN. The training of the GNN was an iterative process and used a subset of our data while a separate validation set was reserved for performance evaluation.

Model building

I designed a GNN to represent the complex vegetative structures of wetland ecosystems which incorporated data augmentation strategies, effectively increasing the diversity of our data set, and bolstering the model's predictive accuracy and generalization capability. I used max-pooling techniques to distill the data further, reducing dimensionality while preserving the most salient features (Grattarola et. al 2021). The GNN's custom aggregation functions then integrated these features, capturing the intricate relationships within the vegetation and their spectral signatures (Sanchez-Lengeling et al. 2021). This approach was particularly suited to dealing with the dimensionality and variability inherent in the LiDAR and multispectral data, enabling a more nuanced understanding of the ecological processes at play. I also applied an affine transformation to each of the .tif files. For geospatial analysis, features must be accurately extracted from the raster data at the exact places where LiDAR data points exist especially when integrating different data sources. The affine transformation pinpoints the raster pixels that correspond to the LiDAR points and extracts data values from those pixels, ensuring that the data is aligned and accurately represents the geographic layout of the study.

Model training

I have adapted the conventional train/test split to suit the interconnected nature of graph data (Mishra et. al 2021). For node-level tasks where each node was associated with a pixel on a graph, I implemented a train/test split by holding out a subset of nodes. Instead of removing these nodes and their edges, which would disrupt the structural integrity of the graph, I retain them within the network during the training phase. However, their labels are concealed or their features are replaced with non-informative dummy data, such as a "MASK" label that indicates the absence of information (Mishra et. al 2021).

This approach preserves the crucial topological structure of the graph, allowing for uninterrupted information. By maintaining full connectivity, the GNN was trained to understand

and leverage the inherent complexities of the graph, enabling a more accurate representation of the ecological relationships and interactions within the wetlands. This withholding of information allows us to robustly validate the model's performance to unseen data, which was critical for advancing our understanding of wetland vegetation characteristics and dynamics.

Model fine-tuning and statistical analysis

The GNN was trained to effectively reflect different aspects of the wetland vegetation's structural complexity as revealed by our UAV surveys (Law et. al 2001) and to predict corresponding carbon fluxes. I will use the mean squared error function as our loss function to evaluate the strength of our model (Wallach & Goffinet 1989). The fine-tuning process will begin with hyperparameter optimization and will also explore advanced feature engineering techniques to incorporate additional ecological variables that may integrate supplementary sources, such as satellite imagery or in situ sensor data, to capture the intricate dynamics of wetland ecosystems.

