

Trends in Forest Ecology and Succession at Howland Forest

Jamis M. Bruening, Ph.D. Student

MnM4SDS Final Project Paper
Instructor: Dr. Taylor Oshan
Department of Geographical Sciences,
University of Maryland

December 16, 2019

Contents

1	Introduction	2
1.1	Background	2
1.2	Objectives	3
2	Literature Review	4
2.1	Ecological basis for methods chosen	5
3	Data	6
4	Methods	7
4.1	Plotting and spatial visualization of variables	7
4.2	Exploratory modeling of successional status	7
4.3	ESDA and GWR	9
5	Results	10
6	Discussion and Conclusions	11
7	Figures	14
7.1	Exploratory Plotting	14
7.2	OLS and Fixed Effects Regression	18
7.3	ESDA	20
7.4	GWR and MGWR	23
8	References	27

1 Introduction

1.1 Background

This project leverages a data set of tree locations and their associated physical attributes to identify and explain ecological trends related to forest succession at Howland Experimental Forest, located in central Maine. Howland Forest (HF) is a secondary forest, meaning it was previously subject to some type of human land use, but at some point in the past that usage changed; the land was no longer ‘used’ by humans and over time it returned to a more natural state (Hurtt et al., 2011). An example of this concept is a secondary forest that establishes and regrows on an abandoned agricultural field, as happened throughout New England in the early 20th century Foster et al. (1998); Canadell et al. (2007). This concept of secondary lands is central to my dissertation research. Pioneered by UMD GEOG’s Dr. George Hurtt, decades of work have shown that the rise of these secondary lands have contributed significantly to the global land carbon sink, and specifically in the northeast United States (Hurtt et al., 2002, 2006; Albani et al., 2006; Houghton et al., 2012; Houghton and Nassikas, 2017).

The world’s leading researchers in the global carbon cycle and terrestrial land fluxes argue that about two-thirds of the ‘missing carbon sink’ (the portion of carbon missing from the atmosphere that cannot be accounted for in top-down global carbon budgeting and thus must have been fixed on the terrestrial land surface) is located in the mid-latitudes, due in large part to the regrowth of secondary forests from bare ground (Schimel et al., 2015; Houghton et al., 2018). My dissertation research is focused on these secondary forested lands in the northeast United States, specifically the future carbon gain and trajectory of above-ground biomass accumulation with continued forest regrowth and natural secondary succession. A central component to this area of study is the age distribution of all the trees in a forest, which is in turn coupled with the succession status and species composition of said forest. As a tree ages it gets larger, and stores an increasing amount of carbon. My interests are in using LiDAR to predict the age of a forest from information about its vertical and horizontal structure, and this estimate the

future carbon gain through continued forest regrowth, across northern New England. To do so, I need a better understanding of forest succession in forests characteristic in the region, and thus I have dedicated this project to interrogating a spatial ecological data set to identify trends in species distribution and forest succession at a site in central Maine. Throughout this course I was exposed to various methods in spatial data science, and I have applied some of them here in an ecological context. As much of the spatial data science community is directed towards answering spatial questions in different fields such as urban dynamics, econometrics, demography and geopolitics, there is substantial room for new applications of these methods in environmental and ecological contexts. This project attempts identify and explain spatial trends in forest succession using cutting edge methods in spatial data science.

1.2 Objectives

Over the course of the semester my project's objectives changed slightly as I worked through the analysis, encountered and addressed problems, and learned new methods. Outlined here are my end-objectives, as were possible in a relatively short amount of time to perform the analysis. When possible I have preserved the words verbatim from my initial project proposal.

1. Use exploratory techniques to identify spatial trends in ecological variables
2. Employ spatial data sciences methods to explain the presence and drivers to ecological trends
3. *Determine the relationship between spatial ecological trends and additional environmental (site) variables, such as temperature, soil moisture, topography.*

Number 3 is italicized because I did not have time to address it given the time constraints on the project.

2 Literature Review

I applied several methods of analysis from this course, based on the specifics of my project's data source and research questions. Some methods and concepts I had been exposed to in previous courses, but I gained a better practical knowledge of them here. The prime example is of spatial autocorrelation (or association), and specifically local indicators (LISA) (Anselin, 1995; Symanzik, 2014). In two previous courses I've learned about spatial association, and the idea of spatial autocorrelation, however in this class I gained a deeper understanding and practical knowledge of the local indicators of these relationships. In his seminal paper on the topic, Anselin (1995) essentially created the field of exploratory spatial data analysis and local spatial associations. This work provided several novel and ground breaking methods to quantitatively assess the extent to which spatial associations exist with a spatial data set, a numerical way to represent Tobler's first Law of Geography. These methods have are used to identify the extent to which 'hotspots' of nonstationarity exist, a measure of how similar a value at one location is to over nearby observations, and the extent to which spatial position influences relationships between processes or variables.

Continuing in this vein and at around the same time, Brunsdon et al. (1996) developed a method of exploring nonstationarity through what they termed as geographically weighted regression (GWR). This method addressed a prominent issue in quantitative geography at the time, that in many instances, a 'global' set of relationships that is identified through aspatial methods may not hold true across and throughout space. This new GWR technique assumes that spatial position is integral to the process being studied, and the relationship between variables may not hold true everywhere throughout the study area. They thus developed a type of linear regression model that accounts for the spatial position of observations, and fits a set of local models for each observation. This technique allows for different coefficients for variables for each observation, thus allowing the model to pick up on spatial variations between response and predictors that a regular, global regression model inherently cannot.

These methods have been further developed over the past 2+ decades, with new

implementations of these concepts for different processing systems or scripting languages, as well as the addition of new functionality and methods. Multiscale Geographically Weighted Regression (MGWR) is one such example (Fotheringham et al., 2017; Oshan et al., 2019). This method builds on the concepts of GWR, but even further reduces some of the spatial modeling assumptions. Similar to how GWR rejects the assumption that a single relationship must hold true across a study area or region and allows for different model coefficients based on local regression fits, MGWR does not assume that the relationships between each predictor-response variable pair operates on the same spatial scale. This method allows for spatial nonstationarity in the same way that GWR does, but also lets each predictor variable's spatial domain vary relative to the other predictor variables (Fotheringham et al., 2017).

2.1 Ecological basis for methods chosen

I chose these above methods based on my understanding of the landscape scale processes that drive ecological succession, and an understanding that the environmental variables that influence tree growth and succession may have different spatial scales on which they operate. For example, gap regeneration theory and shade tolerance are foundational pillars of forest succession theory (Runkle, 1981; Shugart et al., 1984; Yamamoto, 1992). Yet it is also known that growing season climate and edaphic conditions are also crucially important to understanding forest growth and large scale ecological succession (West et al., 2012). A baseline knowledge of landscape ecology necessitates that these processes must operate on extremely different spatial scales, from the very local, regarding gap theory and shade tolerance, to site-based and even regional scales for climate and soil characteristics. Thus I wanted to incorporate methods that allowed for the spatial variation of relationships between individual tree canopy position and the physiological and environmental characteristics of this site.

3 Data

The dataset I used represents every single live tree in the NASA Plot at HF. The NASA plot is a 200 meter by 150 meter plot within HF established in 1989 for long term ecological monitoring. Thus, the spatial positions of every single tree were mapped, and have been periodically updated as new trees establish and old trees experience mortality. I was not looking at temporal trends in this analysis, so I subset the NASA Plot data to include only live trees at the most recent update in 2015, as well as only the most recent tree attributes measured in 2015. My resultant data frame includes all trees with a diameter at breast height (DBH) greater than 5 cm within the plot, the coordinate locations of each tree, along with various physical attributes. These include the DBH, tree height, distance between the base and the live crown, and its canopy class, which is a proxy for its successional status. This variable is the dependent variable in my analysis, and is my primary interest in this project. It is technically an ordinal variable, meaning it is a categorical variable in which the categories are ordered, and describes the position of the tree and its canopy in the forest relative to the neighboring trees. There are six levels to this variable; 1: Super dominant, 2: Dominant, 3: Co-dominant, 4: Intermediate, 5: Overtopped, 6: Suppressed. However, in ecology nothing is cleanly separated into distinct bins, so I've relaxed the assumption that this is a strict ordinal variable and have treated it as a continuous variable to be able to use many of the methods we learned in class. I can justify this, because everything is a gradient in ecology, and almost never are complex organisms such as trees naturally separated into rigid classes based on their physical properties. For example, two trees could both be super-dominant, meaning their canopies extent up beyond most other neighboring trees, yet one could be slightly taller than the other and receive slightly more sunlight than the other. In this case while they are both classified as super dominant, one is still more super-dominant than the other. Thus, I treat the canopy position variable as continuous based on this rationale, and I do not expect this choice to have substantially impacted my results or violated any other assumptions downstream in my analysis.

4 Methods

4.1 Plotting and spatial visualization of variables

My first approach was to plot the tree locations in 2-dimensional space, adding a third and fourth dimension to visualize my dependent and independent variables at the same time, using colors and graduated symbol sizes to represent the physical tree attributes. The Seaborn module is extremely functional for this, and allowed me to produce maps that gave me a very useful visualization of the complex picture of the ecological succession at HF Adding a third and fourth dimension to the spatial plots of the tree locations was especially helpful to start visually identifying spatial trends in these variables, and informed my methods selection later in explicitly quantifying these complex spatial differences. See section 7 for all plots. I chose to put all figures in a separate section at the end to allow for large figure sizes. If I included the figures in-text, I would need to shrink many of them which would reduce their readability. Each plot will also be linked to individually when referenced in the results or discussion section.

Following regular data analysis procedures, I also produced aspatial plots, both in this preliminary exploratory stage to identify aspatial trends in my dependent and independent variables, but also in interpreting the results of the variables models run later in the analysis. Plotting residuals in space and also by canopy type was crucial to my analysis, and provided my initial motivation to include a spatial fixed effect model.

