LINKING ENVIRONMENT, BIODIVERSITY, AND ECOSYSTEM FUNCTIONS IN FOREST ECOSYSTEMS

By

Yu Zhang

A Thesis Submitted in Partial Fulfillment of the Requirements for the Doctor of Philosophy of Science in Forestry

Faculty of Natural Resources Management Lakehead University

August, 2014 ©
# TABLE OF CONTENTS

1 CONTACT INFORMATION IV

2 ABSTRACT V

3 ACKNOWLEDGEMENTS VI

4 NOTE VII

5 LIST OF TABLES VIII

6 LIST OF FIGURES X

7 OVERVIEW IN BIODIVERSITY AND ECOSYSTEM FUNCTIONING 1

8 CHAPTER 1: MULTIPLE DRIVERS OF PLANT DIVERSITY IN FOREST ECOSYSTEMS 5

9 INTRODUCTION ................................................................................................................................... 5

10 MATERIALS AND METHODS ................................................................................................................ 7

11 Study area and dataset 7

12 Species diversity 8

13 Explanatory variables 9

14 Statistical analysis 10

15 RESULTS ........................................................................................................................................... 13

16 Canopy tree species richness and Simpson’s dominance index 13

17 Total plant species richness and understory stratum-specific richness 14

18 DISCUSSION ...................................................................................................................................... 16

20 CHAPTER 2: INDIVIDUAL SIZE INEQUALITY LINKS FOREST DIVERSITY AND PRODUCTIVITY 20

21 INTRODUCTION ................................................................................................................................... 20

22 MATERIALS AND METHODS .............................................................................................................. 22
CONTACT INFORMATION

Student contact information

Yu Zhang, Faculty of Natural Resources Management, Lakehead University
955 Oliver Road, Thunder Bay, ON, Canada P7B 5E1
e-mail: yzhang5@lakeheadu.ca
phone: (807) 346-7768

PhD supervisor contact information

Dr. Han Y.H. Chen, Faculty of Natural Resources Management, Lakehead University
955 Oliver Road, Thunder Bay, ON, Canada P7B 5E1
e-mail: hchen1@lakeheadu.ca
phone: (807) 343-8342

PhD supervisory committee contact information

Dr. Kevin Crowe, Faculty of Natural Resources Management, Lakehead University
955 Oliver Road, Thunder Bay, ON, Canada P7B 5E1
e-mail: kevin.crowe@lakeheadu.ca
phone: (807) 343-8348
ABSTRACT

The positive biodiversity and ecosystem functions relationship (BEF) has been well established, especially in controlled experiments. However, a coherent study to examine the multiple relationships in BEF in natural ecosystems is lacking. Using the observational data of Canadian forests, I aim to examine the species diversity-productivity relationships in natural forest ecosystems across forest strata, under the influences of the climate factors and local site conditions. I found the positive effects of energy on tree species diversity while accounting for potentially confounding evolutionary effects and show evidence for the effects of local site conditions and secondary succession on diversity. The positive species diversity effects on productivity were common across forest strata but to different extent while the canopy tree diversity had negative effects on understorey plant biomass. Furthermore, tree size inequality may be a central process for the positive diversity effects on productivity, and potentially as the shared mechanism in regulating productivity and species diversity simultaneously via interactions among individuals in natural forests. This study sheds lights to deepen our understanding of the key features of natural ecosystems. I highlight that relationships between biodiversity and ecosystem functions and services are multifaceted, subject to several covariates including climate, local site conditions, time since stand-replacing disturbances, and the reciprocal interactions between diversity and ecosystem functions and services. However, our findings call for the complementary experiments to evaluate those potential mechanisms and drivers.

Keywords Biodiversity-ecosystem functioning, species-energy relationship, richness, evenness, life-history traits, productivity, aboveground biomass, soil carbon storage, stand age, soil drainage class, climate, growth form, boosted regression trees, structural equation modeling
ACKNOWLEDGEMENTS

As with the writing of this thesis, there are many people to whom I owe a great deal of thanks. First and foremost, I wish to thank Dr. Han Y.H. Chen, my supervisor, for keeping me up to or exceeding the high performance standard required by the PhD program in Lakehead University. I cannot express enough thanks to my committee for their continued support and encouragement: Dr. Kevin Crowe and Dr. Chander Shahi.

