Factors influencing red oak (*Quercus rubra*) seedling survival at the northern edge of the species range

Sage Wentzell-Brehme

Submitted in Partial Fulfillment
of the
Prerequisite for Honors
in Environmental Studies
under the advisement of Jaclyn Hatala Matthes

May 2021

© 2021 Sage Wentzell-Brehme
# Table of Contents

Acknowledgements ........................................................................................................... 3

1. Executive Summary ........................................................................................................ 5

2. Introduction ..................................................................................................................... 7
   2.1. Global scale ............................................................................................................ 7
   2.2. Disturbance regimes ............................................................................................... 8
   2.4. Historical Context ................................................................................................ 12
   2.5. Climate migrants & seedling survival .................................................................. 16
   2.6. Northern red oak ................................................................................................. 18
   2.7. Importance of studying seedlings ........................................................................ 20
   2.8. Questions ............................................................................................................... 21

3. Methods .......................................................................................................................... 24
   3.1. Site description ...................................................................................................... 24
   3.2. Data collection ...................................................................................................... 26
   3.3. Data processing .................................................................................................... 29
   3.4. Data analysis ........................................................................................................ 30

4. Results ............................................................................................................................ 33
   4.1. Valley-wide Patterns ............................................................................................ 33
   4.2. Environmental variables & plot level survival .................................................... 36
   4.3. Individual seedling variables & individual survival for a single year ................. 37
   4.4. Multi-year survival modeling ............................................................................. 40
   4.5. Winter climate data ............................................................................................. 43

5. Discussion ......................................................................................................................... 46
   5.1. Dispersal and distribution dynamics .................................................................... 46
   5.2. Plot-level seedling survival interacting with environmental variables .............. 50
   5.4. Modeling multi-year survival ............................................................................. 53
   5.5. Interannual variation in frost depth, snow depth and seedling survival ............ 54

6. Conclusions ...................................................................................................................... 56

7. References ....................................................................................................................... 58

8. Appendix ........................................................................................................................ 75
Acknowledgements

To my fabulous advisor Jackie Matthes, I truly could not have done it without you. Thank you for the constant support and encouragement throughout this whole process (and for troubleshooting my R code more times than I can count). I always left our weekly meetings energized and ready to tackle my research again. Your excitement about doing science and endless ecology knowledge made me fall in love with ecology research. You are such an incredible mentor, and I am so grateful that I have had the opportunity to work with you the last few years.

To Alden Griffith & Jay Turner, thank you for being such supportive members of my thesis committee. Jay - your Calderwood class helped improve my ability to write about science and my research in a less technical manner so much. Alden - thank you for introducing me to R and being such a supportive first teacher. That was the first time I learned code without being immediately intimidated. And to Dan Chiasson, thank you for being my honors visitor! It feels very full circle to have you here at the beginning and end of my Wellesley journey.

I would also like to thank all the incredible professors I have had the privilege to work with throughout my Wellesley career. Whether your classes are directly tied into this thesis or not, you all have helped me become a stronger, more knowledgeable person and I am excited to see what I do after Wellesley. And a special shout out to the Environmental Studies Department for welcoming and supporting me throughout my time at Wellesley.

Warm thanks to my incredible advisor, Nat Cleavitt, without whom this project would not have been possible. Thank you for introducing me to the magic of red oak seedlings and teaching me how to do field research. I learned so much from our summer together and I am so excited to be working with you again this year!

Thank you as well to John Battles for countless afternoons lugging around that camera and being endlessly patient while teaching me how to take canopy photos.

Shoutout to Veg Crew 2019: Katie & Alex, thank y’all for being my field partners! You guys kept it light and fun when the fieldwork got long and hard. I could not have measured as many red oak seedlings or taken as many canopy photos without you both. I am so grateful for the entire Hubbard Brook research community. Your warm welcome and encouragement helped convince me that I really could be a scientist.

Thank you to the Hubbard Brook LTER grant NSF-1637685 and the Wellesley Science Center Summer Research Program for their financial support that gave me the opportunity to pursue this project.
To my fellow Ecolab members, Dayna, Sarah Elizabeth, Jaime, and Marisy, thank you for being so supportive this year. I am going to miss our weekly lab meetings a whole lot. It has been so lovely to chat with you all (albeit virtually) every week. Geeking out about plants or commiserating over wrestling with R with y’all kept me sane. And to my former fellow Ecolab mates, Eva & Emma, thank you for both for inspiring me to pursue a thesis and showing me that it could be done.

Thank you to my Counterpoint family for providing respite and a chance to talk about books, music, and inane pop culture. Getting to see you all at our weekly meetings kept me connected to the Wellesley community during this incredibly strange year and I am so grateful for that. And thank you to Sam English for inspiring me to consider writing a thesis and running for EIC of Counterpoint at all. I hope I could fill your shoes at least a little.

And thank you to my other Wellesley people for being bright spots in a crazy year: Zaria, Maggie, and Caroline – I miss you all! I would not have made it through this process or enjoyed Wellesley so much without you all <3

To my two incredible roommates and best friends, Delaney Robertson and Mira Carlininia, thank you for putting up with me this last semester. You never once shook me and told me to stop babbling about my thesis in either excitement or distress and it was truly so appreciated. Thank you for the late-night study sessions, the random hugs, and all-around moral support. Laney—thank you for always being up for an impromptu late night shopping trip or hours of binging terrible reality television to escape our own messy lives. Hopefully we’ll be #thriving soon! Mira—thank you for sending me random cat TikToks all the time and starting me on the path to obsession with 1D: our comfort men really saved me this year. And to my OG roommate and best friend, Natalie Marshall, thank you for always believing in me. I truly never could have survived Wellesley without you. Our Zoom study sessions and your endless supportive texts kept me going this year. Thank you for being the best roommate a girl could ask for. I miss our late-night talks and slightly-tipsy-on-sleep-deprivation dance parties more than I can say. I am so grateful to have the three of you in my life: love you to the moon and to Saturn.

And finally, thank you to my parents, Kristina Wentzell and Chris Brehme, for being my first and fiercest supporters. Thank you for inspiring my love of forests by taking me hiking before I could walk. Or maybe the science fair project in 4th grade really cemented my love of counting trees; twelve years later and I’m still doing it! Thanks in particular to Papa for helping me make a bomb map! To my brother Zeb, thank you for always sharing your random nature facts (even when they are about spiders). I am so glad that you are similarly obsessed with the outdoors. And to Grammy, thank you for always being entirely confident in my abilities. I could not have done it without you all <3
1. Executive Summary

*Quercus rubra* (northern red oak) is expected to migrate into northern hardwood U.S. forests with climate change impacts. Warming temperatures and shifting precipitation patterns across the northeastern U.S. region as a result of climate change (Janowiak et al., 2018) are changing the ideal ranges for tree species (Iverson et al., 2004a). The warmer temperatures and potential droughts will likely be easier for red oak to withstand than many of the more prevalent native species across the northeast such as sugar maple (Iverson et al., 2008). This could create opportunities for red oak to outcompete other species and expand their range into northern hardwood forests.

More indirect effects of climate change such as the spread of hemlock woolly adelgid and emerald ash borer invasive insect pests (Eschtruth et al., 2006; Kovacs et al., 2010) will also affect the success of red oak establishment. The subsequent loss of eastern hemlock and white ash trees will likely shift the northern hardwood forest composition (Dukes et al., 2009; Kovacs et al., 2010) which could allow for red oak to become an established novel species. Seedling regeneration is crucial to understanding the future of our forests as the migration of tree species into new environments is determined by seedling growth and survival, but little research is done on the seedling life stage.

Forests in New England and across the world provide ecosystem services including carbon sequestration, watershed protection, and species habitat. In New Hampshire, where red oak is moving northwards towards the White Mountains, forests offset about 38% of the state’s total carbon dioxide emissions (Domke et al., 2020). The conservation of forests offers economic value through carbon sequestration, natural beauty and cultural importance. NH forests also have essential cultural and economic value through the timber and tourism industries (Hall et al., 2011). The factors affecting forest composition are changing at such a fast rate that they have the potential to outpace management options, requiring forestry and tourism to adapt, which will impact rural communities throughout New England. Information as to the future of these forests is essential to maintaining the livelihoods and economic stability of these communities. My research offers a window into the future of northern hardwood forests. An integrated understanding of forests in both the past and the present will allow us to improve their resiliency.

To help contextualize this study, I considered how historical land use changes or shifts in species composition impacted current forests. The interacting timber and tourism industries in New Hampshire ultimately shaped our current forests which now face even more threats, including climate change and increasing exurban development.