4.2 Exploratory modeling of successional status

In plotting the spatial locations of the trees and their associated physical attributes, it became apparent that there was a high degree of something I call ecological noise (this is a non-technical term). Ecology is complex and messy, and sometimes an effective technique to reduce this noise and identify underlying processes or trends is to slightly aggregate data to visually identify a signal that is present. Thus, I delineated a 5 meter by 5 meter grid, overlaid on the NASA plot, to use both to segment the plot into regularly sized areas, and to aggregate the attributes of the trees within each grid cell to smooth out

the ‘ecological noise’. The size of the grid here is important; too large a grid cell and I will start to smooth out an underlying signal, and too small a grid cell and I will not reduce any noise and have many grid cells without any trees in them. This is a somewhat arbitrary decision, based in part on my intuition and moderate knowledge of forested systems. In a maturing forest, a 25 meter squared area may have a handful of trees, depending on their size and succession status, and this would be enough to reduce the noise levels present in the individual trees and bring out any underlying ecological signals present. To help represent the ecological characteristics of the NASA plot at this new resolution of 5 meters, I aggregated the physical tree attribute data at this resolution, as well as calculated several additional ecological variables for each 5 meter grid cell in the plot. The resultant gridded variables are as follows:

- Average height, average diameter (mean of all trees in cell)
- Mode successional status (most common successional status of all trees in cell)
- Species richness (number of different tree species in each cell)
- Tree density (total number of trees in each cell)
- Heterogeneity Index (species richness divided by total number of species)

With the plotting and data munging completed, I transitioned to some initial, basic modeling. I ran a set of two initial models, an aspatial OLS model and an OLS with spatial fixed effects by including the gridded transect variable as a grouping variable. In both models (and all others in the analysis) my dependent variable is the canopy position of each tree, and the independent tree-attribute variables are tree diameter (DBH), height (HT), and distance between the base and live crown (BLC), explained above in section 3. I chose this initial two-model approach because it allowed me to partition the influence of aspatial physical attributes of the tree and the spatial component of ecological succession. By comparing the results between these two models, I was able to identify the presence of a spatial trend in addition to the explanatory power of both DBH and HT in the

successional status of each tree. These models were run at the individual tree level, with the gridded transect variable added to represent space in the fixed effects model.

4.3 ESDA and GWR

Following the spatial fixed model, I ran a LISA analysis on each of the main variables in my study, aggregated at the 5 meter grid level, as well as several other variables I calculated to help represent the characteristics of the NASA plot at the new resolution of 5 meters. I used queen contiguity weights, a decision that was again informed by ecological theory. There is no *a priori* distance at which ecological processes operate in a forest, and arguably different processes happen at different scales (see section 2.1), so this eliminated a distance-based weights matrix. In the choices of contiguity based weights, queen weights make the most sense ecologically. The square grid I generated has now ecological basis other than an arbitrary way to segment my plot into smaller regions. There is no ecological process that happens in orthogonal space (rook or bishop based weights); the most appropriate scheme is to consider all neighboring areas similarly examining ecological trends, thus necessitating a queen contiguity based weighting scheme.

For each variable (physical tree-attributes such as DBH and HT and locale-specific attributes such as species richness or tree density) I ran a LISA plot analysis to identify patterns of spatial dependence on the variables. My rationale for doing this was to explore the potential for any of these variables to match the effect that seen in the spatial fixed effects model, in addition to better understand the spatial-ecological trends in each independent variable, and how these may relate to the ecological succession of the trees in the NASA plot. I tested the significance of Moran's I for each variable, and plotted each variable in space aggregated at the 5 meter grid level.

In a further attempt to assess the extent to which any of my tree or grid level variables capture the effects of the grid variable seen in the spatial fixed effects model, I ran both a GWR and MGWR. The rationale for this step was that I could envision a scenario where the drivers of succession vary based on the location within a forest. Land-use history

is such an important driver of ecological succession, and I do not have any variables that provide any such information. If part of this plot was disturbed more recently than another, the effects on the successional status of the trees there would almost certainly be present, and not captured under the assumption that the relationship between the independent variables and crown positions the same regardless of position within the plot.

5 Results

In the aspatial OLS model, DBH and HT explained roughly 50% of the variation in canopy class ($R^2 = 0.54, p_{dbh} = 0.00, p_{ht} = 0.00, p_{intercept} = 0.00$). The intercept term is 6.2, indicating that when HT and DBH are = 0, the successional status of the tree is at the lowest level, as would be expected. Here, and in other models, BLC has little-no effect and is not significant, and thus has been largely excluded from the rest of my interpretation and discussion. The spatial fixed effect model explains almost an additional 30% of the variation in canopy type ($R^2 = 0.80, p_{dbh} = 0.00, p_{ht} = 0.00$). A particularly interesting result from this model is the variation and significance of the intercept terms for each grid cell. While similar to the regular OLS, the intercept terms to fluctuate between 4-7, indicating spatial heterogeneity in canopy type in the absence of the effects of DBH and HT. This can be seen in mapping the residuals of both the OLS and spatial fixed effects models figures 5 and 7. The OLS residuals have large, positive, significant spatial autocorrelation ($I = 0.25, p = 0.001$), while the spatial fixed model has a much lower level of negative spatial autocorrelation ($I = -0.08, p = 0.001$). Visually, there is a clear spatial trend in the OLS residuals, while there is little or no obvious trend in the spatial fixed effect models residuals. Interestingly, both models more accurately predict the successional status of more mature trees than those earlier on in the successional trajectory, indicating more variance in the physical attributes of younger trees than more older, established individuals (figures 6 and 8).

The two variables with the largest spatial autocorrelation at the 5 meter grid cell level

were tree density ($I = 0.367$, $p = 0.001$) and height ($I = 0.356$, $p = 0.001$), followed by the heterogeneity index ($I = 0.249$, $p = 0.001$), diameter ($I = 0.199$, $p = 0.001$), and species richness ($I = 0.172$, $p = 0.001$). Mode successional status had a relatively low value ($I = 0.083$, $p = 0.001$), which isn't surprising given that it is more difficult to meaningfully aggregate ordinal variables than continuous variables. Due to this, I was not able to get any meaningful information from any modeling exercise performed at the grid cell level. It seems that the mode successional status of all trees in a 5 meter grid cell isn't as meaningful an ecological attribute as considering the status of each tree individually.

The geographically weighted regression shows a similar picture to the aspatial OLS module, when comparing the GWR's intercept term to the residuals in the aspatial OLS. When variables are standardized, trends in the intercept term of a GWR typically indicate a process or set of processes that have not been captured by the variables in the model. Similarly, in OLS when the residuals have a spatial pattern, it indicates the absence of a variable in the model that explains the trend. To quantify the similarities between these two approaches, I ran a linear model predicting the OLS residual as a function of the GWR intercept term, stratified by tree canopy type figures [18](#) to [23](#).

6 Discussion and Conclusions

Ecology is both inherently aspatial and spatial, and the spatial processes operate at various scales. This has been demonstrated through the two modeling exercised in this project, which both point to similar trends in forest succession as a function of physical tree characteristics (aspatial) and presumably site-related characteristics not represented by variables explicitly in these models, but identified by the intercept terms in both the spatial fixed effects and geographically weighted regressions. A logical next step in this analysis is to include site-related environmental variables that can account for the spatial variation in these intercept terms in these models. There are several that come to mind, mainly edaphic characteristics, as well as climatic variables such as localized average temperatures, soil moisture, all of which are influenced to varying degrees by the local

topography. It is well known that different tree species prefer different types of soil and climatic growing conditions; the same species growing on two different soil types may experience different growth trajectories, and tree successional status is both limited on poor growing sites and enhanced in favorable ones. Thus, providing site-characteristics that are directly related to a tree's ability to grow healthily and compete with neighboring trees may very well help to explain the spatial trends not accounted for in this analysis thus far.

Modeling the successional status of trees is best done at the individual tree level. While aggregating some tree and forest attributes to a grid may help visualization of ecological trends across a forest landscape, there is little information to be gained when looking at a composite measure of succession. This is likely due to several factors that have to do with successional processes. There are two main mechanisms by which trees reach maturity in an already established forest such as HF (there are more than only two but for the purposes of this paper I'm focusing on the two primary mechanisms that are most relevant here). The first is that when a large tree falls over, it creates a gap in the canopy. This gap is quickly filled by fast growing, intermediate shade tolerant trees. These individuals out compete their slower growing neighbors and quickly fill the gap. While it may take years to reach full maturity and peak successional status of a dominant or super dominant tree, there can easily be trees that comprise of different diameters and heights and even canopy levels growing in close proximity, as a function of the spatiotemporal dynamics of gap openings and closing throughout a relatively small forested area. Conversely, shade tolerant trees can establish and begin maturing in a very shaded understory. When a neighboring tree falls and creates a partial gap, these slower growing trees are poised to fill its place quicker than the faster growing trees because they've had a head start and have less growing to do to close and fill the gap than the intermediate shade tolerant trees that can only start growing once the gap opens and enough light penetrates through the canopy to the ground. Both these processes mean that it is very common to have trees of very different successional stages, sizes, and heights in close proximity to one another, and thus an aggregate measure of the

successional status of an area is probably less useful than looking at the individuals. This analysis shows the complexity of succession in a forest, which is influenced by many different factors that vary at different rates in both time and space.

This project has provided an initial picture of the site-specific characteristics of forest succession at HF, as well as several potential avenues for further research. It is obvious from both theoretical and quantitative perspectives that nonstationarity exists in the drivers of ecological succession at HF, and I'd like to provide a better identification of the spatial process that is captured in the GWR intercept term. Further work should also probably include species-specific approaches, or segmenting species by their shade tolerance, as the relative proportions of the three shade tolerances (tolerant, intermediate, and intolerant) is also important for the successional trajectory of that forest and what mechanisms might be at play (environmental factors vs light availability and gap regeneration, for example). This work provides an interesting application of spatial data science methods for assessing a well studied ecological phenomena which may be useful for further research in forest succession.