RESULTS

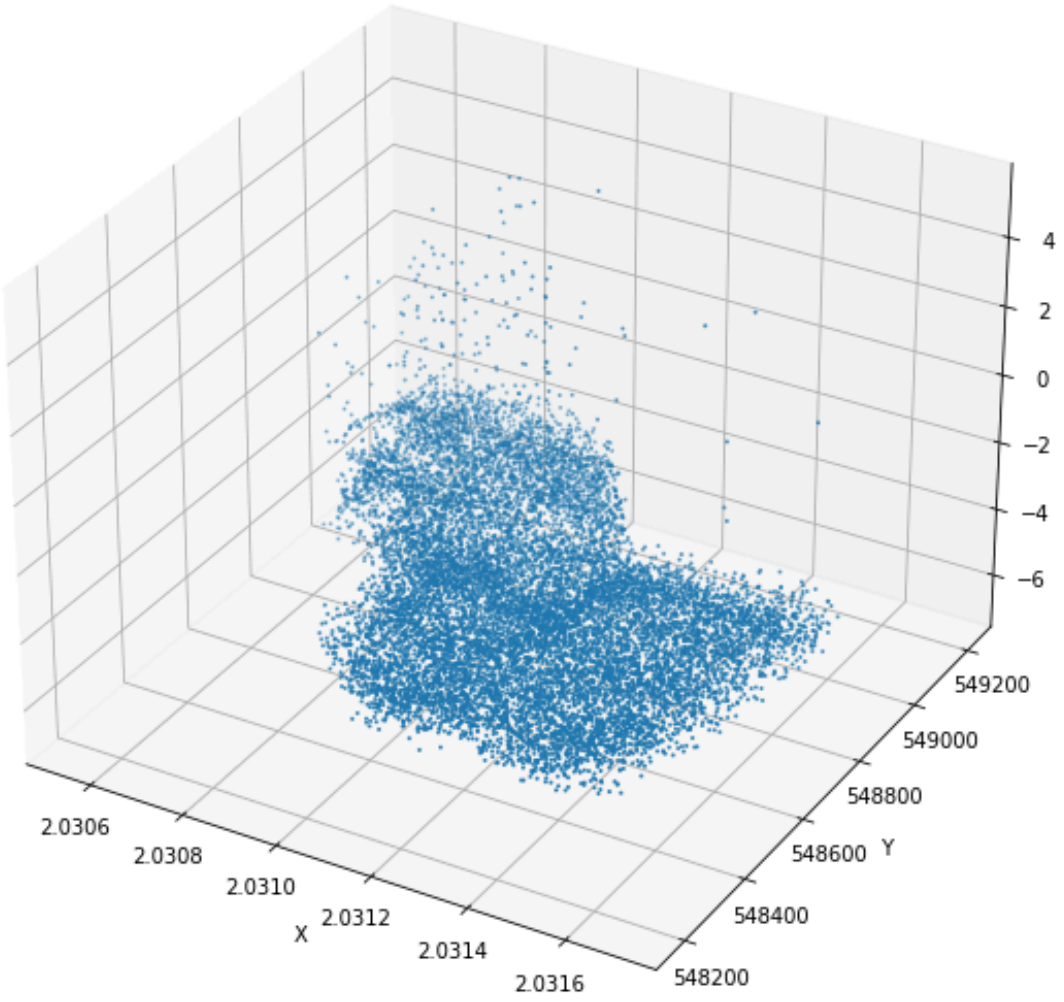

Figure 2. Sample of Twitchell Island Points. This graph shows a plotted systematic sample of 14,000 points (every 10,000). This structure shows the uneven clustered distribution of lidar points in this data.

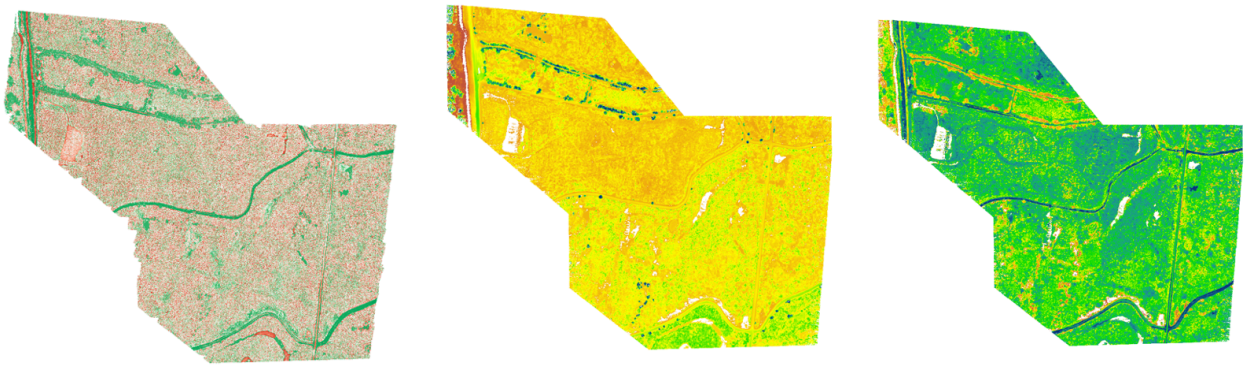

Figure 3. a) Twitchell NDVI, b) Mean Elevation, c) Mean Intensity. Each of the above maps shows different characteristics of Twitchnell Island. (a) NDVI, (b) mean elevation, (c) mean intensity. Twitchell Island shows the strongest relationships. Lower levels of NDVI are shown in red which shows lower vegetation presence while green signifies higher levels of vegetation. In the elevation map, yellows and oranges indicate lower elevation areas, while blues and greens mark higher elevations. In the intensity map, colors like greens and blues indicate that the surfaces reflect more light whereas darker areas might indicate lower reflectance.