My completion of this thesis could not have been accomplished without the support of others. First, I would like to thank John Daisley and Lane Gelhorn at Weyerhaeuser Timberlands, the Canadian Forest Service, and to Dr. E.H. (Ted) Hogg for sharing data used in this thesis. I would like to extend my special thanks to Dr. Michael Huston, Dr. Richard Field, Eric Searle, Amber Brant, and Dr. Anthony Taylor for their valuable comments. Financial supports from the Natural Sciences and Engineering Research Council of Canada (DG283336-09 and STPGP428641) and the Early Researcher Award from the Ontario Ministry of Research and Innovation are gratefully acknowledged. I also thank Forest Ecosystem Science Co-operative Inc. and the Natural Sciences and Engineering Research Council of Canada for a doctoral research scholarship.

Finally, I extend my deepest gratitude to my lovely wife, Weilun. Your encouragements during the rough time are much appreciated and fully noted. Knowing that you were willing to provide management of all our household activities while I completed my work was an incredible relief and comfort. For that, I owe you the rest of my life.
NOTE

All chapters were written individually according to varying publication requirements of selected peer-reviewed journals. Efforts have been made to integrate those chapters into one coherent thesis with caution. However, the styles for writing, references, and organizations of figures and tables may slightly differ between chapters.
list of tables

Table 1-1 Results from boosted regression tree analysis (BRT) of diversity indices. Predictors’ relative influences show the relative contributions of predictors to the accounted variation of each BRT model. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, NPP – net primary productivity, approximated by mean annual biomass increment of canopy trees (Mg ha\(^{-1}\) yr\(^{-1}\)), SA – stand age (yrs), PS – plot size (m\(^2\)), MH – management history (managed vs. unmanaged), TC – tree complexity, \(n\) – number of sample plots, and PE – model prediction error. Moran’s \(I\) global tests were conducted on the residual for each fitted model.

Table 2-1 Direct, indirect, and total standardized effects on aboveground biomass of canopy trees based on structural equation models (SEMs). Significant effects are at \(P < 0.05\) (*), < 0.01 (**), and < 0.001 (***)