7 Figures

7.1 Exploratory Plotting

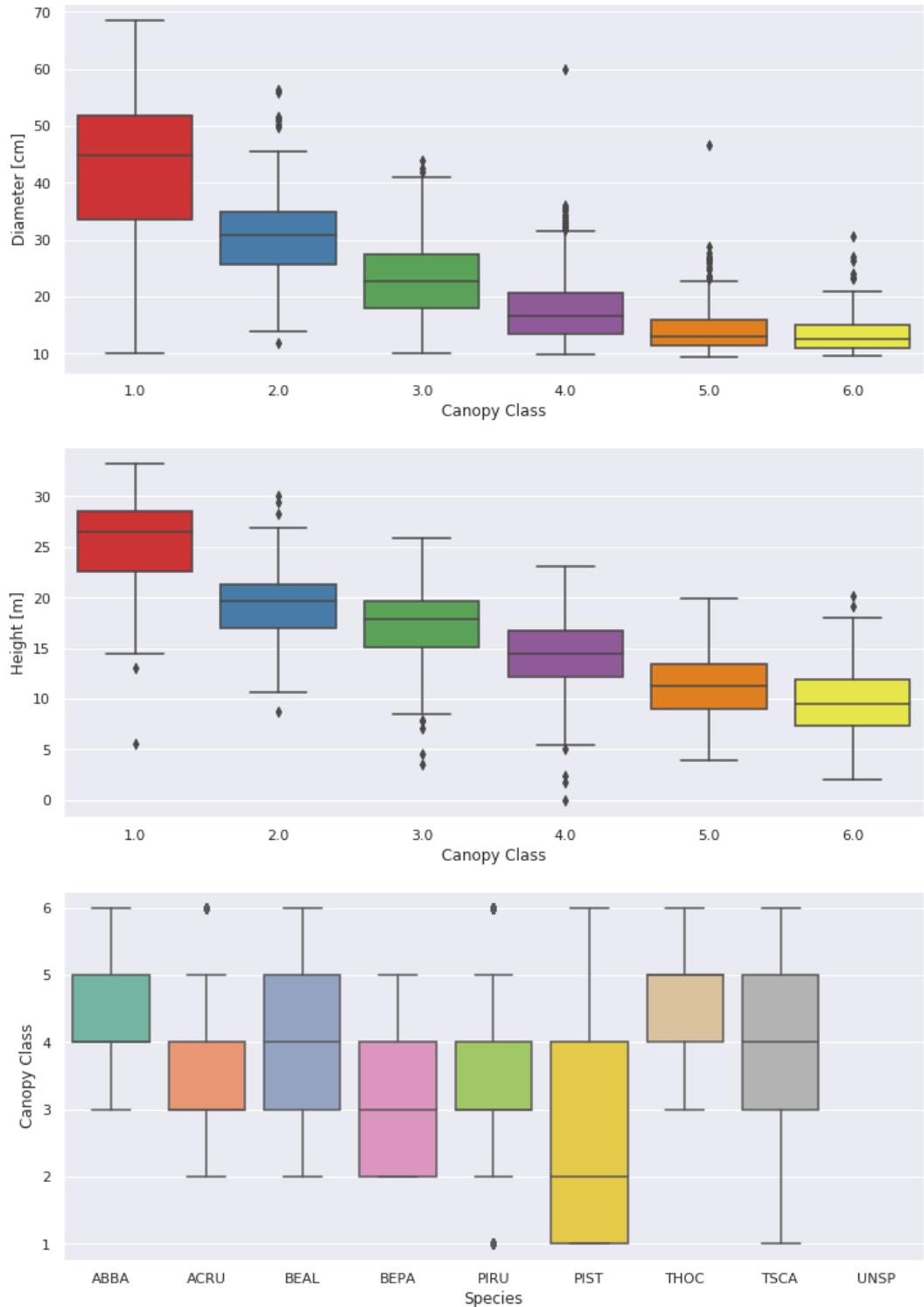


Figure 1: Boxplots of the dependent variable canopy class vs tree diameter (top), tree height (middle), as well as canopy class distributions by species (bottom).

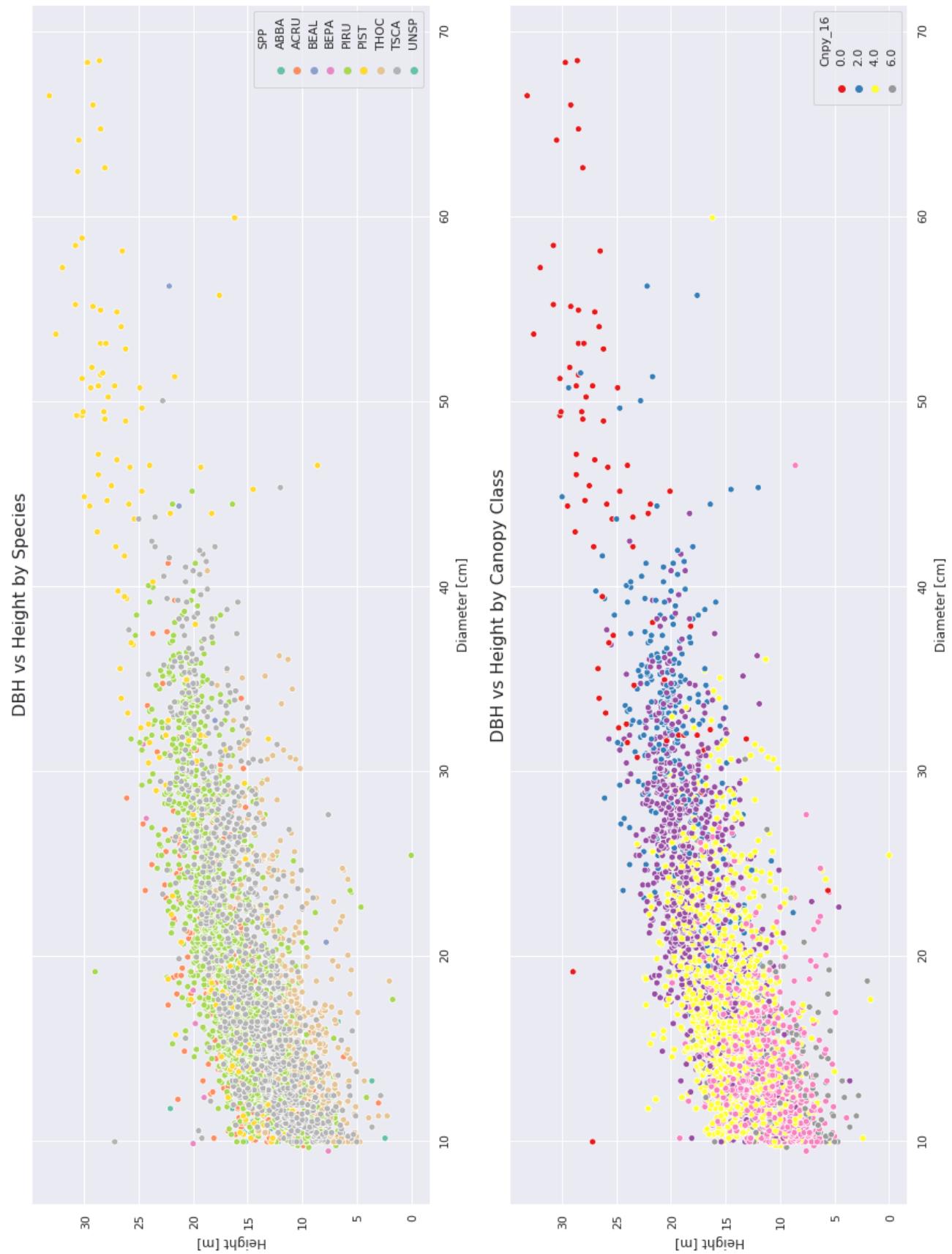


Figure 2: Scatter plots of tree diameter vs height, showing variations within species (top) and canopy class (bottom).