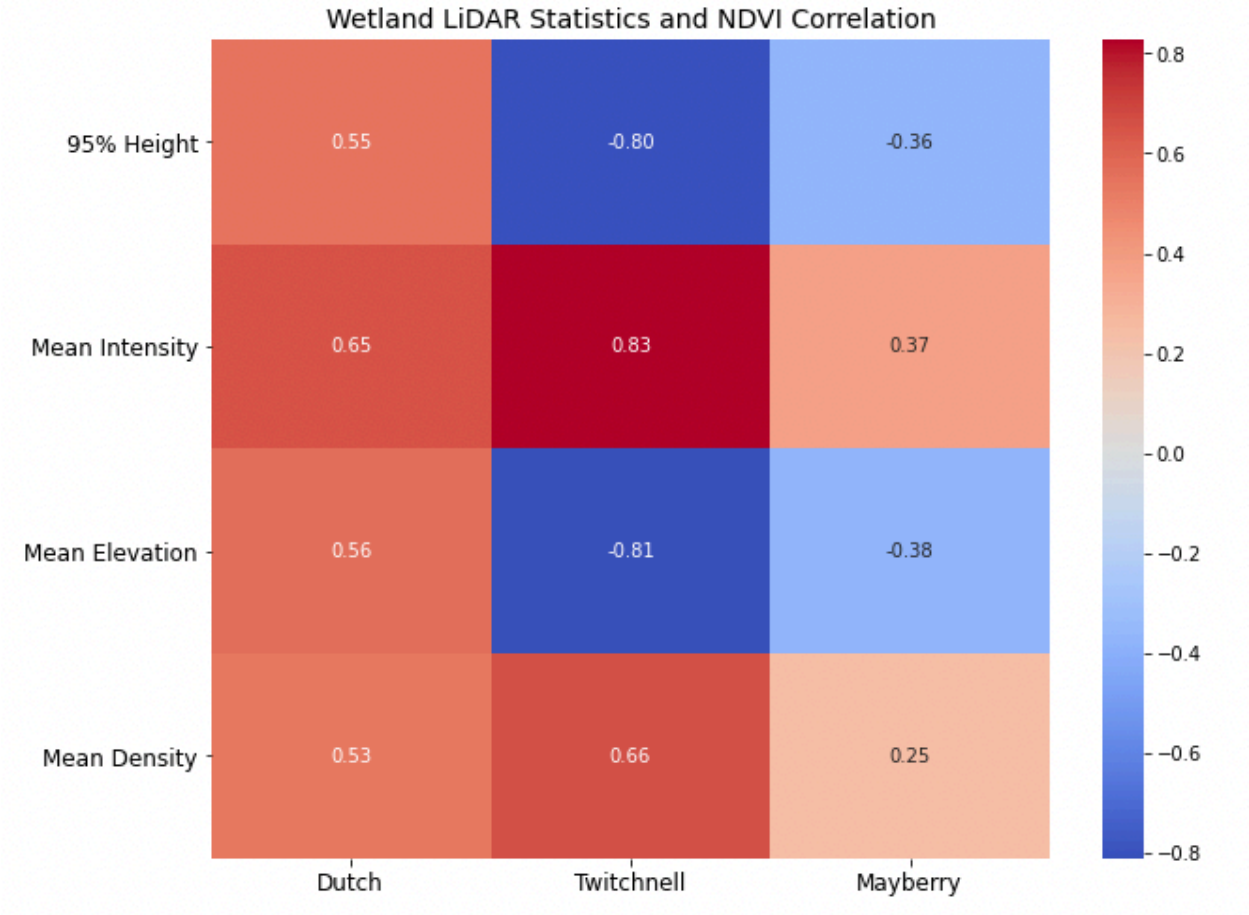

Figure 4. Intrasite Variable Correlations with NDVI. This correlation heat map shows four lidar statistics in mean intensity, 95% height, mean elevation, and mean density over the three different wetland sites of Dutch Slough, Twitchnell Island, and Mayberry Slough and their effects with NDVI.