Some studies have found red oak seedling survival to be dependent on light (Brose, 2011; Frey & Ashton, 2018; Major et al., 2013), while others have found local competition, initial seedling size, nutrient availability, and other factors to be more impactful to red oak seedling success (C. D. Canham et al., 1996; Coomes & Grubb, 2000; Hartman et al., 2005; Willis et al., 2015). Particularly along the northern edge of the species range, where red oak is entering new territory, it is difficult to predict seedling success. This makes red oak an important case study.
for understanding the changes northern temperate forests are facing. My research examines how different factors influence red oak seedling survival at the northern edge of its range.

I looked at red oak seedling survival on a local scale by comparing survival across sites at the Hubbard Brook Experimental Forest (HBEF) in Woodstock, NH, which is a Long-Term Ecological Research (LTER) site at the northern edge of the red oak species range. During Summer 2019, I collected data on surrounding shrub cover, the mature tree community composition and size distribution, ground cover, and light availability for ~600 red oak seedlings. I combined these new data with long-term seedling survival monitoring that began in 2011 with measurements of seedling age, size, number of leaves, and associated leaf damage from herbivores. Statistical modeling was used to identify the most important environmental and seedling characteristic variables influencing red oak survival.

I found that leaf number was an important predictor of seedling survival over multiple years. This positive relationship between leaf number and survival supports the importance of size and resource availability as essential to understanding seedling survival (García & Houle, 2005; Moles & Westoby, 2004). It is important to continue this long-term research as even comparing the effects of various individual characteristics on seedling survival in a single year versus multiple years demonstrates that not all variables have consistent relationships with survival. Although both age and seedling leaf damage were significant predictors of survival in a single year, the strength of this relationship diminished when looking at a longer time frame.

My results showed that the distance from mature red oak trees did have a positive effect on survival as expected (Connell, 1971; Janzen, 1970; Jevon et al., 2020). I did not find support for light being important to predicting oak seedling survival, which went against my hypothesis and several studies (Brose, 2011; Frey & Ashton, 2018; Major et al., 2013). This could be explained by the extremely low light available in areas with a hemlock-dense canopy. This research emphasizes the need to further research the interaction of hemlock die-off with red oak regeneration. Expanding the dataset of environmental factors could provide insights to how red oak seedling survival will evolve with the changing forest composition and climate.
2. Introduction
   2.1. Global scale

Anthropogenic climate change is one of the most important issues of the 21st century. The main drivers of climate change are increasing atmospheric greenhouse gas concentration as a result of human emissions of carbon dioxide, methane gas, and nitrous oxide (IPCC, 2013). But climate change is going to have vastly different impacts based on location. The northeastern United States has already seen a 2.4 °F increase in average annual temperature between 1901 and 2011 (Janowiak et al., 2018). And climate models predict that the speed of this warming is only going to increase (Fan et al., 2014). Warm temperature anomalies are particularly high in winter months (Contosta et al. 2019), likely tied to the dramatic warming in the Arctic (Cohen, 2016).

While precipitation is projected to increase overall across the year, the distribution of precipitation will likely change. New England specifically is projected to have both drier summers and wetter autumns and winters (Janowiak et al., 2018).

But climate change is also bringing increased variability in many weather patterns. Global climate patterns are controlled by the regional jet streams which are heavily influenced by the polar vortex. The polar vortex is maintained by extreme cold in the Arctic and the loss of these typical climate controls, such as warming in the Arctic, is creating much greater variation in climatic conditions in the northern hemisphere (Cohen, 2016). The rapid warming of the Arctic, where temperatures are increasing at twice the global average rate, is thought to be unbalancing global wind patterns, leading to a greater frequency of extreme weather events in midlatitudes (Francis & Vavrus, 2012). This means greater intensity in weather events, which includes larger, more destructive storm events, as well as heightened erraticness in general weather patterns, from unpredictable precipitation to wild temperature swings.
An increase in temperature variability, particularly during the winter, can have drastic consequences for forest ecosystems. But of particular concern is the changes during the winter season. The northeastern United States is seeing winter temperatures warming three times faster than throughout the rest of the year (Garlick et al., 2020). Climate models show that warming temperatures and changing precipitation patterns will only worsen in northern temperate forests over the next century (Fan et al., 2014).

2.2. Disturbance regimes

Climate change promises shifts in species composition of forests across the world. The northeastern United States is already seeing northwards tree species migration with significantly higher presence of seedlings in northern forests in comparison to southeastern United States locations (Woodall et al., 2009). And climate change projections predict even more tree species moving northward or experiencing stress and mortality in present locations over the next century (Iverson et al., 2008).

Global trends in forest dynamics show a progression towards younger stands with high turnover while the area covered by old-growth forests is shrinking (McDowell et al., 2020). This is likely due to climate change causing a combination of long-term stressors on forest ecosystems as well as an overall increase in episodic forest disturbances (Dale et al., 2001). Land-cover change such as deforestation and regrowth is another important factor when looking at global trends in forest dynamics (McDowell et al., 2020). Figure 1 shows how this combination of chronic and episodic stressors can lead to species immigrations (Smith et al., 2009). Many of the secondary impacts of climate change, such as increased spread of invasive species, or increased frequency of extreme weather events, are altering the disturbance regimes of forests (Dale et al.,
Different disturbances can cause widespread tree mortality, or more localized, species-specific mortality. Disturbance patterns are essential to understanding forests. Whether natural or anthropogenic, these disturbances determine the future demography and dynamics of forests. Northern temperate forest disturbance regimes rely on small canopy gaps to facilitate changes in species composition and forest dynamics through time (DeGraaf and Miller 1996; Worrall et al. 2005). But this disturbance regime pattern is likely to be interrupted by the increase in disturbance events in the northeastern United States as an impact of climate change (Dale et al., 2001).

Figure 1. From Smith et al. 2019 conceptually shows how episodic disturbances such as extreme weather events have the potential to lead to species immigrations.
An increase in extreme weather events such as ice storms or microbursts has the potential to be extremely influential to northern temperate forests. These forests typically rely on small, canopy gap disturbances caused by the natural death or breakage of one or two trees (Seymour et al., 2002). As the weather intensifies it creates many more gaps that are fundamentally transforming these forests. Of particular concern is the predicted increase in ice storms. These can be extremely damaging to forest canopies, as the ice accumulation causes branches to break off or whole trees to fall (R. T. Fahey et al., 2019). These resultant canopy gaps are part of the natural disturbance regime of New England forests, but climate change is drastically increasing the scale and frequency of these storms (R. T. Fahey et al., 2019; Klima & Morgan, 2015). Similar disturbance patterns are created by other weather events including windstorms and microbursts. These sudden columns of downward sinking air create extreme localized wind events at low altitudes that can cause severe damage. Research is ongoing into the long-term impacts of the 2013 microburst at my study site of Hubbard Brook Experimental Forest (HBEF).

Warmer winters means an increase in the abundance and spread of invasive pests such as the hemlock woolly adelgid with the loss of bitterly cold days that would kill them off (Contosta et al., 2019). Hemlock woolly adelgids are slowly wiping out *Tsuga canadensis* (eastern hemlock) across the eastern United States (Paradis et al., 2008). Other invasive species, such as emerald ash borers, ticks, and Asian tiger mosquitos, will be similarly affected by these warming temperatures (Garlick et al., 2020). This unprecedented spread of numerous invasive pests all at once is placing tremendous strain on northeastern forests. Insect pests and pathogens will have varying impacts on changes to the forest structure and there is uncertainty when considering how much influence climate change will exert (Dukes et al., 2009). But the pressure of numerous
species-specific invasive pests (such as hemlock woolly adelgid and emerald ash borer) has the potential to wipe out multiple species at once, causing massive shifts in the community composition of northeastern forests (Canham 2020). Regardless of the specific future forest composition, forest insect pests can have major implications for long-term carbon and nitrogen cycling (Crowley et al., 2016).