Howland Forest: NASA Plot DBH by PFT

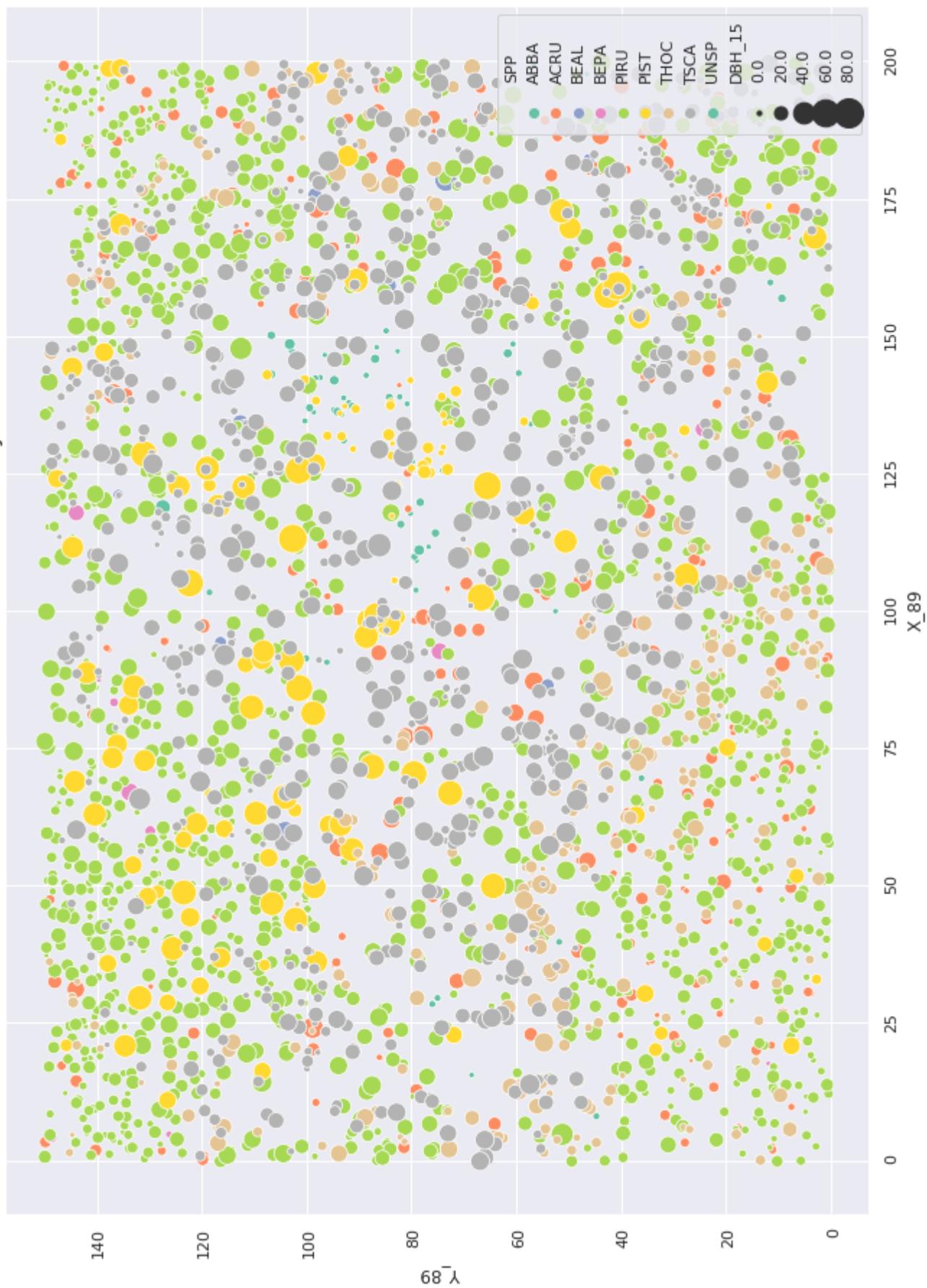


Figure 3: Spatial position of each tree within the NASA plot, colored by species and sized by tree diameter.

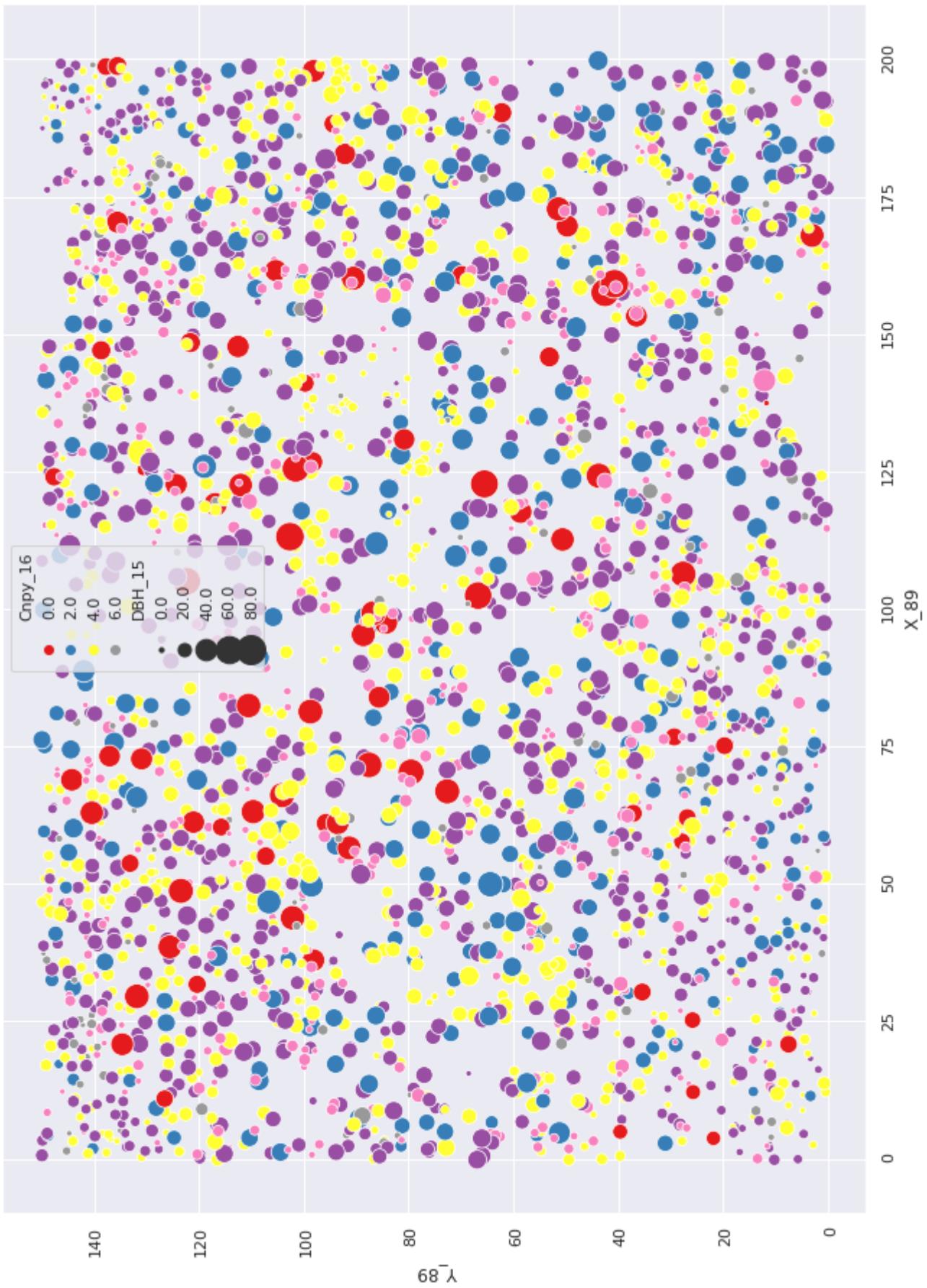


Figure 4: Spatial position of each tree within the NASA plot, colored by canopy class and sized by tree diameter.

7.2 OLS and Fixed Effects Regression

Aspatial OLS Model Residuals



Figure 5: Spatial position of the trees within the NASA plot, colored by the residuals of the aspatial OLS model.

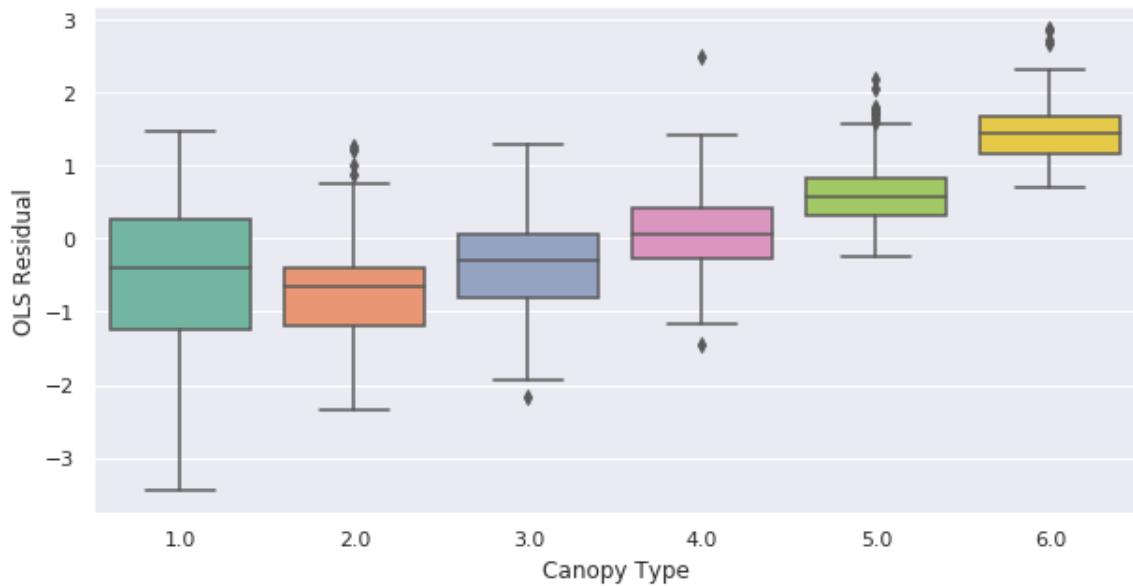


Figure 6: Boxplot indicating the distribution of the aspatial OLS residuals by canopy class. Notice the increasing trend, indicating the spatial model misses some trend relating to canopy class of the younger, less mature trees (higher canopy class values).