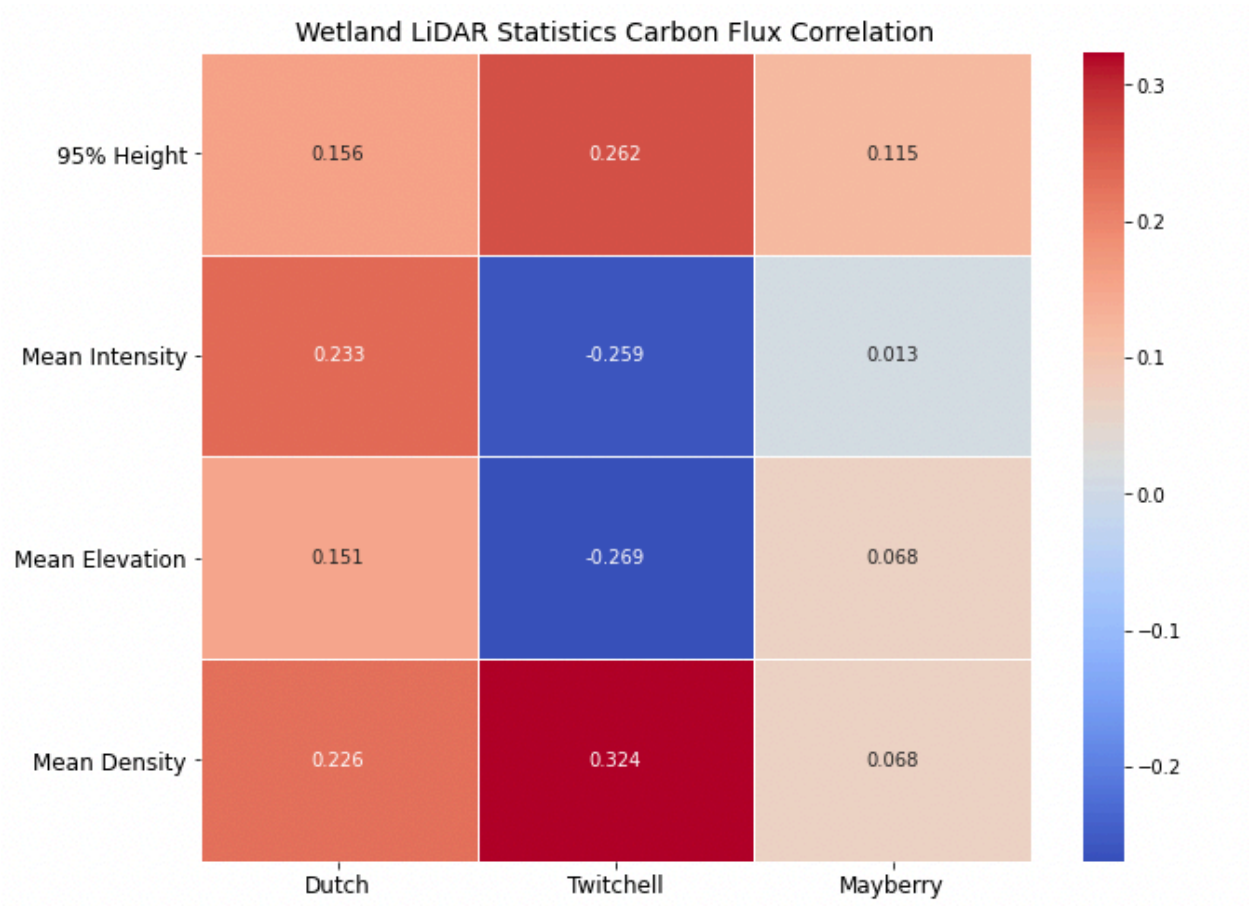

Figure 5. Intrasite Variable Correlations with NDVI. This correlation heat map shows four lidar statistics in mean intensity, 95% height, mean elevation, and mean density over the three different wetland sites of Dutch Slough, Twitchnell Island, and Mayberry Slough and their correlations with carbon flux.

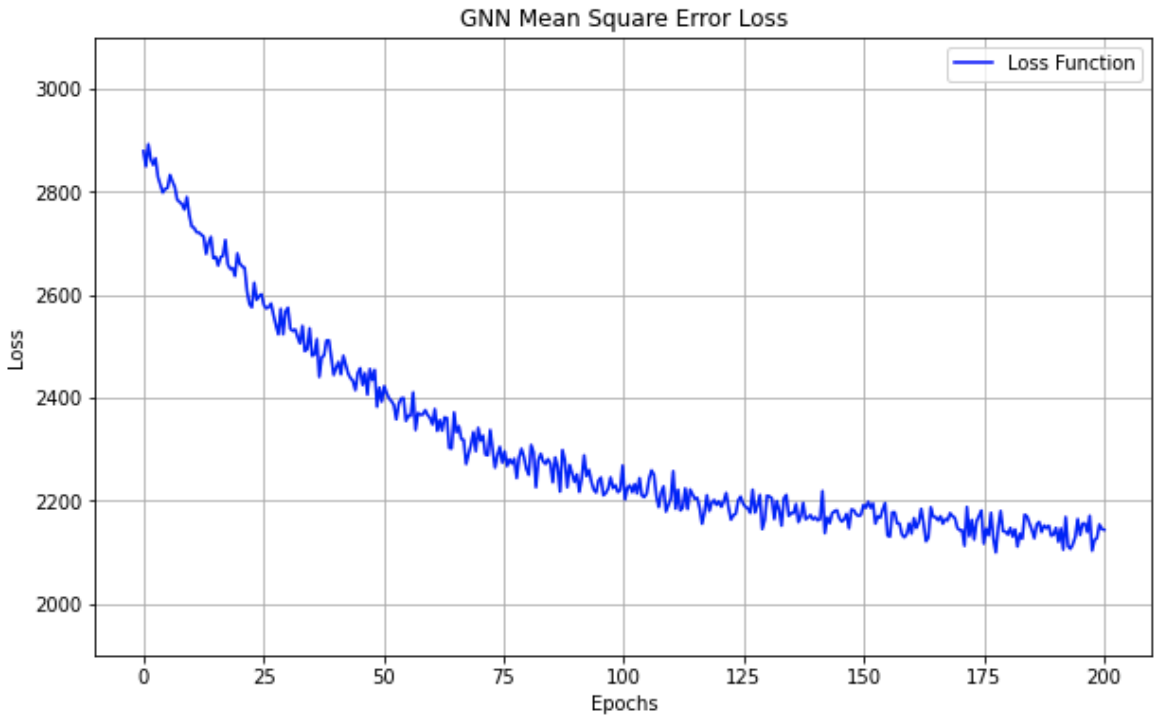

Figure 6. GNN Loss Function. This figure depicts the loss function of the GNN, which uses a scaled-down form of mean squared error from 2900 to 2175 over 200 epochs.

DISCUSSION

This discussion addresses the strategic selection of pixel sizes for data sparsity, minimizing biases in ecological analyses. It also discusses the correlation between NDVI, topography, and intensity, highlighting the strong relationships with Twitchell Island. I then compare LiDAR features across different wetland sites to highlight the variability and commonalities in ecological patterns and the results of the GNN model in predicting carbon flux. Finally, I emphasize the importance of a richer and more diverse dataset for a deeper understanding and effective management of wetland ecosystems and highlight the vast potential this research holds for fast and effective carbon monitoring.