### 2.3. Importance of light to seedlings

Within a mature forest, canopy gaps allow for forest dynamism; the shifts and changes as new species move in and other species or older individuals die out. Northern temperate forest disturbance regimes rely on these small canopy gaps to facilitate the movement of species (DeGraaf & Miller, 1996; Worrall et al., 2005). These projected changes in disturbance patterns as a result of climate change will have an impact on species composition of forests. The most common linkage between disturbances and shifts in species composition is the increased light associated with forest disturbances (Seymour et al., 2002). Whether an entire forest canopy is wiped out by an ice storm or a single tree falls after a pest infestation, it creates increased light for the forest floor. Light has been found to be a crucial variable for seedling survival, although how seedlings respond to increases in light is somewhat species dependent (Beaudet & Messier, 1998; Catovsky & Bazzaz, 2000). As Bolton & D’Amato point out, natural disturbances such as canopy gaps lead to increased seedling abundance, but not necessarily increased diversity (2011). Only certain species are able to take advantage of different resources like light. Current forest disturbances are critical to forming a complete understanding of northern temperate forests, but a historical background is also extremely useful in providing a context for these modern changes.
2.4. **Historical Context**

Pleistocene era climate change led to the retreat of the glaciers across the northeastern United States around 13,000 years ago. Early regeneration of forests in the region as the climate warmed, around 12,000 years in the past, saw high abundances of conifers (C. Canham, 2020). The beginning and end of the Younger Dryas, a 1000-year cooling period caused by shifts in sea-surface temperature, saw rapid changes in vegetation across North America (C. Canham, 2020; Oswald et al., 2018; Shuman et al., 2002). One study found that some taxa migrated hundreds of kilometers in mere centuries (Shuman et al., 2002). These rapid ecosystem responses to sudden climatic changes are particularly useful when forecasting future shifts in forest composition in response to anthropogenic climate change. Across the northeastern United States, conifers gave way to birch at the end of the Younger Dryas (C. Canham, 2020). Further warming about 10,000 years ago allowed *Quercus* (oak) and *Pinus rigida* (Pitch pine) to spread throughout southern New England while cooler conditions in northern New England maintained the dominance of *Pinus strobus* (Eastern white pine) (Oswald et al., 2018). Rising precipitation over the next two millennia allowed for *Betula* (birch), *Fagus grandifolia* (American beech), and *Tsuga canadensis* (eastern hemlock) to overtake it. Sandy glacial deposits across New England were continuously dominated by *Quercus* (Oswald et al., 2018). *Quercus* remained a strong presence across southern New England aside from a decline in coastal areas concurrent with a cooling period 5000 to 5500 years ago (C. Canham, 2020; Oswald et al., 2018). Major shifts in forest vegetation composition were clearly common in North America before European settlement, perhaps signaling that such changes are to be expected.

The timber and tourism industries are more recent historical drivers of compositional shifts in forests across the northeastern United States. New Hampshire is the second most
forested state in the United States, with 81% of its land cover comprised of forests (Morin et al., 2015). These forests offset roughly 38% of the state’s total carbon dioxide emissions (Domke et al., 2020). But small-scale sustainable timber harvesting does not necessarily mean deforestation or even a net loss of carbon storage. The timber and forest products industries contribute about $1.5 billion to the state’s economy every year (The Economic Importance of New Hampshire’s Forest-Based Economy, 2011). But the conservation of forests also offers important economic value. Forests are the backbone of the tourism industry throughout New England, and especially in popular outdoor recreation areas such as northern New Hampshire. Tourism contributes $5 billion annually to the economy in New Hampshire alone (2011). This interplay of forests as an essential revenue source for the state through forestry or tourism has been the case for over a century.
There was heavy deforestation in New England following European settlement due to farming. Forest loss peaked between 1830 and 1880 with clear-cutting and agricultural expansion (Fig. 2) and 80% of old-growth or pre-settlement forests were lost (D. R. Foster, 1992). Reforestation of New England has been occurring for the last 150 years as the region urbanized. It hit a plateau in recent years as development has begun to expand into reforested areas (D. Foster et al., 2017).

The natural beauty and stark landscape of the White Mountains made it a tourism destination beginning in the 1840s (Gosselin, 1995). The completion of a railroad through Gorham, NH in 1851 drove a boom in tourism, as the number of visitors jumped from hundreds...
to tens of thousands per year (1995). The forest ecology and fresh mountain air of the region helped birth the tourism industry in New England. Throughout the 19th and into the 20th century, the White Mountains and other undeveloped areas across New England became popular vacation spots for people looking to escape the bustle of city life in Boston or New York (Wallace, 1995). Wealthy people came to experience the beautiful nature and undeveloped wilderness, especially at the start of the twentieth century, with a shift in public consciousness towards an appreciation of the sublime. This shift in public consciousness prompted a move towards preservation of forests. In fact, it was wealthy tourists who saw the environmental damage that logging was wreaking on the forests of central New Hampshire and began organizing to create the White Mountain National Forest (Govatski & Johnson, 2013). Although the nature of tourism has shifted over the years, the White Mountains still fulfills this purpose today. The White Mountains see 6 million or more visitors each year (White Mountains, 2018). Current reliance on tourism could be damaging to the forest as the very visitors that are a reason for its preservation are placing stress on the environment with sheer numbers. This was made evident in 2020 as COVID-19 closures of other tourist destinations brought thousands of visitors to the White Mountains, effectively overwhelming capacity and staff (Hurley, 2020; Leclerc, 2020). The large number of people, many of who were new to outdoor recreation, overwhelmed trail capacity. This led to trampling of undergrowth and alpine vegetation along the trails, as well as trash and even human waste being littered, all of which is damaging to the forest ecosystem.

The study site for this research is the Hubbard Brook Experimental Forest (HBEF) in the White Mountains region of central New Hampshire. Historical records show that pre-settlement era HBEF forest was composed of *Acer saccharum* (sugar maple), *Betula alleghaniensis* (yellow
birch), and *Fagus grandifolia* (American beech) (Chittenden, 1905), although historically there was less *A. saccharum* and more *Picea rubens* (red spruce). *P. rubens* was selectively harvested in the 19th century which led to its decline. The modern forest at HBEF mostly dates to the early 20th century as heavy logging occurred throughout the valley in the first two decades (Schwarz et al., 2001). The 1938 hurricane was the other major disturbance that shaped forest composition and structure of the valley (Peart et al., 1992). Although age and size distribution varies across the White Mountains, the forest at HBEF is similar in composition to a nearby remnant of old-growth forest (Schwarz et al., 2001). I am investigating this particular forest and region to see how modern changes are playing out at the study site.

2.5. *Climate migrants & seedling survival*

The arrival of new seedlings could be the first sign of a shift in species composition for northern temperate forests because of climate change. These tree species that are moving northwards into new ecosystems are considered climate migrant species; essentially, they are expanding their range and entering a new ecosystem as a result of changing climatic conditions. Climate migrants differ from invasive species in that they are not introduced through human behavior and they are native to a nearby geographic area (Urban, 2020; Wallingford et al., 2020). A single climate migrant does not spell disaster for forests, but climate change is leading to many new interactions between species (Simberloff, 2000). Thinking about seedlings (or other new plants) becomes particularly important in the context of climate change and shifting range dynamics. The introduction of several climate migrants, all at once, has unknown consequences for biodiversity (Scheffers & Pecl, 2019). Research is already showing the northward movement of various tree species (Woodall et al., 2009). Looking at overall migration patterns in tree
species across the United States requires careful selection of which species to monitor (Woodall et al., 2010). My study is considering the fate of seedlings on the edge of this northern range migration.

The emerald ash borer has not yet been detected at HBEF, although it is in close vicinity (Emerald Ash Borer Management Zones, 2021) and expected to begin impacting the forest in the next few years. Hemlock woolly adelgid has also not been detected, but is expected to slowly wipe out *T. canadensis* in the forest over the next several decades. *T. canadensis* will die out slowly, thinning with the increase of hemlock woolly adelgids. But at HBEF the forest succession is even more unclear, as *F. grandifolia*, which has been a successional species in other locations, is struggling as a result of beech bark disease (Hane, 2003). Disturbance regimes have important implications for many environmental factors essential to seedling survival and growth.

Climate change is dramatically altering the overall global climate, with much larger impacts during the winter season (Burakowski et al., 2008). A study by Contosta et al. (2019) looking at meteorological data across the northeastern United States over the past century found a decrease in snow and cold weather with time. This has conflicting implications for tree health. It is a benefit in terms of less fine root mortality and nutrient loss as a result of winter frost (N. Cleavitt et al., 2008), but is a detriment in terms of the increase in pests and northward movement of invasive species. A survey of just HBEF detected changes in climate variables over the last 50 years with similar patterns (Hamburg et al., 2013). Climate, and particularly winter climate, can have major impacts on seedling survival. The depth of snowpack can determine whether seedlings are crushed or if the soil is frozen and seedlings die because of root damage (Harrison et al., 2020). These climate variables with interannual variation likely impact seedling
survival. The multitude of direct and secondary impacts of climate change make the net change on forest composition extremely uncertain.

2.6. **Northern red oak**

![Map of current *Quercus rubra* species range in the United States from Forest Inventory and Analysis data (Peters et al., 2020).](image)

**Figure 3.** Map of current *Quercus rubra* species range in the United States from Forest Inventory and Analysis data (Peters et al., 2020).
*Quercus rubra* (northern red oak) is predicted to migrate into northern hardwood U.S. forests with the impacts of climate change, and the success of its spread is potentially tied to shifting seasonality and canopy gaps created by invasive pests (Richardson et al., 2006; van Doorn et al., 2011). Climate change is placing many pressures on northern temperate forests including invasive pests, warming temperatures, increased frequency of extreme weather events, fire, and more (Dale et al., 2001). The spread of *Q. rubra* is driven by many of these factors. As a novel component of northern hardwood forests (Fig. 3), the migration of *Q. rubra* may create more pressure on other tree species that are historically a component of the forest and are already experiencing stress.