Spatial FE Model Residuals

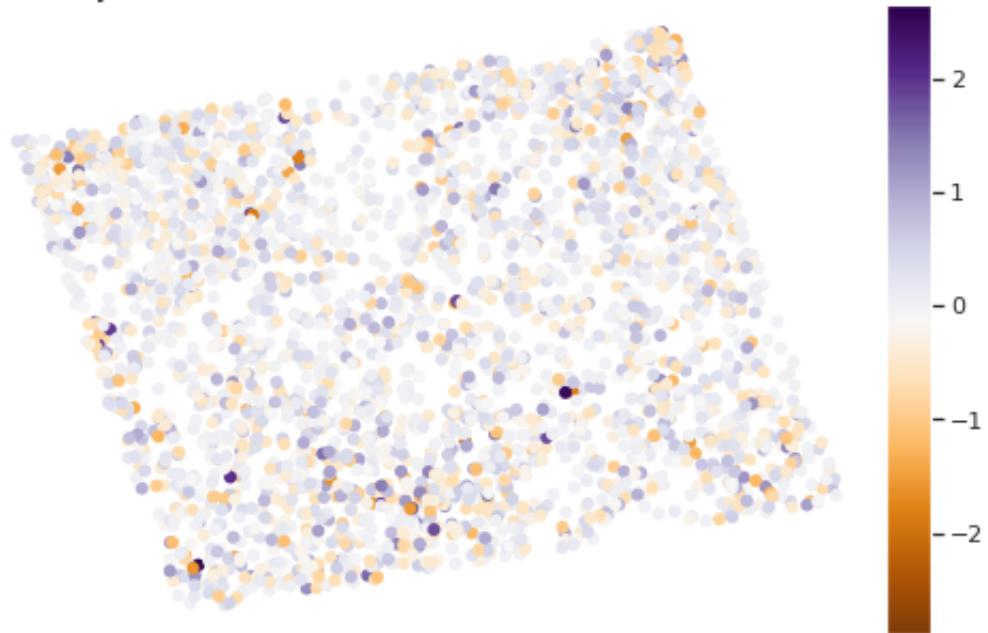


Figure 7: Spatial position of the trees within the NASA plot, colored by the residuals of the spatial fixed effects OLS model.

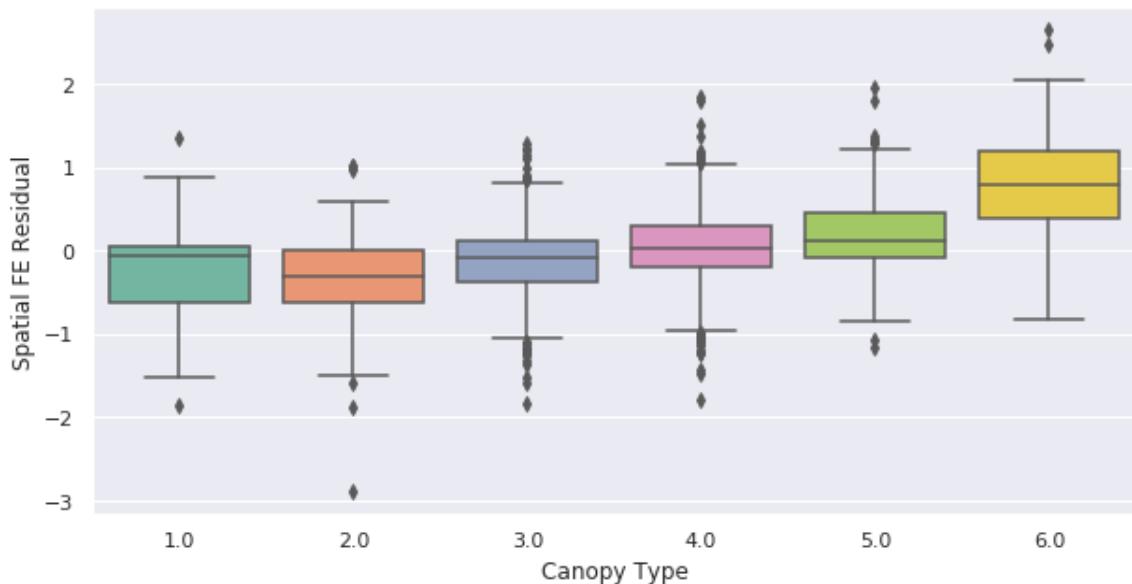


Figure 8: Boxplot indicating the distribution of the spatial fixed effects OLS residuals by canopy class. Notice the increasing trend is still present here, but must less pronounced than in the aspatial model, indicating a both spatial and aspatial component of the residuals in the OLS models.

7.3 ESDA

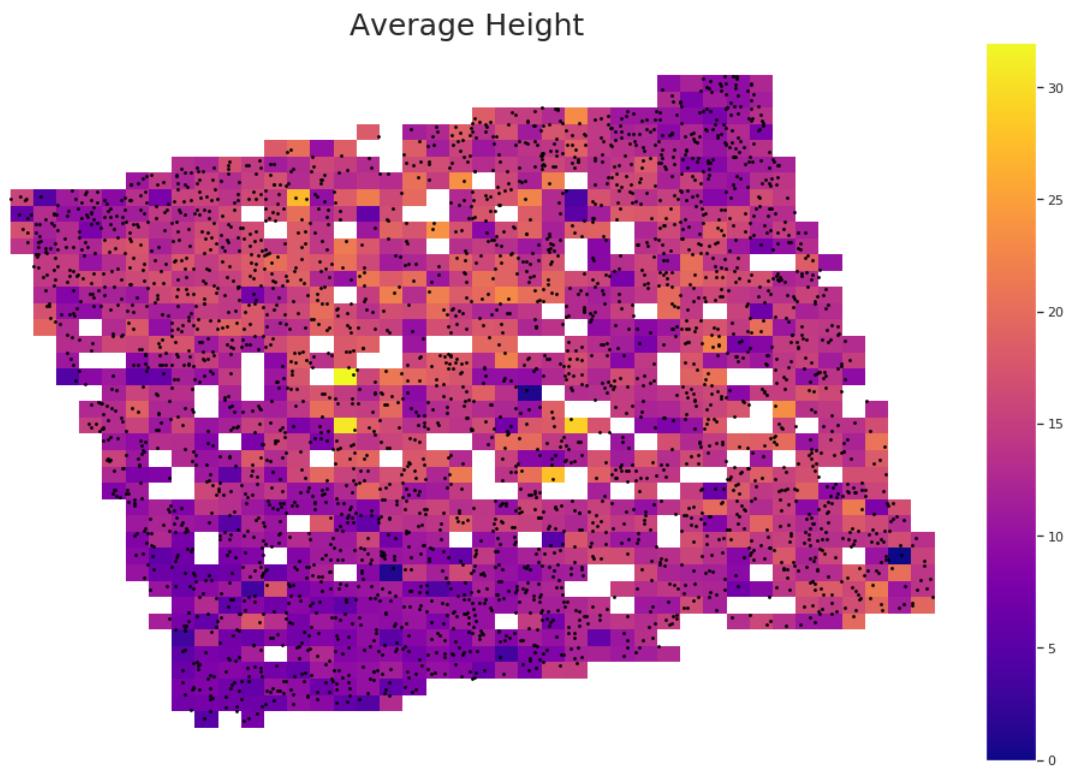


Figure 9: Gridded plot of average tree height at the 5 meter scale.

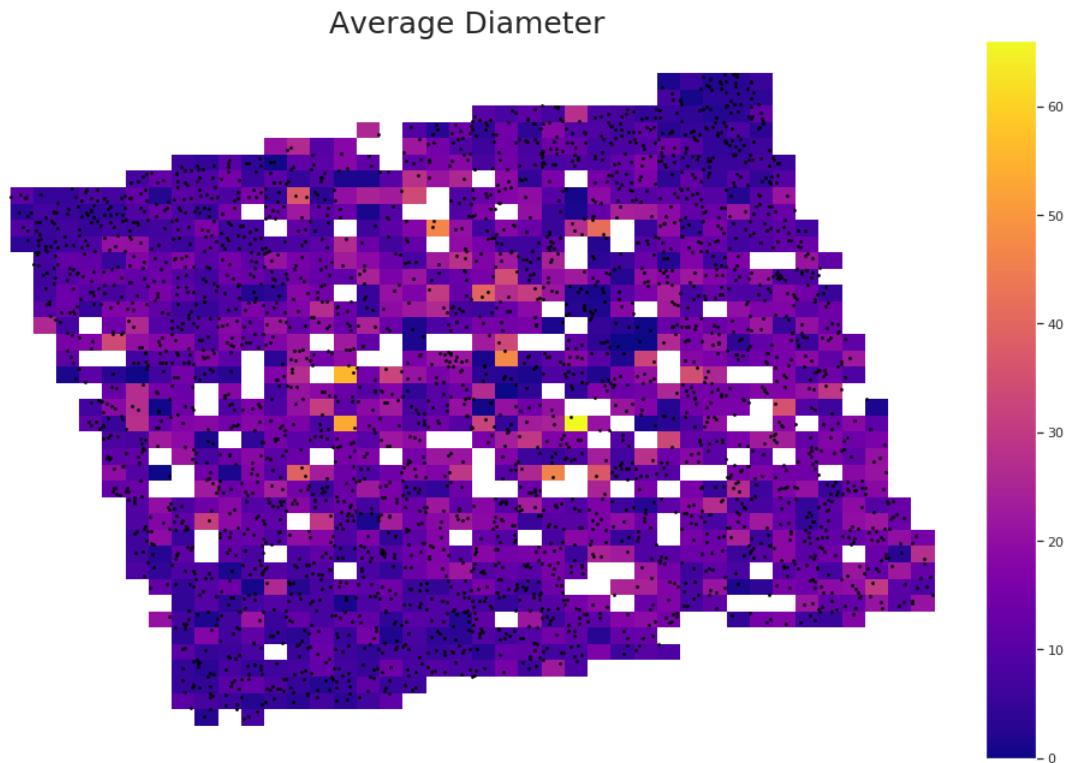


Figure 10: Gridded plot of average tree diameter at the 5 meter scale.