LiDAR data sample visualization

The plot shows data points densely clustered in specific areas, justifying selecting larger pixel sizes for coarser resolution. This approach helps mitigate the impact of data sparsity and

uneven distribution seen across the plot. Such dense clustering indicates that smaller pixel sizes could lead to biased representations, especially in less populated regions. Additionally, the sparse data along the z-axis is characteristic of wetland areas, where even subtle elevation variations are crucial for environmental analyses. This sparsity poses challenges in masking specific areas for detailed studies, as some regions are significantly data-deficient, potentially introducing bias (Jin et. al 2020, Sun et. al 2020). Thus, data was aggregated from all three wetland sites due to ecological similarities across sites, allowing the model to identify common unifying patterns rather than overfitting on specific features that might dominate in one wetland area (Zhang et. al 2020).

Vegetation and topographic influence

The Twitchell Island maps show stronger correlations over the other two wetland sites. I first present the NDVI as a control map, followed by two additional maps of LiDAR statistics. This helps better contextualize the ecological significance of the values observed across the different datasets. Figure 2a) serves as a critical indicator of vegetative health, with NDVI highlighting areas of living green vegetation (USGS). In this visualization, red colors signal regions with scant vegetation, whereas green colors denote dense plant life. The wetlands present a mosaic of red and green, mirroring the complex interactions between aquatic and terrestrial ecosystems. This temporal snapshot also reveals that much of the alfalfa fields are not in peak condition (Tedesco et. al 2022).

Figure 2b) provides a topographic profile of Twitchell Island where yellows and oranges indicate lower elevations and blues and greens signify higher elevations (Hooijer & Vernimmen 2021). This map shows a strong negative correlation of -0.81 between elevation and NDVI, indicating that lower-lying areas typically show higher NDVI levels. The lowest points usually represent wetland areas and the highest points often correspond to trees adjacent to roadways. Although trees in some ecosystems might exhibit higher carbon flux levels than wetlands (Pangala et al., 2012), in this case, the dense alfalfa wetlands display significantly higher carbon flux due to their vegetative density.

Figure 2c) is a mean intensity map of Twitchell Island and shows the distribution of how different areas reflect light (Lang & McCarty 2009). Brighter greens and blues on the map

indicate higher reflectance levels, typically associated with surfaces rich in moisture or having specific material properties (Miguez & Fernández 2012). Conversely, darker yellows and reds suggest lower reflectance, which could indicate drier vegetation, water bodies, or man-made structures. Observing a correlation coefficient of 0.83 confirms this interpretation: areas with brighter greens correspond to higher levels of NDVI which is similarly shown in green hues.

Synthesis of lidar features across wetlands in NDVI

Figure 4 illustrates the correlations among intrasite variables across various wetland sites, revealing a consistent pattern in all but Dutch Slough. This suggests that certain ecosystems share common underlying characteristics, which could be crucial for broader ecological understanding and modeling (Xu et. al 2020). The data from Twitchell and Mayberry, show correlations that may provide generalizable insights into wetland ecology. In contrast, Dutch Slough exhibits an almost uniform distribution across its variables, deviating from the patterns observed in Twitchell and Mayberry. This divergence suggests that relying solely on correlation analysis is not sufficient for discerning broader, generalizable patterns across different wetland environments (Ko et. al 2023). Recognizing these patterns allows for a better understanding of the variance and consistency of wetland features across different sites and can be helpful for the model to be able to recognize (Jafarzadeh et. al 2022).

Synthesis of lidar features across wetlands in carbon flux

Figure 5 highlights moderate correlation values between LiDAR statistics and carbon flux across three wetland areas: Dutch Slough, Twitchell Island, and Mayberry Slough. Correlation coefficients range from -0.3 to 0.3, indicating weak to moderate relationships. Twitchell Island stands out as it exhibits the most significant correlations, suggesting some linear relationship between LiDAR variables and carbon flux, akin to the stronger correlations observed with NDVI compared to the other wetland sites. This might indicate more discernible patterns at Twitchell compared to the other sites. In contrast, Dutch Slough consistently shows positive correlations across all LiDAR attributes, whereas Mayberry Slough demonstrates very weak to negligible correlations, highlighting the variability in environmental structure and

carbon flux across sites. Both of these sites show little to no patterns between NDVI and carbon flux correlations.

This variance suggests challenges in generalizing findings across different wetlands, underlining the complexity of ecosystem interactions. Thus, a GNN approach might be particularly advantageous, as a GNN can capture spatial relationships and leverage proximity data, which could be crucial for identifying complex or non-linear patterns that traditional correlation analyses will miss (Mishra et. al 2021, Zhou et. al 2018). However, before adopting a GNN, enhancing the dataset with additional features or more comprehensive data could be beneficial to provide a stronger dataset foundation (Maharana et. al 2022).