Much previous research on the impacts of anthropogenic changes to forests relies on a forestry perspective: essentially how to manage the species composition and age distribution of a forest to facilitate regeneration and growth (Bolton & D’Amato, 2011; Coates, 2000; Dey et al., 2012; Kern et al., 2013). But there is less research into how unmanaged forests are handling natural (or climate change induced) shifts in species. The spread of *Q. rubra* is closely tied to changes in other tree species within the forests such as *Tsuga* (hemlock) or *Fraxinus* (ash), which will likely die out in the coming years with the spread of hemlock woolly adelgid and the emerald ash borer (Eschtruth et al., 2006; Kovacs et al., 2010) creating potential opportunities for *Q. rubra* to become an established novel species within the forest. It is still difficult to predict how *Q. rubra* will fare in this northern location. Increased light availability once *T. canadensis* die out does not ensure space for more species (Kern et al., 2013). Previous studies have identified light as the most significant predictor of *Q. rubra* seedling success (Brose, 2011; Frey & Ashton, 2018; Major et al., 2013), but others have found nutrient availability, local competition, negative density dependence, or initial seedling size to be equally important.
variables when considering survival (C. D. Canham et al., 1996; Coomes & Grubb, 2000; Hartman et al., 2005; Willis et al., 2015). The Janzen-Connell hypothesis which expects seedlings to have higher survival rates when near heterospecific tree species (Connell, 1971; Janzen, 1970), has been shown to hold true in temperate ecosystems (Comita et al., 2014), but it is still only one factor on survival. Predicting the future disturbance regimes of northeastern forests is difficult which further complicates any understanding of the potential establishment of *Q. rubra*.

### 2.7. Importance of studying seedlings

On a broader level, studying *Q. rubra* seedlings is a starting point for studying regeneration, the establishment and survival of the species within the forest tree community. The migration of tree species is controlled by seedling growth and survival. Seedling regeneration is crucial to understanding the future of our forests, but little research is done on the seedling layer of forests. Most forest ecology research instead focuses instead on recruitment within forests for saplings that have crossed the 10 cm diameter threshold. These saplings are a more immediate resource to foresters and conservationists. But seedlings are a measure of future population recruitment. Understanding the environmental factors and relative importance of those factors to the survival of *Q. rubra* seedlings offers insight into other species and informs the factors that influence the early stages of establishment of this novel species.

Although seedlings have little immediate influence on the mature tree population growth rate, they are the eventual future of the forest. Seedlings typically experience high rates of mortality as a result of various individual and environmental stressors. But if they are experiencing unexpectedly high or low mortality rates for new reasons (Hartman et al., 2005),
that could be useful information both for understanding the changing forest ecology of the region and for more practical management concerns in the New England region.

2.8. Questions

The main question my thesis will address is how do different environmental factors and seedling characteristics related to growth and damage influence the survival of *Q. rubra* seedlings at the northern edge of its range? I will look at this on a local scale by comparing *Q. rubra* seedling survival across sites of varying forest composition at a study site in northern New Hampshire. This study addresses this question at two temporal and spatial scales: first on a plot level for a single year to understand how the environmental factors including light and competition might influence seedling survival, and then on an individual seedling level across a decade to understand ontogenetic characteristics describing individual development as determinants of future growth and survival.

These two scales address different angles of hypotheses for patterns in seedling survival. The plot level scale is looking at localized microsite conditions, such as local competition and substrate, that might also contribute to *Q. rubra* seedling survival. But individual ontogenetic data such as seedling age, leaf number, and health can help to establish the size of the acorn and the presence of any pathogens influencing seedling growth. I am also incorporating winter climate data to elucidate the potential impact these annually variable factors have on seedling survival. Seedlings with a larger initial size and leaf number indicates they started with more resources, essentially a larger acorn. Large-seeded species such as *Q. rubra* tend to have first year seedlings that are somewhat insensitive to changes in abiotic factors as they have more...
initial resources from the seed (García & Houle, 2005). This does not always hold true though, and other studies have concluded that light is an essential factor for *Q. rubra* seedling survival (Brose, 2011; C. D. Canham et al., 1996). These various factors are heavily connected and can often exert pressure on each other (Fig. 4).

There is no simple answer for variables that influence seedling survival, but this study hopes to use multiple time scales and extensive environmental data to help better understand the interactions of these various factors and provide insight into how climate change may be shaping the survival of *Q. rubra* seedlings.
Figure 4: Concept map of different interacting factors on *Q. rubra* seedling survival
3. Methods

3.1. Site description

The Hubbard Brook Experimental Forest (HBEF) is a Long-Term Ecological Research (LTER) site with an area of 3,160 hectares located in the White Mountain National Forest in central New Hampshire (43° 56’N, 71°45’W) (Fig. 5). The topography is a bowl-shaped valley with Hubbard Brook running along the bottom (T. Fahey, n.d.). HBEF is primarily a northern hardwood forest, composed of a mixture of *Acer saccharum* (sugar maple), *Fagus grandifolia* (American beech), and *Betula alleghaniensis* (yellow birch). At lower elevations in the valley there are dense, nearly conspecific stands of *Tsuga canadensis* (eastern hemlock) (Battles, n.d.). At higher elevations it transitions to spruce-fir forests with a combination of *Picea rubens* (red spruce) and *Abies balsamea* (balsam fir). Soils are primarily well-drained, acidic spodosols formed from glacial till of varying thicknesses (Bailey, 2003). Elevation at HBEF ranges from 220 to 1,015 meters above sea level. Most of the plots for this study are in the 200-to-500-meter range, with higher elevations being present at the edges of the valley on Mt Kineo and Mt Cushman (Fig. 6).
Figure 5: Map of New England region showing location of Hubbard Brook Experimental Forest in New Hampshire, USA.

This location sits at the northern edge of the *Q. rubra* range and provides a study site where seedlings persist without the presence of parent trees (Peters et al. 2020). Climatic conditions typically mean long, cold winters and short, cool summers. Mean January and July temperatures are -9 °C and 18 °C respectively. HBEF has an average annual precipitation of 1400 millimeters, about one-third of which falls as snow (Bailey, 2003). Snowpack is typically present between December and mid-April.
3.2. Data collection

Since 2011, data has been collected monitoring the spread of *Q. rubra* seedlings into HBEF. Part of the long-term data collection at HBEF relies on a network of 371 Valley-wide plots that cover the entire valley (Fig. 6) apart from the experimental watersheds. Vegetation data has been monitored on these 0.05 ha circular plots since 1995 with resurveys every ten years (van Doorn et al., 2011). This study used a subset of these plots where *Q. rubra* seedlings were present and created transects using the Valley-wide plots as the center point. Transects cover a

Figure 6. Map of HBEF depicting study transects in orange dots.
2000 square meter area total, with a 100-meter distance extending to the east and west of the plot center and 5 meters to the north and south of this transect line (Fig. 7). Seedlings were typically defined as anything under 1 meter in height, but upon starting the *Q. rubra* project in 2011 seedlings of up to 2 meters in height were included. Seedlings are tagged with unique flags and mapped on the transect to be found in subsequent years. These transects are surveyed once a summer, usually in late August. New seedlings are also added to the dataset at this point. A small percentage of seedlings are not found each year as flags may be moved by animal activity, tree falls, or equipment activity along the road’s edge. We only included seedlings tagged in 2019 or earlier in the survival analysis as they must have survived at least one full year. Individual seedling characteristics such as size and damage from pests or pathogens may also influence mortality, so seedling height and leaf number have also been surveyed each year since 2011. Additional seedling variables including branch number and leaf damage have been collected since 2015.

We collected additional data on the environmental conditions in the summer 2019 field season on a subset of 10 of the original 26 oak seedling transects. These transects were in areas with hemlock-dominated canopies, defined as over 50% of the mature tree basal area in the plot. These transects were chosen to provide a baseline measure of the environment, as *T. canadensis* stands are predicted to decline within the next few years with the arrival of the invasive hemlock woolly adelgid (Paradis et al., 2008). This predicted increase in canopy openings with *T. canadensis* mortality will likely impact oaks greatly as their seedlings rely heavily on light availability.
Figure 7. Plot design for environmental data collected in summers 2019 and 2020. Displays the Valley-wide transect on which oak seedling survival data was collected and zooms into the plots along this transect at which the environmental variables were collected.

We collected this environmental data in association with the location of *Q. rubra* seedlings (n=338), although not directly adjacent to seedlings so as to prevent trampling during data collection. As (Fig. 7) shows, this data was collected within a 5 square meter distance of seedlings. These variables include visual estimates of the shrub cover to account for light availability on a local scale. During Summer 2020, shrub cover data was collected on the remaining 16 transects as preliminary data analysis showed that shrub cover may influence plot level survival of seedlings. We also visually estimated the types of ground surface cover within the plot as a way of measuring available ground surface for seedlings to sprout in. The litter layer is measured by inserting metal pins at the north and south of the 1-meter square plot and
counting the various species of deciduous leaves. We used a basal area prism (size 2.5cm) while standing at the center of the 1-meter plot to estimate the mature tree composition. Both the litter layer and basal area prism measurements were intended to be a metric of competition from the mature tree community.