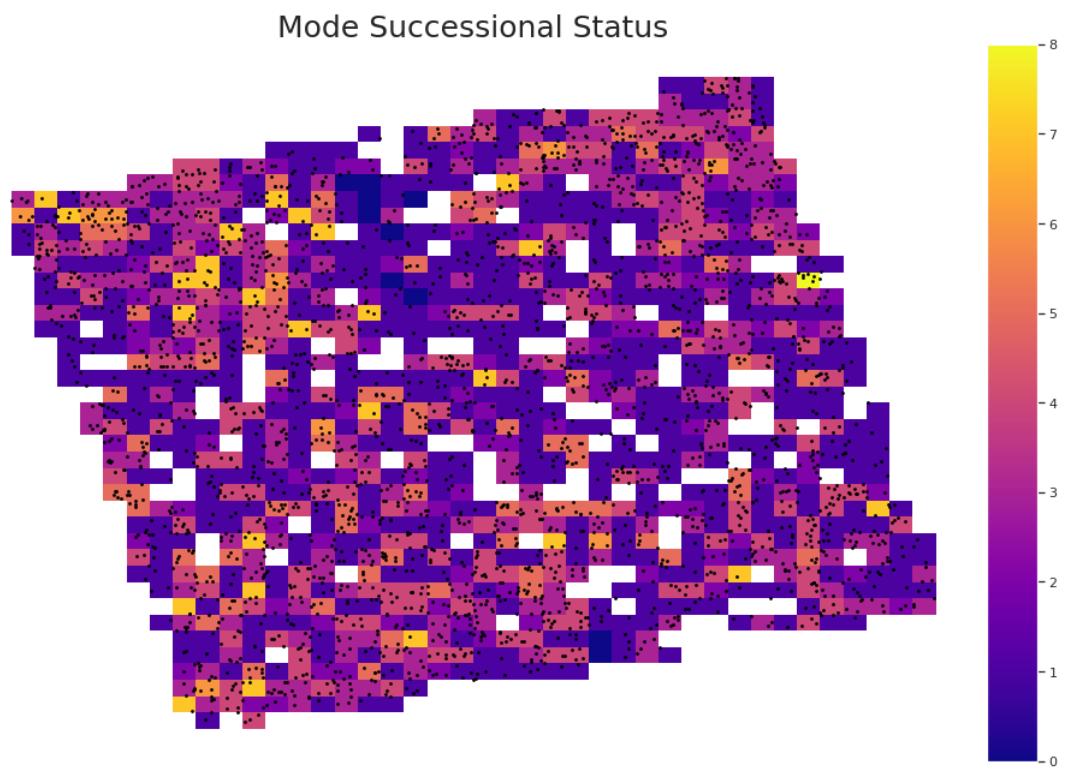


Figure 11: Gridded plot of mode canopy class at the 5 meter scale.

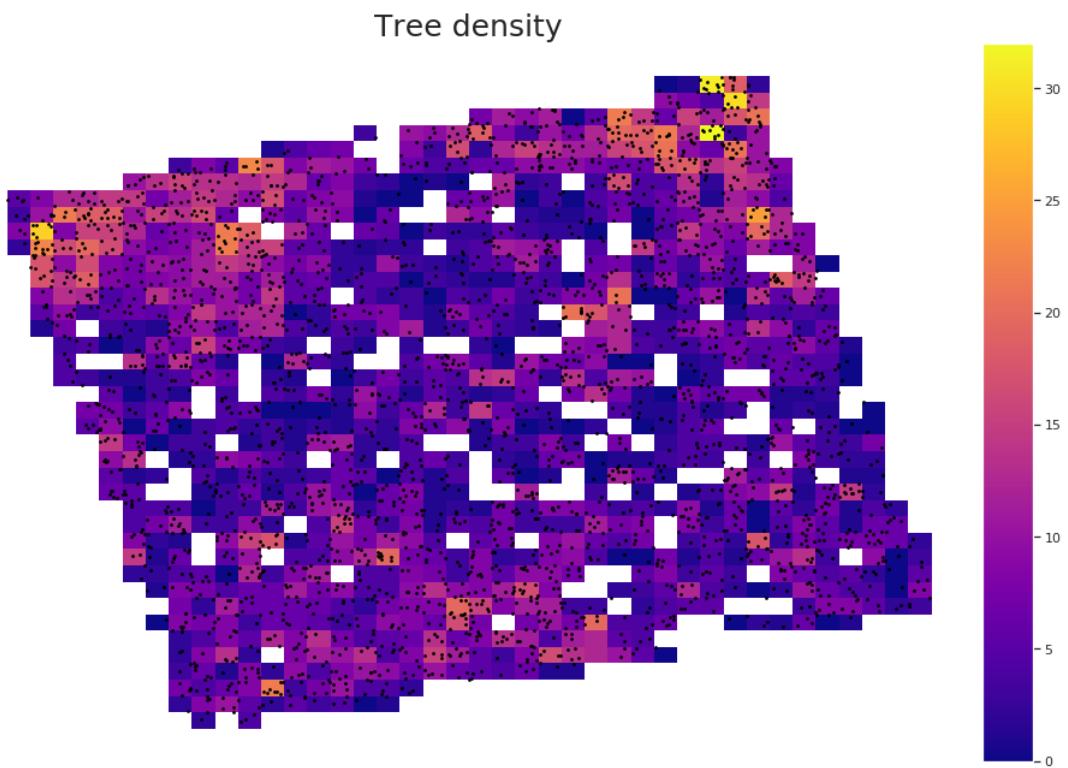


Figure 12: Gridded plot of tree density at the 5 meter scale.

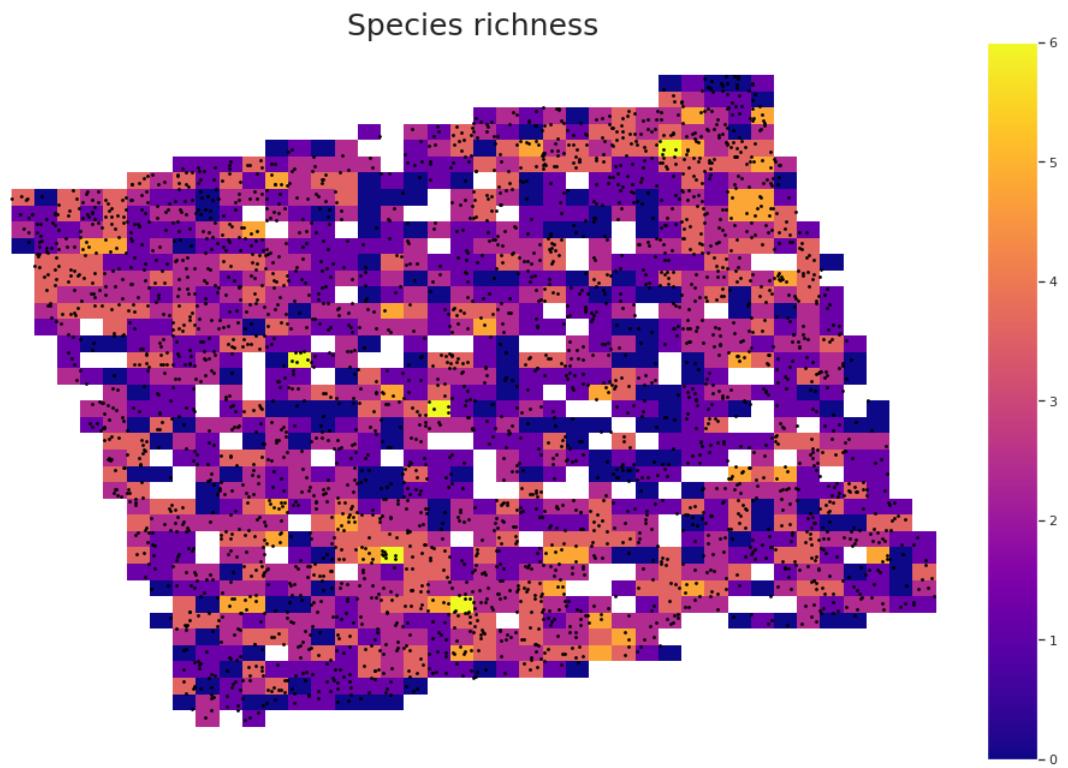


Figure 13: Gridded plot of species richness at the 5 meter scale.

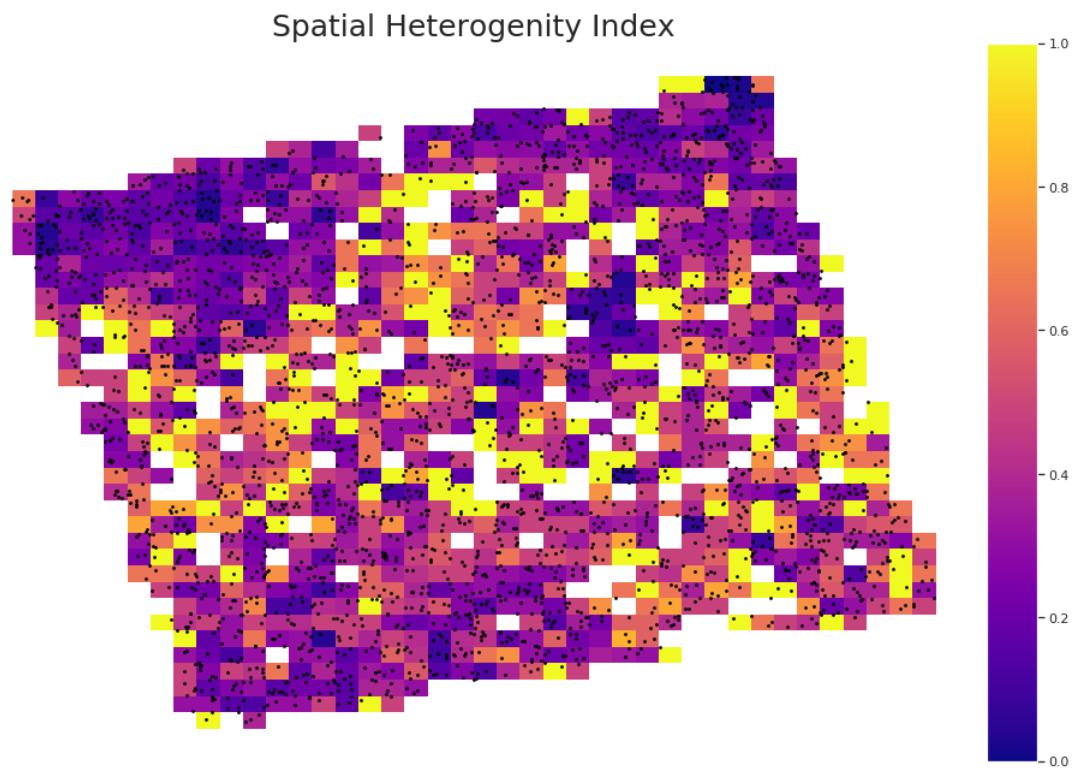


Figure 14: Gridded plot of spatial heterogeneity at the 5 meter scale.