GNN loss function

According to Figure 6, the GNN still exhibits high loss after 200 epochs and struggles to converge to a reasonable loss value, thereby failing to provide accurate predictions on carbon flux data (Zhao et al., 2015). The difficulty in achieving convergence suggests that with the current data and features, it is difficult to learn significant underlying patterns within the data. Thus, expanding the dataset or refining feature selection might be necessary. Furthermore, the current GNN architecture, which selects the 26 nearest points (forming a 3x3x3 cube) to define a neighborhood (Kang et al., 2021), may be biased due to the clustered nature of LiDAR data, which often shows stark density variations as observed in Figure 2. The nearest neighbor approach may disproportionately favor clustered areas and lead to poor performance in sparser regions where the 'nearest' neighbors could be relatively distant and potentially irrelevant to local flux dynamics. Thus, imposing a distance limit on these neighborhood selections can be crucial to improving the model. Additionally, if a point does not have enough neighbors, those points can be considered outliers and excluded or handled differently to avoid skewing the model.

Another solution can be to adapt the GNN architecture to dynamically adjust the number of points included from each area based on the local point density instead of using a fixed 26 (Ye and Ji 2019). Specifically, we can add weighted edges in the graph to represent distances and allow the model to emphasize more significant connections based on proximity (Gong & Cheng 2019). This adjustment could potentially enhance the model’s ability to handle variance and

sparsity of ecological data, leading to more accurate and robust ecological predictions (You et. al 2019).

Limitations and future directions

The model's performance underscores the need for a more expansive and robust dataset to improve model convergence and enhance the accuracy of carbon flux predictions. The dataset should expand to include more temporal snapshots that can reduce the variance in carbon flux measurement. Although the data was averaged over the day for this model, extending the data to span months or years could mitigate the high diurnal variance observed in daily carbon flux measurements. Adding temporal diversity can also help the model capture more of the ecological characteristics of wetlands. Furthermore, the current method of estimating carbon flux—utilizing pointwise data from a flux tower, extrapolated using wind vectors and interpolation techniques—may introduce biases that compromise the model's predictive capabilities. Although the Kiljun method for carbon flux estimation is robust, it still relies on interpolations that can be problematic if the underlying assumptions are uncertain. Another concern is the dataset's composition, which predominantly captures a phase of deceased alfalfa vegetation. This could risk biasing the model's output by disproportionately representing dead vegetation, potentially skewing the results toward these conditions. Thus, the dataset should expand to more diverse vegetative types and states, increasing the model's structural complexity. By incorporating more features and environmental variables, the dataset will better represent and reflect the diverse dynamics within wetland ecosystems.

Broader implications and conclusion

Our research contextualizes observed patterns within a broader ecological framework, highlighting the potential of LiDAR technology for environmental management. As more comprehensive data becomes available, LiDAR could potentially supplant the need for expensive flux towers that require continuous monitoring. This shift would significantly expand the scope and efficiency of LiDAR-based studies. In areas where constructing flux towers is impractical, LiDAR offers an effective alternative, facilitating rapid and extensive environmental surveys

without the logistical challenges and time constraints. With access to expanded remote sensing data, measurements of additional environmental variables, and a larger dataset of wetlands, this research could be invaluable in monitoring the carbon source and sink dynamics of these ecologically sensitive areas. Furthermore, the implications of this research could extend beyond wetlands to all ecosystems, making LiDAR a key tool in holding communities and nations accountable for their carbon emissions and environmental commitments.

ACKNOWLEDGMENTS

I would like to extend my deepest gratitude to Professor Dennis Baldocchi and the biometeorology lab for their invaluable support throughout my research. Professor Baldocchi's generosity in sharing his data and insights into carbon footprint analysis has been instrumental to my study. Special thanks are also due to Joseph Verfaillie for his patience and expertise, which greatly enhanced my understanding and processing of the complex data. I am also grateful to Patina Mendez for her exemplary mentorship in the ESPM thesis class. Her guidance has been a cornerstone of my academic journey, helping to shape a constructive and engaging research environment.

REFERENCES

Baldocchi, D. 2014. Measuring fluxes of trace gases and energy between ecosystems and the atmosphere – the state and future of the eddy covariance method. Global Change Biology 20:3600–3609.

Bauer, P., and M. Nagy. 2024. Flight-Data-Based High-Fidelity System Identification of DJI M600 Pro Hexacopter. Aerospace 11:79.

Bian, L., A. M. Melesse, A. S. Leon, V. Verma, and Z. Yin. 2021. A Deterministic Topographic Wetland Index Based on LiDAR-Derived DEM for Delineating Open-Water Wetlands. Water 13:2487.

Buffington, K. J., C. N. Janousek, B. D. Dugger, J. C. Callaway, L. M. Schile-Beers, E. Borgnis Sloane, and K. M. Thorne. 2021. Incorporation of uncertainty to improve projections of tidal wetland elevation and carbon accumulation with sea-level rise. PLoS ONE 16:e0256707.