Canopy photo data was collected at 20-meter intervals along the transects with an additional data point at the center of the transect. Some transects have fewer photos and data points because of natural features in the landscape such as a brook or gorge. 70% of intervals were missing at least one canopy photo due to the presence of a natural feature (e.g., gorge, stream) that prohibited data collection. Canopy photos were taken using a digital camera with a hemispherical fisheye lens.

3.3. Data processing

The photos were then analyzed using Gap Light Analyzer, an open source program developed specifically to calculate gap light transmission indices through a forest canopy from true-color hemispherical photos (Sallaway & Marinakis, n.d.). We used a quality control methodology, where every tenth photo analysis was performed again by a second person. If the results differed significantly, the initial person performed the analysis again.
**Figure 8.** Methodology for taking and analyzing canopy photos. (left) using hemispheric fisheye lens to capture photo, (right) example of photo being analyzed using Gap Light Analyzer program (Sallaway & Marinakis, n.d.).

### 3.4. Data analysis

I conducted data cleaning and statistical analysis in R (R Core Team, 2014). I used the tidyverse package (Wickham et al., 2019) to organize and subset the data, the ggplot2 package (Wickham, 2016) and the cowplot package (Wilke, 2020) to create figures, and the lme4 package (Bates et al., 2015) to run linear mixed-effects models. I used the raster package (Hijmans, 2020) to incorporate the spatial data and the lubridate package (Grolemund & Wickham, 2011) to organize date variables. I then used the broom.mixed package (Bolker & Robinson, 2020) to view model summaries and compare model parameters.
3.4.1. **Environmental plot-level**

To assess how the environmental variables impacted seedling survival I first visually examined pattern variation in the data using the ggplot2 package to create figures. I considered how the various environmental variables including canopy light availability, shrub cover, ground cover, and distance from mature parent trees, impacted plot level seedling survival across the 2018-2019 year. I used linear regressions to show the relationships between individual variables and seedling survival. I then looked at the overall relationship between plot-level seedling survival and environmental variables related to light and competition by using generalized linear models. I tested different combinations of canopy light availability, shrub cover, ground cover, and distance from mature parent trees to see which hypothesis of various independent variables was the most useful in modeling survival.

3.4.2. **Long-term seedling dataset**

To look at multi-year survival I created a binomial generalized linear model with a random intercept effect. I tested multiple models with differing combinations of fixed effects representing biotic seedling characteristics including age, number of leaves, leaf damage, and ratio of live to dead branches. Year interval (e.g., summer 2019 to summer 2020) was used as a random intercept effect to help account for interannual variability in the environment. I compared models using Akaike Information Criterion (AIC) values and determined the best-fit model to have the lowest AIC value of at least two below the AIC values of the other models.

3.4.3. **Winter climate**

To examine other environmental variables that might help to explain the random effect interannual variation, I tested for statistical correlations between the random effects and winter
climate variables. If a correlation was strong (greater than 0.5), then I tested whether adding the winter climate variable to the models improved the model performance (defined as a decrease in AIC by 2 or more). I tested multiple winter climate variables including median snow depth, median frost depth, the number of days with a snow depth over 10 centimeters, and the number of days with a frost depth over 5 centimeters.

Winter climate data has been collected on a roughly weekly basis throughout the winter season each year at HBEF from 1955 to the present (USDA Forest Service, 2020). Winter is defined by the ice-in and ice-out dates for neighboring Mirror Lake. The variables are collected along a snow course at ten points spaced in 2-meter intervals within a designated 0.25 hectare area adjacent to the standard rain gages. Snow depth is collected at all ten of these points along a snow course and averaged for each collection date. Frost depth is collected by probing the ground with a ski pole at two points adjacent to the ten sample points. At the first two points for which frost is detected, the soil is dug up and frost depth is measured. Data from one snow course is averaged for each collection date. There are 5 sampling sites throughout HBEF with data from 2011 to the present, one at Headquarters and the rest in the experimental watersheds.

This study used the data from Headquarters as that is closest in distance and elevation to the seedling transects. Between 2011 and 2020 there were an average of 19 collection days per winter. I also tested correlations between the data from Watershed 6 and Watershed 1 to ensure that the patterns at Headquarters did not diverge from other measurement sites. Median snow depth was rescaled to follow a distribution with mean of 0 and standard deviation of 1 as the data ranged from 7 to 728 millimeters, which is multiple orders of magnitude larger than the other variables included in the model.
4. Results

4.1. Valley-wide Patterns

I assessed *Q. rubra* seedling abundance across the HBEF in relation to distance to the valley entrance, which is also where parent *Q. rubra* trees are present (Fig. 9). Seedlings were more abundant near the valley entrance and decreased in number as the distance from the parent trees grew.

![Figure 9](image-url)

**Figure 9.** Distance to Hubbard Brook Experimental Forest entrance in meters by number of seedlings present in 2019 (n = 655) at a plot transect (n = 41) level. More seedlings can be found closer to the eastern valley entrance (where parent trees are present).
To get a better sense of how the *Q. rubra* seedling survival data is changing across this 9-year time scale from 2012 to 2020, I plotted the proportion of seedling survival by the year (Fig. 10). The average percentage of *Q. rubra* annual seedling survival between 2012 and 2020 was 78% of seedlings. There is a slight downward trend in annual seedling survival with time from 83% in 2012 to a low of 70% in 2019.

![Graph showing seedling survival proportions over years](https://via.placeholder.com/150)

**Figure 10.** Proportion of *Q. rubra* seedling survival (between 0 and 1) by year from 2012 to 2020. Seedlings (n = 1369) are re-measured every August.
Table 1: Mean and standard deviation of all seedling and environmental variables classified by the spatial scale at which it is measured (individual or plot-level). Variables are also sorted by the years in which that data was collected and number of observations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition and description</th>
<th>Years Covered</th>
<th>Mean ± SD</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>leafNumber</td>
<td>Number of leaves on seedlings</td>
<td>2012-2020</td>
<td>3.27 ± 2.08</td>
<td>3189</td>
</tr>
<tr>
<td>yearsAlive</td>
<td>Seedling age</td>
<td>2011-2020</td>
<td>3.86 ± 3.08</td>
<td>3811</td>
</tr>
<tr>
<td>seedHeight</td>
<td>Seedling height</td>
<td>2011-2020</td>
<td>16.3 ± 8.3 cm</td>
<td>2912</td>
</tr>
<tr>
<td>seedDamage</td>
<td>Categorical rating of leaf damage by percentage</td>
<td>2014-2020</td>
<td>1.3 ± 1.34</td>
<td>2373</td>
</tr>
<tr>
<td>brchLvD</td>
<td>Ratio of live and dead branches</td>
<td>2017-2020</td>
<td>0.77 ± 0.28</td>
<td>1381</td>
</tr>
<tr>
<td>PerCnpyOpenTotal</td>
<td>% of canopy opening to allow light through</td>
<td>2019</td>
<td>4.63 ± 1.48</td>
<td>91</td>
</tr>
<tr>
<td>ShrubCover</td>
<td>% of plot covered by shrubs and herbaceous plants</td>
<td>2019-2020</td>
<td>13.6 ± 18.4</td>
<td>463</td>
</tr>
<tr>
<td>Litter</td>
<td>% of plot covered by leaf litter or available ground surface cover</td>
<td>2019</td>
<td>88.3 ± 16.9</td>
<td>463</td>
</tr>
<tr>
<td>HQdist</td>
<td>Distance of plot transect from Hubbard Brook valley entrance</td>
<td>2019</td>
<td>1308 ± 676</td>
<td>41</td>
</tr>
</tbody>
</table>
4.2. Environmental variables & plot-level survival

Figure 11. Environmental variables by proportion of seedling survival between 2018 and 2019 at a 20-meter plot level (a) Percent shrub cover by proportion of seedling survival. Results correlation between the variables shows shrub cover is a marginally significant predictor of survival ($r(55) = 0.22$, p = 0.09). (b) Percent available ground surface cover for seedlings to sprout. (c) Distance to Hubbard Brook Experimental Forest entrance in meters by proportion of seedling survival between 2018 and 2019 at a plot transect level. Seedlings further west into the valley had higher rates of survival. (d) Percent canopy light availability. Light availability is not a significant predictor of seedling survival.
Initial correlations between environmental variables and plot-level survival from 2018 to 2019 showed that light and ground surface cover were not statistically significantly related to *Q. rubra* seedling survival. Shrub cover explained 21.7% of the variation in annual survival with marginal significance, where shrub cover was positively related to seedling survival \((r(55) = 0.22, p = 0.09)\) (Fig. 11a). As shrub cover in summer 2020 had a nearly identical trend with survival I did not include it in a separate analysis (see Appendix Fig. A3). Distance from valley entrance and parent trees explained 29% of the variation in annual survival (Fig. 11c), where survival was more likely farther from the entrance \((r(39) = .29, p = 0.07)\). For plot-level seedling survival, the model with the distance from the valley entrance \((t = 1.75, p = 0.09)\) and mean shrub cover \((t = 1.72, p = 0.09)\) was better than any other predictors, although not more useful than the null model \((t = 5.48, p\text{-value} < 0.001)\). Both distance and shrub cover had p-values below 0.1 and were marginally significant predictors of survival. A model including canopy light \((t = -0.95, p = 0.34)\) and ground cover \((t = -1.08, p = 0.28)\) did not explain significant variation in *Q. rubra* survival from 2018 to 2019.