7.4 GWR and MGWR

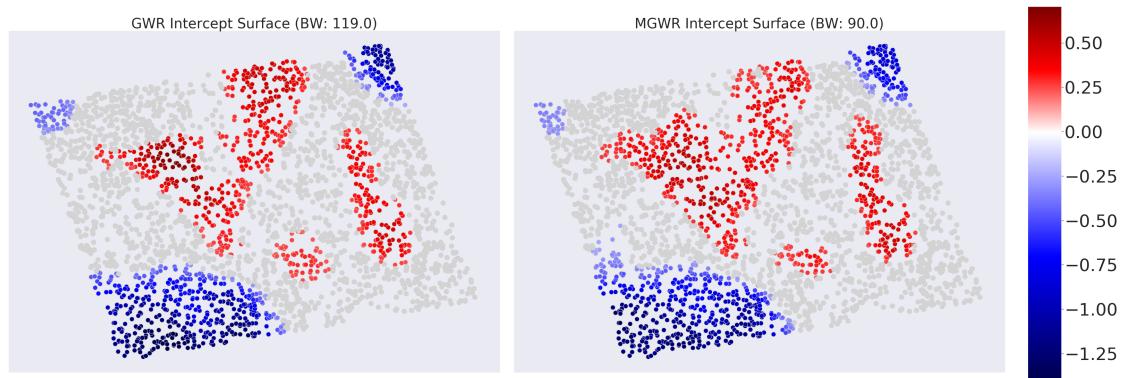


Figure 15: GWR and MGWR intercept surfaces. Colored points are significant at the 95% confidence level.

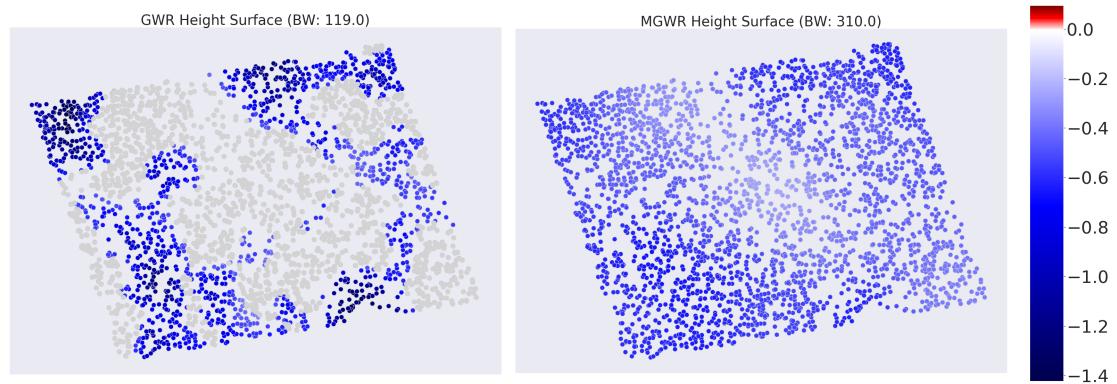


Figure 16: GWR and MGWR coefficient surfaces for height. Colored points are significant at the 95% confidence level.

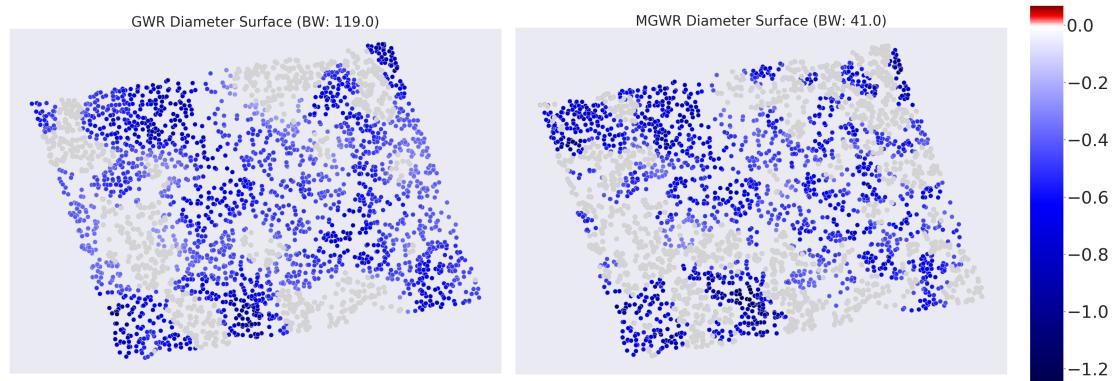


Figure 17: GWR and MGWR coefficient surfaces for diameter. Colored points are significant at the 95% confidence level.

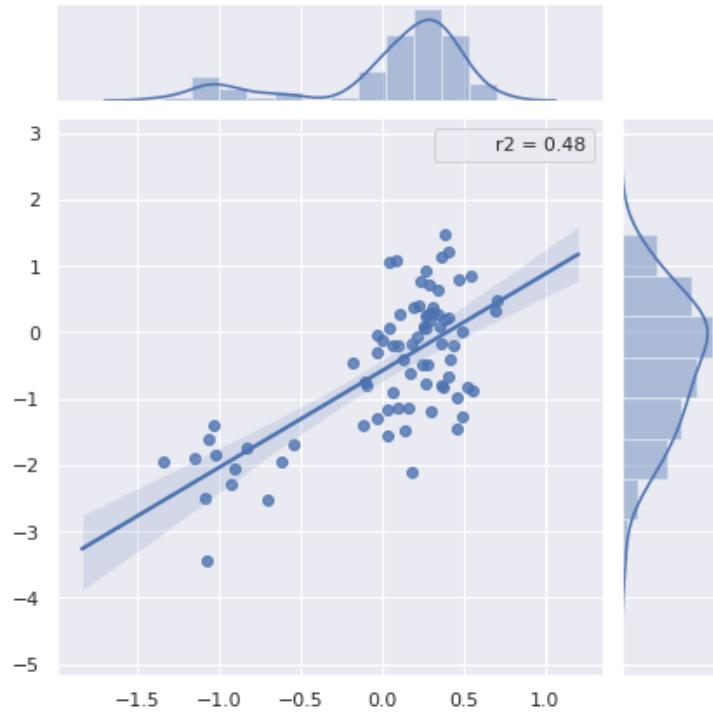


Figure 18: Scatter plot of the OLS model residuals vs the GWR intercept term, for super-dominant trees.

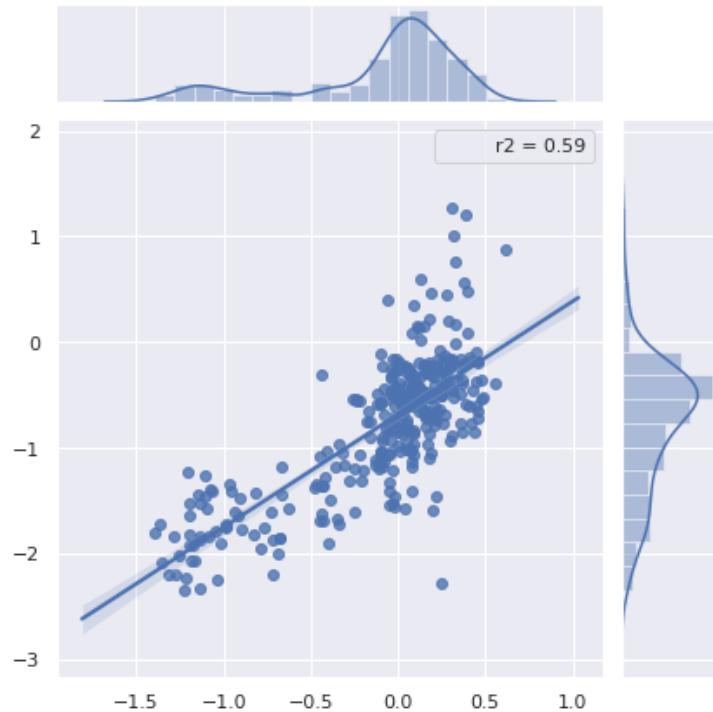


Figure 19: Scatter plot of the OLS model residuals vs the GWR intercept term, for dominant trees.

