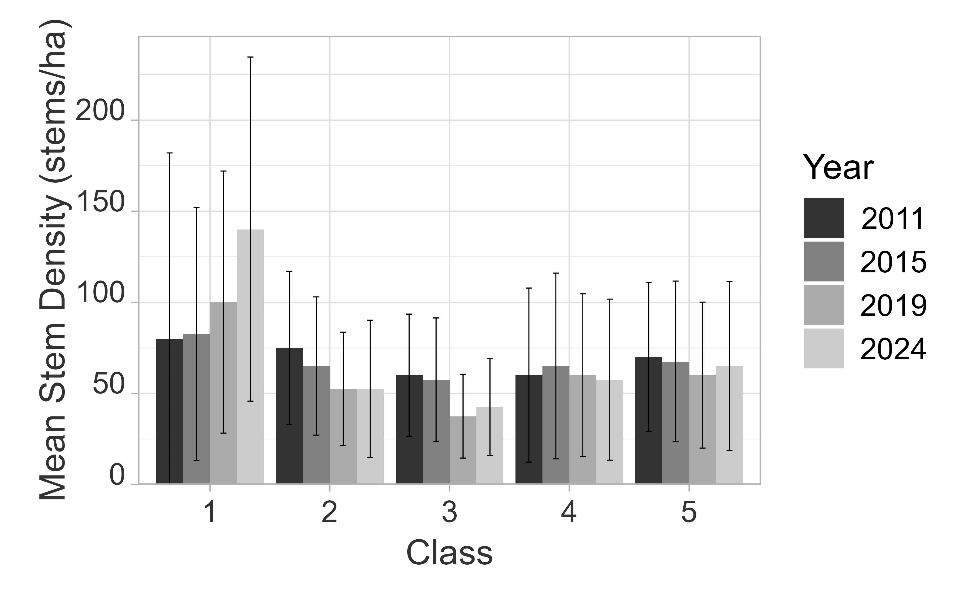
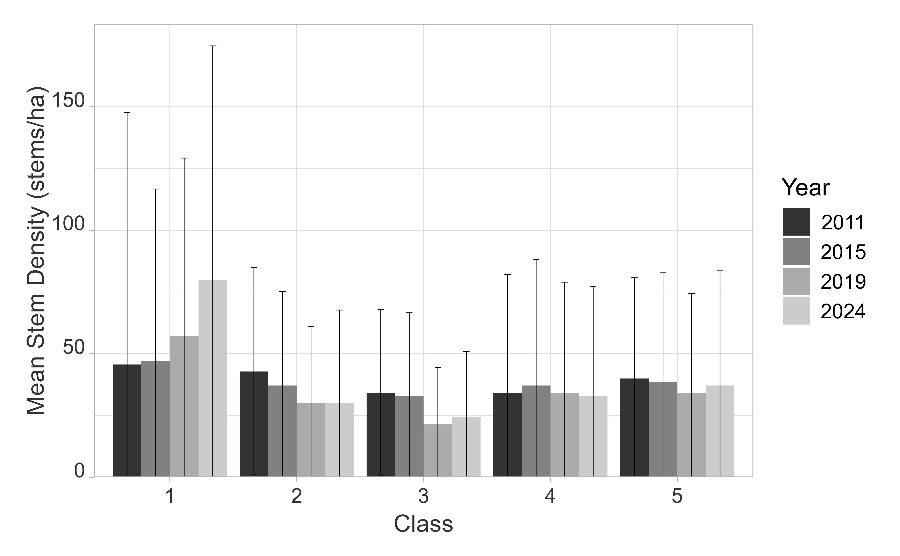
Notes on updating Vegmon overstory code from Kate Miller’s Comments

In April 2025, Kate Miller provided comments on Vegmon overstory code that helped to develop OverstoryRegenCanopyV2.5.Rmd. Her complete comments (HTLN\_veg\_code\_revew) and a version of the R script she coded (Overstory\_Regen\_Canopy\_v2.2\_kmm.Rmd) are located in dev folder in the overstory R folder.

NOTE: The changed made in OverstoryRegenCanopyV2.5.Rmd altered the confidence interval values previously calculated because calculations were shifted to bootstrap (sites>=7) or max/min values (sites < 7). One mean value was altered by the improved code: mean overstory density by size class. The previous code did not correct this computation to divide by the total number of sites even if a given size class was not present. This script has calculated correct values and those values were double checked. Values of all other metrics were also double checked for this error, but none were found. This changed some values for the in the following outputs: DensityClass.csv, DensityClassD.csv, DensityClass.jpg, and DensityClass.jpg for the 2024 LIBO report (no impacts to HOSP report). This did not impact the total average density by year value.

Updated LIBO density by class



Previous LIBO density by class

Almost all of Kate’s comments were adopted and used to improve the script. A few suggestions were not used or modified to fit project/reporting needs.

1. Changed names of output tables and graphs to camel-case.
2. “Year” was removed from params at the beginning of the document. There was no use of this “Year” value apart from helping to identify the document. It was not used again in the code and removed to reduce confusion.
3. Kate suggested using data directly from the Access database and doing preliminary cleaning in R. After discussion with project lead, we will not be using this method and cleaning will continue in excel. Going forward, it will be important to note that columns from the import csv files will need to remain consistent and the data files will have to be saved to specific naming conventions (i.e. HOSPoverstory.csv, HOSPregeneration.csv, and HOSPcanopy.csv). Notes of these requirements are located in the README file in the overstory R folder.
4. Several constants that Kate created for BA/density analysis (a dataframe with all combos of site and year, a dataframe with all combos of year and class, and setting the number of replications for bootstrapping) were moved into the setup chunk to remove redundancies later in data.
5. For Basal Area (live and dead), Density (live and dead), Canopy, and Regeneration analysis if/else statements were added to use Kate’s bootstrapping methods of calculating confidence interval if the data has >= 7 sites and to use max and min values as Confidence interval if <7 sites. This was based on comments from Kate in the document (HTLN\_veg\_code\_review\_20250328.docx) and over email.
6. We implemented Kate’s suggestion to use line graphs for year data and remove even spacing from graphs where year is on the X axis. A few minor changes were made to ggplot graphs: theme was switched to theme\_minimal(), when year was plotted on the x-axis breaks and tick marks were marked evenly through time and not just on sampling years, the x-axis line was forced to 0, margins were added, and dpi was set to 600. These changes did not impact the data.
7. Graphs of dead basal area and density were extended to show the same y-axis extent as the graphs of live basal area and density. This was requested for reporting.
8. For overstory species analysis, there was a situation where first/current year columns swapped titles but subtracted correctly. I changed this to improve readability but it did not change the values calculated.
9. Added common names to occurrence table
10. Bootstrapping was not used to calculate confidence interval for graph displaying mean canopy by individual site. After discussing with project lead, this graph did not require any CI bars.
11. Regeneration correction factor was changed from x10 to x100 because regen in measured in plots.
12. For the regeneration tables, made slight alterations so each species only appeared in one row in the wider pivot table. Calculations were all correct but table was pivoted wider but still took three rows. Also added a check to ensure that regeneration graphs only displayed changed for species that were present in that class in at least one year, even is change was 0.