4.3. **Individual seedling variables & individual survival for a single year**

For individual seedling survival over the same 2018 to 2019 period as the plot-level models, seedling age, leaf number, and leaf damage explained the most variation in *Q. rubra* seedling survival. Both age \((z(588) = 1.39, p = 0.17)\) and leaf number \((z(636) = 5.06, p < 0.001)\) had a positive relationship with survival, such that as age or leaf number increased, survival also increased (Fig. 12). Although leaf number had a stronger relationship with survival, thus explaining more of the variation in seedling survival. Leaf damage rating had a negative effect on survival such that greater leaf damage was associated with lower survival probability.
although only damage category 3 ($z = -3.48, p < 0.001$) and category 4 ($z = -2.64, p = 0.01$) were statistically significant (Fig. 12c). The Tukey HSD test post hoc comparisons showed mean survival for seedlings with the leaf damage category 3 ($m = 0.35, sd = 0.49$) and category 4 ($m = 0.13, sd = 0.35$) were significantly different from the damage category 0 ($n = 181, m = 0.71, sd = 0.45$), category 1 ($m = 0.78, sd = 0.41$), and category 2 ($m = 0.65, sd = 0.48$). However, pairwise comparisons within the lower damage categories (0, 1, and 2) and higher damage categories (3 and 4) respectively did not find significant differences. This strong difference in survival probability from the low damage categories to the high damage categories means seedlings with less leaf damage were more likely to survive than seedlings with significant leaf damage. Survival probability averaged 78% for leaf damage less than 50% but dropped to an average of 23.5% for leaf damage greater than that (see Appendix Fig. A2).
Figure 12. (a) Seedling age by survival between 2018 and 2019. Line of best fit shows a positive relationship between variables. Higher age is associated with a higher probability of survival. (b) Number of leaves on individual seedlings by survival between 2018 and 2019. More leaves is associated with a higher probability of survival. (c) Leaf damage rating category by seedling survival proportion between 2018 and 2019. Letters indicate statistical significance of pairwise comparisons from Tukey HSD post-hoc ANOVA test. Lower damage rating is associated with a higher probability of survival.
4.4. Multi-year survival modeling

4.4.1. Model selection

Figure 13. Distribution of seedling ages (n=1369) for each year interval. 2011 and 2015 were mast years for Q. rubra.

Table 2: Model selection for multi-year seedling survival model. All models being compared are generalized binomial linear models with year intervals as a random effect.

<table>
<thead>
<tr>
<th>Model for Individual Seedling Survival (survival)</th>
<th>AIC</th>
<th>p-value: leaf number &amp; (yearsAlive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>~ 1 (null model)</td>
<td>4134</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>~ leafNumber</td>
<td>553</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>~ yearsAlive</td>
<td>4108</td>
<td>(&lt; 0.001)</td>
</tr>
<tr>
<td>~ leafNumber + yearsAlive</td>
<td>553</td>
<td>&lt; 0.001 (0.19)</td>
</tr>
</tbody>
</table>

To examine the relationship between individual seedling survival during the 2011-2020 period and various ontogenetic variables about the Q. rubra seedlings, I created generalized linear models. All models included year as a random intercept effect, which captured the concept
that the survival probability in any one year might be higher/lower due to external effects across all the seedlings in that year (e.g., impacts of weather). I tested models with varying combinations of fixed dependent variables that I hypothesized could explain the variation in the independent variable of survival based on the relationships in the correlation matrix (see Appendix Fig. A1). I chose not to include the variables of live to dead branch ratio or leaf damage as they only covered 3 and 6 intervals respectively (see Appendix Fig. A2), and the more data included in a multi-year survival model, the better chance of accuracy. As age and leaf number had both been statistically significant when considering a single year of survival in 2018-2019 (Fig. 12), I hypothesized that the same would hold true when looking at survival across multiple years. I compared models using Akaike’s Information Coefficient (AIC) to test the usefulness of the model, and the one with leaf number alone or leaf number and age had the same lowest AIC value (Table 2). I did find seedling age to be a statistically significant predictor in a single model \((z(3808) = 5.12, p < 0.001)\), although not in the combined model with leaf number \((z(3185) = -1.30, p = 0.19)\) (Table 2). Age is also highly correlated with seedling height, which was a variable not included in the models \((r(577) = 0.32, p < 0.001)\). In the final model, leaf number was the only statistically significant predictor of seedling survival \((z(3185) = 16.8, p < 0.001)\). But both variables had a positive effect on survival, such that a higher leaf number and an older age were associated with a higher proportion of survival (Fig. 14). Leaf number and age are not strongly correlated \((r(3187) = 0.2, p < 0.001)\), so in order to estimate the combined effects of the two variables I chose the model that included both variables. A model with leaf number as the only predictor of survival might be just as effective, for example if it was impossible to determine seedling age in field conditions.
4.4.2. Visualization of final model

**Figure 14.** (a) Number of leaves on individual seedlings by survival (n = 1369) for multiple year intervals between 2012 and 2020. Line of best fit shows a strongly positive relationship between variables. More leaves are associated with a higher probability of survival. (b) Seedling age by survival (n = 1369) for year intervals between 2012 and 2020. Line of best fit shows a positive relationship between variables. Higher age is associated with a higher probability of survival.
4.5. Winter climate data

Figure 15: Winter climate variables summarized by year, (a) median frost depth in millimeters by year, (b) median snow depth in millimeters by year, (c) sum of days that winter with a frost depth over 50 millimeters, (d) sum of days that winter with a snow depth over 200 millimeters. Winter is defined by the dates of first and last frosts. The winter year refers to spring of the named year and the fall of the previous year (winter 2013 = 2012-2013). (n=170) Dotted line on a, c, and d refer to the average value across all winters.
Figure 16: Random effect values by winter climate summary variables (n=17), (a) median frost depth in millimeters by year, (b) median snow depth in millimeters, (c) sum of days each winter with a frost depth over 50 millimeters, (d) sum of days with a snow depth over 200 millimeters. Random effect values from the survival model including age and leaf number of individual seedlings (Table 2). Winter is defined by the dates of first and last frosts. The winter year refers to spring of the named year and the fall of the previous year (winter 2013 = 2012-2013).
To more closely investigate the variation in year-to-year survival captured by the random effect of year on *Q. rubra* seedling survival, I looked at winter climate variables at HBEF for the same 2012 to 2020 period. I specifically looked at frost depth and snow depth variables, as these are hypothesized to potentially influence seedling survival. There was significant interannual variation in these variables over the 9 years (Fig. 15). When looking at these variables in connection to the random effects of the year interval from the generalized linear model, I found median snow depth and number of days with a frost depth over 50 millimeters significantly explained interannual variation in seedling survival (Fig. 16). This demonstrated that snow depth and frost depth explain much of the variation in year as a random effect on survival. However, adding these variables back into the original generalized linear model as fixed effects had minimal impact on the overall model improvement.
5. **Discussion**

This study was designed to explain patterns in *Q. rubra* seedling survival patterns by investigating associations with ontogenetic factors at the individual level, environmental factors at the plot level, and winter climate factors at the regional level. Overall, individual biological variation in seedlings and herbivory pressure had more impact on survival than environmental factors such as light limitation and competition. Contrary to expectations, light availability had no noticeable impact on seedling survival while distance from parent trees and shrub cover had slight positive relationships with survival. The multi-year dataset showed that leaf number, or seedling size, was the most useful predictor of survival. Further analysis showed that most of the random variation in seedling survival from year to year could be explained by winter climate variables. I looked at individual seedling survival on a multi-year scale to help account for year-to-year variation and to aid in comparisons to other long-term studies. Environmental variables were considered on a plot level scale to study survival on a coarser scale. These environmental measures such as light availability were collected in hemlock-dominated areas so as to provide baseline measures of the *T. canadensis* forest before it is decimated by the hemlock woolly adelgid.