Burba, G. 2013. Eddy Covariance Method for Scientific, Industrial, Agricultural and Regulatory Applications: A Field Book on Measuring Ecosystem Gas Exchange and Areal Emission Rates.

Dai, W., S. Chen, Z. Huang, Y. Xu, and D. Kong. 2022. LiDAR Intensity Completion: Fully Exploiting the Message from LiDAR Sensors. Sensors (Basel, Switzerland) 22:7533.

Extract by Mask (Spatial Analyst)—ArcGIS Pro | Documentation. 2024. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/extract-by-mask.htm.

Fan, P., W. Li, X. Cui, and M. Lu. 2019. Precise and Robust RTK-GNSS Positioning in Urban Environments with Dual-Antenna Configuration. Sensors (Basel, Switzerland) 19:3586.

Ground Control Points | U.S. Geological Survey. 2021. https://www.usgs.gov/landsat-missions/ground-control-points.

Hu, W., Z. Lu, F. Meng, X. Li, R. Cong, T. Ren, T. D. Sharkey, and J. Lu. 2020. The reduction in leaf area precedes that in photosynthesis under potassium deficiency: the importance of leaf anatomy. New Phytologist 227:1749–1763.

Huang, H., A. Yang, Y. Tang, J. Zhuang, C. Hou, Z. Tan, S. Dananjayan, Y. He, Q. Guo, and S. Luo. 2021. Deep color calibration for UAV imagery in crop monitoring using semantic style transfer with local to global attention. International Journal of Applied Earth Observation and Geoinformation 104:102590.

Inertial Explorer GNSS+INS post-processing. 2024. https://novatel.com/products/waypoint-post-processing-software/inertial-explorer.

Jafarzadeh, H., M. Mahdianpari, and E. Gill. 2022. Wetland Mapping by Jointly Use of Convolutional Neural Network and Graph Convolutional Network. Pages 2219–2222 IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium.

Jin, W., Y. Ma, X. Liu, X. Tang, S. Wang, and J. Tang. 2020, June 27. Graph Structure Learning for Robust Graph Neural Networks. arXiv.

Jones, C., W. A. Christens-Barry, M. Terras, M. B. Toth, and A. Gibson. 2020. Affine registration of multispectral images of historical documents for optimized feature recovery. Digital Scholarship in the Humanities 35:587–600.

Kljun, N., P. Calanca, M. W. Rotach, and H. P. Schmid. 2015. A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP). Geoscientific Model Development 8:3695–3713.

Ko, H., I. Praca, and S. G. Choi. 2024. Anomaly detection analysis based on correlation of features in graph neural network. Multimedia Tools and Applications 83:25487–25501.

Lang, M. 2014. Light Detection and Ranging (LiDAR) for Improved Mapping of Wetland Resources and Assessment of Wetland Conservation Practices.

Li, L., J. Weiner, D. Zhou, Y. Huang, and L. Sheng. 2013. Initial density affects biomass–density and allometric relationships in self-thinning populations of Fagopyrum esculentum. Journal of Ecology 101:475–483.

Maharana, K., S. Mondal, and B. Nemade. 2022. A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings 3:91–99.

Mateo-García, G., L. Gómez-Chova, J. Amorós-López, J. Muñoz-Marí, and G. Camps-Valls. 2018. Multitemporal Cloud Masking in the Google Earth Engine. Remote Sensing 10:1079.

Miao, J., and L. Niu. 2016. A Survey on Feature Selection. Procedia Computer Science 91:919–926.

Mishra, P., Piktus, A., Goossen, G., & Silvestri, F. 2021. Node Masking: Making Graph Neural Networks Generalize and Scale Better. arXiv.

Morehead, A., W. Chantapakul, and J. Cheng. 2022, March 23. Semi-Supervised Graph Learning Meets Dimensionality Reduction. arXiv.

Nahlik, A. M., and M. S. Fennessy. 2016. Carbon storage in US wetlands. Nature Communications 7:13835.

Pangala, S. R., S. Moore, E. R. C. Hornibrook, and V. Gauci. 2013. Trees are major conduits for methane egress from tropical forested wetlands. New Phytologist 197:524–531.

Petras, V., A. Petrasova, J. B. McCarter, H. Mitasova, and R. K. Meentemeyer. 2023. Point Density Variations in Airborne Lidar Point Clouds. Sensors (Basel, Switzerland) 23:1593.

Remote Sensing | Remote Sensing on Alfalfa as an Approach to Optimize Production Outcomes: A Review of Evidence and Directions for Future Assessments. 2022. https://www.mdpi.com/2072-4292/14/19/4940.

Rey-Sanchez, C., A. Arias-Ortiz, K. Kasak, H. Chu, D. Szutu, J. Verfaillie, and D. Baldocchi. 2022. Detecting Hot Spots of Methane Flux Using Footprint-Weighted Flux Maps. Journal of Geophysical Research: Biogeosciences 127:e2022JG006977.