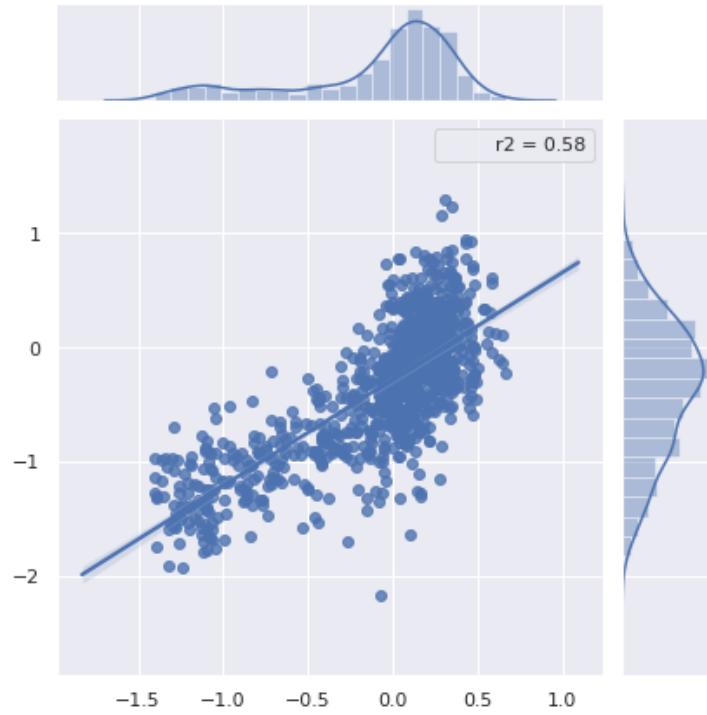


Figure 20: Scatter plot of the OLS model residuals vs the GWR intercept term, for co-dominant trees.

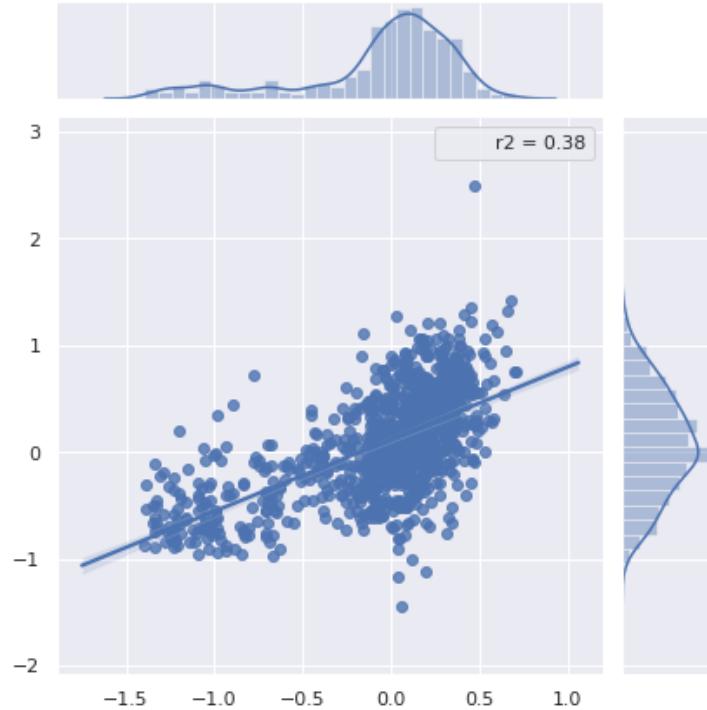


Figure 21: Scatter plot of the OLS model residuals vs the GWR intercept term, for intermediate trees.

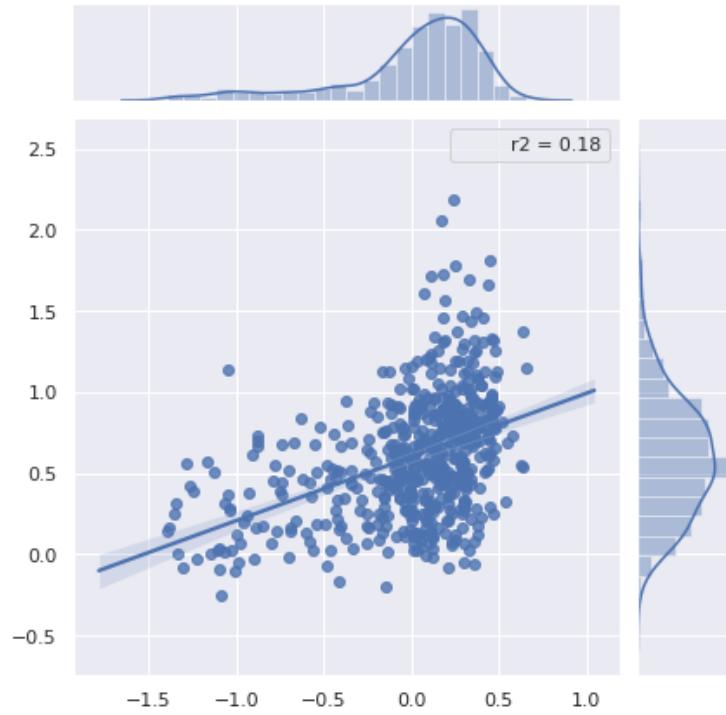


Figure 22: Scatter plot of the OLS model residuals vs the GWR intercept term, for over-topped trees.

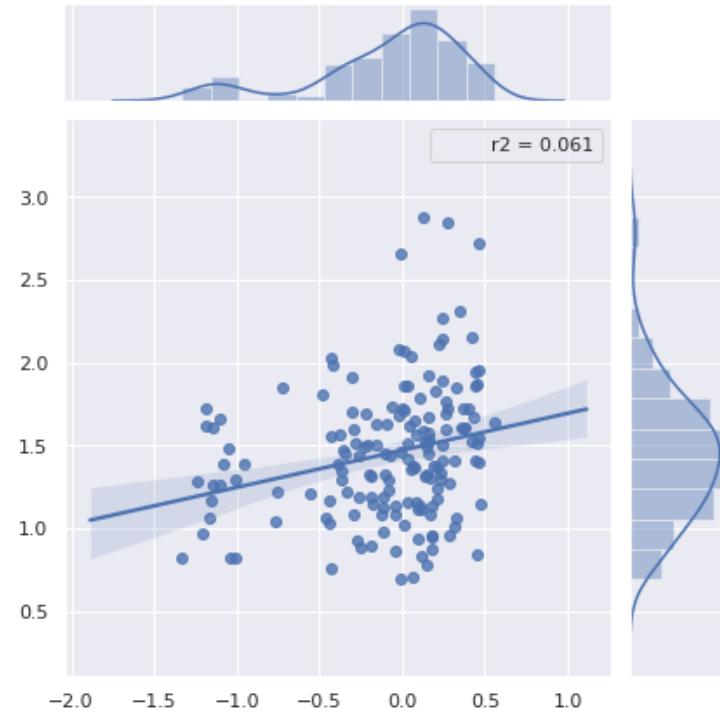


Figure 23: Scatter plot of the OLS model residuals vs the GWR intercept term, for suppressed trees.

8 References

- Albani, M., Medvigy, D., Hurt, G. C., and Moorcroft, P. R. (2006). The contributions of land-use change, co₂ fertilization, and climate variability to the eastern us carbon sink. *Global Change Biology*, 12(12):2370–2390.
- Anselin, L. (1995). Local indicators of spatial association?lisa. *Geographical analysis*, 27(2):93–115.
- Brunsdon, C., Fotheringham, A. S., and Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis*, 28(4):281–298.
- Canadell, J. G., Pataki, D. E., Gifford, R., Houghton, R. A., Luo, Y., Raupach, M. R., Smith, P., and Steffen, W. (2007). Saturation of the terrestrial carbon sink. In *Terrrestrial ecosystems in a changing world*, pages 59–78. Springer.
- Foster, D. R., Motzkin, G., and Slater, B. (1998). Land-use history as long-term broad-scale disturbance: regional forest dynamics in central new england. *Ecosystems*, 1(1):96–119.
- Fotheringham, A. S., Yang, W., and Kang, W. (2017). Multiscale geographically weighted regression (mgwr). *Annals of the American Association of Geographers*, 107(6):1247–1265.
- Houghton, R. A., Baccini, A., and Walker, W. S. (2018). Where is the residual terrestrial carbon sink? *Global change biology*, 24(8):3277–3279.
- Houghton, R. A., House, J., Pongratz, J., Van Der Werf, G., DeFries, R., Hansen, M., Quéré, C. L., and Ramankutty, N. (2012). Carbon emissions from land use and land-cover change. *Biogeosciences*, 9(12):5125–5142.
- Houghton, R. A. and Nassikas, A. A. (2017). Global and regional fluxes of carbon from land use and land cover change 1850–2015. *Global Biogeochemical Cycles*, 31(3):456–472.

- Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R., Feddema, J., Fischer, G., Fisk, J., Hibbard, K., Houghton, R., Janetos, A., et al. (2011). Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic change*, 109(1-2):117.
- Hurtt, G. C., Frolking, S., Fearon, M., Moore, B., Shevliakova, E., Malyshev, S., Pacala, S. W., and Houghton, R. (2006). The underpinnings of land-use history: Three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands. *Global Change Biology*, 12(7):1208–1229.
- Hurtt, G. C., Pacala, S., Moorcroft, P. R., Caspersen, J., Shevliakova, E., Houghton, R., and Moore, B. (2002). Projecting the future of the us carbon sink. *Proceedings of the National Academy of Sciences*, 99(3):1389–1394.
- Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., and Fotheringham, A. S. (2019). mgwr: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geoinformation*, 8(6):269.
- Runkle, J. R. (1981). Gap regeneration in some old-growth forests of the eastern united states. *Ecology*, 62(4):1041–1051.
- Schimel, D., Stephens, B. B., and Fisher, J. B. (2015). Effect of increasing co2 on the terrestrial carbon cycle. *Proceedings of the National Academy of Sciences*, 112(2):436–441.
- Shugart, H. H. et al. (1984). *A theory of forest dynamics: the ecological implications of forest succession models*, volume 278. Springer-Verlag New York; Berlin; Heidelberg; Tokyo.
- Symanzik, J. (2014). Exploratory spatial data analysis. *Handbook of regional science*, pages 1295–1310.
- West, D. C., Shugart, H. H., and Botkin, D. (2012). *Forest succession: concepts and application*. Springer Science & Business Media.

Yamamoto, S.-I. (1992). The gap theory in forest dynamics. *The botanical magazine=Shokubutsu-gaku-zasshi*, 105(2):375–383.