5.1. **Dispersal and distribution dynamics**

I found that as the distance from the parent trees grows there are fewer *Q. rubra* seedlings present. *Q. rubra* is known to rely mostly on animal dispersal of seeds, primarily by jays and gray squirrels (Bossema, 1979). With their heavy seeds, *Q. rubra* creates these distribution patterns as acorns fall close to parent trees and are picked up and moved by animals to germinate elsewhere (Sork, 1984). Both jays and gray squirrels distribute *Q. rubra* seeds as many of the
acorns they cache and forget end up germinating (Bossema, 1979; Smallwood et al., 2001). This dispersal method is beneficial to *Q. rubra* as buried acorns are also more likely to germinate than acorns that fall on the forest floor (R. J. Barnett, 1977; García et al., 2002). Multiple studies have shown that a layer of leaf litter or soil improves germination ability of acorns (Crow, 1988, p. 19; Gribko et al., 2002). Distance to parent trees also had a positive relationship on a single year of *Q. rubra* seedling survival between 2018 and 2019. In my study, seedlings further west into the Valley and farther from parent trees had 20% higher rates of survival than more easterly transects. This relates back to the Janzen-Connell hypothesis that seedlings at a distance from conspecifics or parent trees have an advantage for survival as competition and presence of pathogens is lessened (Connell, 1971; Janzen, 1970). Seed dispersal patterns could be a secondary explanation. The Bossema study found that jays often choose the larger acorns to distribute longer distances (1979). They also will not choose diseased or broken acorns to distribute. Thus larger, healthier acorns are often being distributed longer distances. If they selectively carry larger acorns longer distances, those acorns at a longer distance have a higher initial probability of survival as they have more initial resources (Moles & Westoby, 2004). In contrast, one study looking at seedling recruitment more broadly found the availability of suitable substrate for germination to be an essential predictor variable, suggesting that germination requirements are more important than dispersal in determining structure of temperate forests (Caspersen & Saprunoff, 2005). This may be true for northern temperate forests generally, but *Q. rubra* is a new addition to HBEF and there are many other added pressures on forests that could change structure and composition in new ways.

Studies have found distance and density dependent seedling survival is prevalent amongst plants across a broad range of latitudes, showing support for the Janzen-Connell hypothesis in
northern temperate forests as well as tropical locations (Comita et al., 2014; Hille Ris Lambers et al., 2002). In contrast to these individual findings, a meta-analysis study of the Janzen-Connell hypothesis did not find distance to be an important predictor of survival in temperate latitudes (Hyatt et al., 2003). My research agreed with results of the former studies perhaps because it focused on a single species rather than summarizing numerous studies and locations. The results of my generalized linear modeling showed that distance from valley entrance and parent trees explains 29% of the variation in seedling survival, $((r(39) = .29, p = 0.07)$. This distance dependent change in seedling survival is likely caused by a combination of seed dispersal patterns as well as natural enemies such as herbivores or pathogens (Comita et al., 2014; García et al., 2002). Survival was lower closer to conspecifics which is also where a higher prevalence of species and density specific herbivores or pathogens are present, as the distance from conspecifics increases so too does the distance from herbivores and pathogens. This negative density-dependent impact of conspecific adults on survival was supported in a recent study looking at Q. rubra at Harvard Forest in Massachusetts (Jevon et al., 2020).

My results saw a slight downward trend in overall seedling survival with time over the eight years from 83% in 2011-2012 to 72% in 2019-2020. The total average percentage of Q. rubra seedling survival between 2012 and 2020 is 78% of seedlings which does appear to be a high survival rate for seedlings. Although oak seedling survivorship has been found to vary significantly depending on site conditions. One study in Wisconsin found that the removal of tall understory vegetation increased 5-year Q. rubra and Q. alba seedling survival from less than 30% to over 90% (Lorimer et al., 1994). Another study in Wisconsin looking at Q. rubra seedling survival over a 6-year period found seedlings with a partial overstory to have a 56% survival rate (Crow, 1992). My results seem to fit better with other studies that have found Q.
rubra to have high rates of early survival regardless of environmental conditions (Gribko et al., 2002). This partly can be explained by large-seeded species such as Q. rubra having greater initial resources to start with and thus being more insensitive to variation in abiotic factors (García & Houle, 2005). Another possibility could be that Q. rubra seedlings have an advantage at the edge of the range as they are experiencing less intraspecific competition and are more suited to the changing climate than some of the more prevalent species (Fisichelli et al., 2014; Gribko et al., 2002; Iverson et al., 2004b, 2008). The survival rates within this study are higher than survival rates found at Harvard Forest over the same years, 62% between 2017 and 2019 (Jevon et al., 2020). This location in central Massachusetts is firmly in the center of Q. rubra range (Iverson et al., 2008), supporting our hypothesis that being on the edge of the species range gives seedlings an advantage.

Although seedling age distribution is influenced by mast years, survival rates do not seem unduly influenced by the influx of new seedlings following mast years in 2011 and 2015. As significantly more new seedlings are added to the population it skews the age values for the subsequent years to be much lower (Fig. 13). I did find age to be a statistically significant predictor of survival across a single year (from 2018 to 2019) (Fig. 12). This is unlikely to be related to mast years as there were none recently. But when looking at seedling age as a predictor of survival over multiple years, from 2012 to 2020, it was no longer significant (Table 2, Fig. 14). The positive relationship of older seedlings having a greater chance of survival remained the same, but age was not as strong of a predictor of seedling survival.
5.2. **Plot-level seedling survival interacting with environmental variables**

Most plot-level environmental factors relating to light and local competition were not significant predictors of survival. These findings of competition and overstory being unimportant to survival are concurrent with a study looking at first-year Q. rubra seedlings in hopes of improving regeneration (Bardon, 1992). Some correlation between *Q. rubra* survival from 2018 to 2019 showed shrub cover is a marginally significant predictor of survival ($r(55) = 0.22$, $p = 0.09$). I included shrub cover in the model and found a slight positive relationship between shrub cover and seedling survival, such that a higher percentage of shrub cover related to a higher survival (Fig. 1). This seems counterintuitive as I expected higher shrub cover to be associated with lower light availability and thus lower survival. But this is concurrent with my findings about canopy light availability having no impact on survival. Perhaps the already low light availability from the canopy meant that the additional change in light availability posed by shrub cover had no significant effect on survival. Although one study found high shrub cover to have a significant negative effect on *Q. rubra* survival over a long term period (Kern et al., 2012), others have found shrub cover to provide protection against herbivory or frost thus improving survival (Buckley et al., 1998; Ostfeld et al., 1997; Valladares et al., 2016). Prolonged competition for resources could also lead to lower outcomes for seedlings which were not captured in the one year measurement of seedling environmental variables used in this study. Shrubs could also provide physical protection for seedlings, whether by reducing herbivory by hiding them from predators or by buffering against winter climate effects such as soil freezing (N. Cleavitt et al., 2008; García & Houle, 2005; Sullivan & Sullivan, 2008; Templer et al., 2017).

The percent available ground surface cover for seedlings to sprout was not an important predictor of seedling survival. A study by Crow (1992) tracking *Q. rubra* seedlings at a site in
Canada, did find microsite to be an important factor to survival. But available ground surface cover has more influence on *Q. rubra* establishment (Gray & Spies, 1997) than seedling survival as acorns that are dispersed to areas with no available ground cover are not able to germinate. Once seedlings have already germinated their survival is more likely dependent on nutrient availability or microtopography of the site (Crow, 1992; Hartman et al., 2005; Walters et al., 2014). Neither of which is captured by simplifying to simple ground surface cover available for seeds to germinate in.

This research did not find light availability to be a significant predictor of seedling survival. This is contrary to my hypothesis but could be explained by the low light levels we measured (only 3% to 8%) opening. The light measurements were only taken on plots with hemlock-dominated canopies, which are known to be dense and unlikely to allow light through (Catovsky & Bazzaz, 2000). Although previous studies have identified light as the most significant predictor of *Q. rubra* seedling success (Brose, 2011; Frey & Ashton, 2018), they likely encapsulated a wider range of light availability. And light is not the agreed upon most significant predictor of seedling success within the literature (C. D. Canham et al., 1996, p. 199; Hartman et al., 2005; Willis et al., 2015), with some studies finding a large variance in species responses to light availability (Beaudet & Messier, 1998; Catovsky & Bazzaz, 2000). It is also important to consider that this comparison of survival and light availability was looking at only a single year. Light may have a greater impact on survival over time as seedlings consistently have less ability to grow. One study found that early oak seedling survival was not limited by low light availability and only began to show a decline in survival when tracked over multiple years (Crow, 1992). Another study in managed forests in German found *Q. rubra* regeneration to be uninhibited by low canopy light availability (Major et al., 2013).
I also hypothesize that there would be a difference in survival with relation to light availability if sites with hemlock canopy and hardwood canopy are compared. Light availability could increase substantially if *T. canadensis* declines in the future with the arrival of hemlock woolly adelgid at HBEF. As *T. canadensis* tends to grow in dense conspecific patches its decline has been found to significantly increase light availability and create large canopy caps (Orwig & Foster, 1998). In southern New England *Betula lenta* (black birch) appears to be most directly taking over in the absence of *T. canadensis* (Orwig & Foster, 1998; Stadler et al., 2005). But forest species composition is different in northern New England making *Q. rubra* an important contender to succeed *T. canadensis*. In particular, *Q. rubra* seedlings’ preference for higher light environments (Buckley et al., 1998) and increased growth in higher light (Brose, 2011) could allow them to outcompete the more shade-tolerant species (Hartman et al., 2005).