Table 3-1 Results from boosted regression tree analysis (BRT) to test the effects of canopy tree species diversity on aboveground biomass and soil carbon storage. Predictors’ relative influences represent the percentage contributions of predictors in the accounted variation of each BRT model. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, SA – stand age (yrs), S – canopy tree richness, \(J'\) – canopy tree evenness index, FDis – canopy tree life-history trait index, TC – tree complexity, \(n\) – number of sample plots, and PE – model prediction error. All numerical variables were natural log transformed except for SDC.
Table 4-1 Results from boosted regression tree analysis (BRT) of aboveground biomass and soil carbon storage. (a) models with plant richness of each plant layer. (b) models with tree richness as predictor. (c) models with total plant richness as predictor. Predictors’ relative influences represent the percentage contributions of predictors in the accounted variation of each BRT model. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, SA – stand age (yrs), S – plant richness, TC – tree complexity, n – number of sample plots, and PE – model prediction error. All numerical variables were natural log transformed except for SDC.
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>O-1</td>
<td>A conceptual map demonstrates the hypothesized causal paths studies in this thesis.</td>
<td>3</td>
</tr>
<tr>
<td>1-1</td>
<td>Observed and predicted responses of diversity indices to growing degree days, climate moisture index, soil drainage class, and natural logarithm of stand age on the canopy tree layer. (a) Canopy tree species richness. (b) Simpson’s dominance index. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables.</td>
<td>14</td>
</tr>
<tr>
<td>1-2</td>
<td>Observed and predicted responses of diversity indices to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), and natural logarithm of stand age on species richness. (a) Total plant species richness. (b) Shrub layer richness. (c) Herbaceous layer richness. (d) Bryophyte richness. Soil drainage class from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables.</td>
<td>15</td>
</tr>
<tr>
<td>2-1</td>
<td>Univariate relationships between endogenous (dependent) and exogenous (independent) variables (n =448). All variables were natural log-transformed except aboveground biomass. Significant regression lines were plotted using linear regression. Parametric assumptions were checked. (a) $y = 101.2 + 153.1x - 150.2x^2$, $R^2 = 0.07$, (b) $y = 199.8 + 59.2x$, $R^2 = 0.13$, (c) $y = -1.42 + 0.612x$, $R^2 = 0.24$, (d) $y = -1343.7 + 608.4x -61.8x^2$, $R^2 = 0.24$; (e) $y = -2.4 + 0.281x$, $R^2=0.1$; and</td>
<td>28</td>
</tr>
</tbody>
</table>
(f) not significant. All fitted regressions are significant at $P < 0.001$. The assumptions of normality and homogeneous variance were validated for all fitted regressions ($P > 0.05$) with an exception of marginal violation of normality of the fitted regression in c ($P = 0.04$). **Figure 2-2** Structural equation models linking aboveground biomass (AGB) and species diversity. (A) Effects of species diversity, soil nutrient regime, and stand age on AGB. (B) The model with tree size inequality, represented by DBH variation, as the linking mechanism. The coefficients are standardized prediction coefficients for each causal path. Solid lines represent significant paths ($P \leq 0.05$) and dash lines for non-significant paths ($P > 0.05$). The path coefficient marked with ‘±’ indicates a nonlinear (quadratic) relationship. **Figure 2-3** An alternative model for the effects of tree size inequality represented by DBH variation on both aboveground biomass and species diversity. **Figure 3-1** Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to canopy tree diversity indices: canopy tree richness, canopy tree evenness index, and canopy tree life-history trait index. (a) Tree aboveground biomass. (b) Shrub and herb biomass. (c) Bryophyte biomass. (d) Total aboveground biomass. (e) Soil carbon storage. Scatter points are observed values plotted by the respective response and explanatory variables. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, SA – stand age (yrs), S – canopy tree richness, $J'$ – canopy tree evenness index, and FDis – canopy tree life-history trait index. All numerical variables were natural log transformed except for SDC. **Figure 3-2** Results of BRT models showing the observed and predicted responses to
growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC),
and stand age (SA). (a) Tree aboveground biomass. (b) Shrub and herb biomass. (c) Bryophyte
biomass. (d) Total aboveground biomass. (e) Soil carbon storage. Soil
drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly,
poorly, and very poorly drained, respectively. Scatter points are observed
values plotted by the respective response and explanatory variables. All numerical
variables were natural log transformed except for SDC.

Figure 3-3 Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to canopy tree diversity indices: canopy tree richness, canopy tree evenness index, and canopy tree life-history trait index. (a) Aboveground live biomass. (b) Aboveground dead biomass. Scatter points are observed values plotted by the respective response and explanatory variables. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, SA – stand age (yrs), S – canopy tree richness, J’ – canopy tree evenness index, and FDis – canopy tree life-history trait index. All numerical variables were natural log transformed except for SDC.

Figure 3-4 Results of BRT models showing the observed and predicted responses to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), and stand age (SA). (a) Aboveground live biomass. (b) Aboveground dead biomass. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical
variables were natural log transformed except for SDC.

**Figure 4-1** Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), stand age (SA), and species richness of forest strata (S). (a) Tree aboveground biomass. (b) Shrub and herb biomass. (c) Bryophyte biomass. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical variables were natural log transformed except for SDC.

**Figure 4-2** Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), stand age (SA), and total plant species richness (S). (a) Aboveground live biomass. (b) Total aboveground biomass. (c) Soil carbon storage. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical variables were natural log transformed except for SDC.