Roussel, J.-R., D. A. (Reviews the documentation), F. D. B. (Fixed bugs and improved catalog features), A. S. M. (Implemented wing2015() for segment_snags()), B. J.-F.

(Contributed to R. for track_sensor()), G. D. (Implemented G. for track_sensor()), L. S. (Contributed to parallelization management), S. A. (Author of the C. concaveman code), and B. S.-O. (Author of the "chm_prep"function). 2024, March 5. lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications.

Shao, Y., Q. Zhu, Z. Feng, L. Sun, X. Yang, X. Li, A. Wang, F. Yang, and H. Ji. 2024. Temporal and Spatial Assessment of Carbon Flux Dynamics: Evaluating Emissions and Sequestration in the Three Northern Protection Forest Project Areas Supported by Google Earth Engine. Remote Sensing 16:777.

Sharma, L. K., and R. Naik. 2023. LiDAR (Light Detection and Ranging) for Wetland Study. Pages 1–12 International Encyclopedia of Geography. John Wiley & Sons, Ltd.

Sun, K., Z. Lin, and Z. Zhu. 2020. Multi-Stage Self-Supervised Learning for Graph Convolutional Networks on Graphs with Few Labeled Nodes. Proceedings of the AAAI Conference on Artificial Intelligence 34:5892–5899.

Tedesco, D., L. Nieto, C. Hernández, J. F. Rybecky, D. Min, A. Sharda, K. J. Hamilton, and I. A. Ciampitti. 2022. Remote Sensing on Alfalfa as an Approach to Optimize Production Outcomes: A Review of Evidence and Directions for Future Assessments. Remote Sensing 14:4940.

The Basics of LiDAR - Light Detection and Ranging - Remote Sensing | NSF NEON | Open Data to Understand our Ecosystems. 2024. https://www.neonscience.org/resources/learning-hub/tutorials/lidar-basics.

Tocci, F., S. Figorilli, S. Vasta, S. Violino, F. Pallottino, L. Ortenzi, and C. Costa. 2022. Advantages in Using Colour Calibration for Orthophoto Reconstruction. Sensors (Basel, Switzerland) 22:6490.

Triglav Cekada, M., F. Crosilla, and M. Fras. 2010. Theoretical LiDAR point density for topographic mapping in the largest scales. Geodetski Vestnik 54:403–416.

US Department of Commerce, N. O. and A. A. 2024. NCN Homepage - National Geodetic Survey. https://geodesy.noaa.gov/CORS/.

Valach, A. C., K. Kasak, K. S. Hemes, T. L. Anthony, I. Dronova, S. Taddeo, W. L. Silver, D. Szutu, J. Verfaillie, and D. D. Baldocchi. 2021. Productive wetlands restored for carbon sequestration quickly become net CO2 sinks with site-level factors driving uptake variability. PLoS ONE 16:e0248398.

Wallach, D., and B. Goffinet. 1989. Mean squared error of prediction as a criterion for evaluating and comparing system models. Ecological Modelling 44:299–306.

Wang, Y., and H. Fang. 2020. Estimation of LAI with the LiDAR Technology: A Review. Remote Sensing 12:3457.

Xu, X., M. Chen, G. Yang, B. Jiang, and J. Zhang. 2020. Wetland ecosystem services research: A critical review. Global Ecology and Conservation 22:e01027.

Yuan, W., J. Li, M. Bhatta, Y. Shi, P. S. Baenziger, and Y. Ge. 2018. Wheat Height Estimation Using LiDAR in Comparison to Ultrasonic Sensor and UAS. Sensors (Basel, Switzerland) 18:3731.

Zayed, A. M. A., A. Saber, M. A. Hamama, M. Rabah, and A. Zaki. 2023. Evaluation of vertical accuracy of TanDEM-X Digital Elevation Model in Egypt. The Egyptian Journal of Remote Sensing and Space Sciences 26:919–936.

Zhang, H., D. Yang, E. Yurtsever, K. A. Redmill, and Ü. Özgüner. 2021a, March 27. Faraway-Frustum: Dealing with Lidar Sparsity for 3D Object Detection using Fusion. arXiv.

Zhang, Y., J. Yan, X. Cheng, and X. He. 2021b. Wetland Changes and Their Relation to Climate Change in the Pumqu Basin, Tibetan Plateau. International Journal of Environmental Research and Public Health 18:2682.

Zhang, Z., W. Jiang, K. Peng, Z. Wu, Z. Ling, and Z. Li. 2023. Assessment of the impact of wetland changes on carbon storage in coastal urban agglomerations from 1990 to 2035 in support of SDG15.1. Science of The Total Environment 877:162824.

Zhou, J., G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun. 2020. Graph neural networks: A review of methods and applications. AI Open 1:57–81.