5.3. **Individual seedling survival interacting with ontogenetic variables**

Individual seedling characteristics (leaf damage, leaf number, and age) are significant predictors of survival from 2018 to 2019. Comparing the models by AIC values showed that the model with age, leaf number, and leaf damage was the most useful in predicting *Q. rubra* seedling survival. Both age and leaf number had a positive relationship with survival, such that as age or leaf number increased, survival also increased (Fig. 12a, Fig. 12b). Leaf damage had a strong negative relationship, a higher leaf damage rating was associated with a lower survival rate. This was particularly true for higher leaf damage categories, seedlings with less than 50% leaf damage all survived (*see Appendix Fig. A2*). Studies have shown leaf damage weather from folivory or mechanical disruption to be prevalent among *Q. rubra* seedlings (García & Houle, 2005). Although there is intraspecific variation in responses to herbivory and potential recovery,
it is still extremely damaging to seedlings (Byington et al., 1994). Leaf number could be considered a proxy for size which has been shown to be a major predictor of seedling success particularly amongst large-seeded species like *Q. rubra* (García & Houle, 2005; Moles & Westoby, 2004).

### 5.4. Modeling multi-year survival

Similar to the single-year survival modeling, I found a positive relationship between leaf number and age with seedling survival (Fig. 14). This matches my hypothesis that the trends in variable relationships with survival would be consistent between the single year and the multi-year modeling. But the final results were not the same. Leaf number was a far more consistent predictor variable when modeling survival over the 9-year period. One study in North Carolina found that *Q. rubra* seedling survival was only marginally linked with age over a 6-year study period (Beck, 1970; Crow, 1992). Age could also have confounding variables that complicate using it as a predictor of survival. And unlike species such as *Betula* that use wind dispersal and thus have small seeds with high rates of early seedling mortality, *Q. rubra* and most other large-seeded species have high initial rates of seedling survival (García & Houle, 2005).

Height, like leaf number, is another size related variable that has been shown a useful predictor of seedling success (Jevon et al., 2020). Age and height are very highly correlated variables, with height explaining 32% of the variation in age, so I ultimately decided to only include age in the final model ($r(577) = 0.32, p < 0.001$). Height can be a somewhat subjective measure that varies depending on the angle at which the seedling is measured from, the sponginess of the ground cover, or even between field researchers. The depth of leaf litter can also obscure the base of seedlings, leading to false equivalencies in height measures. Variation in
height is likely connected to biomass allocation, which could also cause variation in survival. Leaf number is a different metric considering aboveground biomass allocation of the seedling that might help account for some of that variation within each age class.

5.5. Interannual variation in frost depth, snow depth and seedling survival

Median snow depth over a winter had the most significant (negative) correlation with survival. Both median and snow depth and number of days per winter with a frost depth over 50 millimeters were useful in explaining the variation from year to year. Adding them to the model greatly reduced the random effect variation of the year variable, but it did not improve the overall usefulness of the model. This does make sense on some level, as winter climate has a negative relationship with survival, but not as great of an impact on individual seedling survival as leaf number which is seedling specific. This is perhaps because winter climate variables were generalized to the entire study site and across an entire year. They may not capture some of the fine-scale variation presented by microtopography which could interact with snow depth, or simply are not useful for considering individual seedlings. Depth of snowpack has a complex relationship to seedling survival. On one hand, a deeper, consistent snowpack provides insulation against soil freezing which leads to seedling roots freezing which is extremely damaging as it leaves them unable to uptake water or nutrients the following growing season (N. Cleavitt et al., 2008; Tierney et al., 2001). This explains the strong negative relationship between the number of winter days with a soil frost depth and seedling survival. Climate change brings further complications as most predictions for the northeastern United States foresee increasingly variable winters that will increase the number of freeze/thaw cycles making soil freezing and subsequent fine root damage more common (Harrison et al., 2020; Templer et al., 2017). But
depth of the snowpack also displayed a negative relationship with survival, as snowpack increased survival went down. A deeper snowpack could have the potential to crush or break seedling stems. We might expect snowpack to provide protection from herbivory by rodents or white-tailed deer which is a common limitation to *Q. rubra* regeneration (Dey, 2014; Kern et al., 2012; Ostfeld et al., 1997). But this deeper snowpack could also provide insulation for rodents and force them to tunnel under the snow thus allowing for greater herbivory (Korslund & Steen, 2006). There is a complex interaction between winter snowpack and herbivory on seedling survival that is likely to shift with climate change (Sanders-DeMott et al., 2018).
6. Conclusions

Overall, *Q. rubra* seedling survival appeared most influenced by ontogenetic variables such as leaf number over environmental variables such as light. When looking at seedling survival over multiple years, leaf number is the most useful predictor of survival. Age also has a positive relationship with survival, older and larger seedlings have a higher chance of survival. Modeling the winter climate variables of snow depth and frost depth explained most of the year to year variation in survival although the relationship was not as expected. Higher snow depth was associated with lower seedling survival which could be explained by the mechanical damage to seedlings or the greater presence of rodents under a higher snowpack. This research does not find local environmental factors including light availability and local competition to be significant predictors of seedling survival. I do feel this should be explored further by expanding both the spatial and temporal scale of the dataset. Limiting the study to just the hemlock-dominated areas restricts the variation in the environment, perhaps obscuring some of the relationships between local environmental variables and *Q. rubra* seedling survival. And only looking at survival over a single year could throw off the results depending on other interannual variation. Studying the relationship between light availability and seedling survival in areas with greater canopy light available may provide insight into the future of *Q. rubra* as *T. canadensis* begins to die out. The arrival of hemlock woolly adelgid could mean an increase in *Q. rubra* in the region. But this research emphasizes the need to further understand the interaction of *T. canadensis* die-off with *Q. rubra* regeneration. Future research should also expand the dataset on biotic individual seedling factors such as leaf damage over time. These factors had significant impacts on seedling survival on a single year scale, so understanding whether those impacts remain consistent through time is essential to understanding seedling survival in the context of
recruitment. Some variables such as leaf damage may be more impactful to seedlings depending on their age or size. My research found leaf number which is a proxy for seedling size to be the most useful predictor of survival but expanding the spatial and temporal scale of this study may uncover new relationships between environmental and ontogenetic factors that determine seedling survival. Seedling survival cannot be explained by a single variable.
7. References


Bardon, R. E. (1992). *First-year survival and growth of northern red oak (Quercus rubra L.) seed and seedlings underplanted in upland hardwood stands in south central Iowa* (p. 8913738) [Master of Science, Iowa State University, Digital Repository].
https://doi.org/10.31274/rtd-180813-8096


https://doi.org/10.18637/jss.v067.i01


Natural & Cultural Resources.

https://extension.unh.edu/resources/representation/Resource003178_Rep4610.pdf


https://doi.org/10.1016/j.foreco.2005.05.041


https://doi.org/10.1038/nature00809


Janowiak, M. K., D’Amato, A. W., Swanston, C. W., Iverson, L., Thompson, F. R., Dijak, W.


https://doi.org/10.1111/j.1365-2656.2005.01031.x


https://doi.org/10.1016/j.ecolecon.2009.09.004


Seymour, R. S., White, A. S., & deMaynadier, P. G. (2002). Natural disturbance regimes in

https://doi.org/10.1016/S0378-1127(01)00572-2


https://doi.org/10.2307/1938075


Wilke, C. (2020). *cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2”* (R
package version 1.1.0) [Computer software]. https://CRAN.R-project.org/package=cowplot


8. Appendix

Figure A1. Correlation matrix of individual seedling variables for multi-year survival modeling. Leaf number is slightly correlated with seed damage and branch live/dead ratio.
Figure A2. Structure of seedling survival in relation to branch live/dead ratio and leaf damage variables. Cannot include branch live/dead ratio or seedling damage rating in models as the data are not evenly distributed. Best to include the most data possible (these start in 2018 and 2014 respectively).
Figure A3. Percent shrub cover by proportion of seedling survival between 2019 and 2020 at a 20-meter plot level on all 26 oak transects.