**Figure 4-3** Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), stand age (SA), and tree species richness (S). (a) Aboveground live biomass. (b) Total aboveground biomass.
(c) Soil carbon storage. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical variables were natural log transformed except for SDC.
Over the past 30 years, remarkable progress has been made towards understanding how loss of biodiversity affects the functioning, goods and services of ecosystems (Cardinale et al. 2012). Integration has become the ultimate trends in ecology with the advances in understanding the extent of various branches of ecology and the intrinsic links between them. Multi-causality, multi-directionality, and high-dimensionality between ecosystem functions and biodiversity, and the associated underlying ecological processes demand future studies to be constructed by considering concurrent processes/mechanisms simultaneously, preferably in multivariate analyses. For example, abiotic factors such as climate and physiochemical constraints, biodiversity such as richness/evenness of producer, decomposers, and consumer specie, and ecosystem function such as carbon and nutrient cycling are all interrelated in a coherent framework (Midgley 2012).

Great efforts of syntheses have been made to reconcile conflicting concepts in ecology in the form of theoretical modeling and meta-analysis. A number of influential meta-analyses published to address the maintenance of species richness and species stable coexistence (e.g., Mittelbach et al. 2001, Balvanera et al. 2006, Gillman and Wright 2006, Bartels and Chen 2010, Zhang et al. 2012) and the consequences of species loss in ecosystem function (e.g., Balvanera et al. 2006, Cardinale et al. 2006, Cardinale et al. 2007). However, the conventional approach in ecology has been challenged in several aspects because of the complexity of the natural ecosystem. First, many previous studies have focused on single processes at an isolated level of ecosystem, and have produced often ambiguous and conflicting results and views. Second, the interpretations of the results from manipulative experiments are often skewed with much to be desired in order to demonstrate the causal relationships between variables clearly, and to shorten the gaps between empirical results and
predictions of theories (Carroll et al. 2011). Many of those have not been adequately addressed.

Understanding the ecological processes that drive species diversity and co-existence remains a significant intellectual challenge to ecologists, particularly, as the negative impacts of global biodiversity loss become increasingly apparent (Hooper et al. 2012, Reich et al. 2012). There are urgent needs in ecology to reconcile the somewhat contrasting theories and empirical findings in order to yield meaningful and realistic knowledge in aid to facilitate the efforts in conservation of species diversity and stability of ecosystem function through diversified approaches, including manipulative experiment, field observation and sampling, and utilization of long-term experimental data collecting by various agencies and parties.

It has been increasingly recognized that ecosystem functions and its relationships with biodiversity, and abiotic environment are scale dependent, and changing over time (McGill 2010, Dornelas et al. 2013). Therefore, there are multiple highly dynamic processes along temporal and spatial gradients instead of equilibrium. This nature of ecosystems and its properties demand new approaches to be adapted by ecologists both as novel statistical methods and the analyses of observational data, usually covering large geographical areas and long period of time, impossible to be done in controlled experiments (Wardle et al. 2012). As a result, ecinfomatics has been called for as ecology as an interdisciplinary science evolving to a more accountable and data-intensive state to cope with the complexity of ecosystems (Michener and Jones 2012).
In this thesis, I aim to: (1) quantify the influences of climate and local conditions on plant diversity, (2) identify the patterns and drivers of diversity-productivity relationship in forests in multivariate space, by linking abiotic drivers, biodiversity, and ecosystem functions with the integration of the currently separate lines of studies, i.e., diversity-productivity relationship (DPR) and species coexistence, (3) quantify the relative importance of canopy tree richness, evenness, and life-history trait divergence over multiple ecosystem functions, and (4) explore the magnitude and pattern of biodiversity effects on ecosystem functions across forest strata. I included a conceptual map to visualize the hypothesized causal relationships studied in this thesis for the better clarity (Fig. O-1). The lack of comprehensive tests of multiple processes concurrently as they occur in nature, and the resulting over-simplifications in ecological studies have imposed major challenges yet great opportunity for ecologists to exploit this distinctive dynamics in forest ecosystem due to the large scale, long life-span, and the extremely complex stand dynamic. Therefore, this study can facilitate...
further investigations by identifying the patterns and processes of species diversity-
productivity relationships in natural forest ecosystems.
Introduction

Understanding the ecological drivers of species richness and co-existence has long been a central pursuit in ecology (Connell 1978, Huston 1979). It has become even more pronounced as the negative impacts of global biodiversity loss on the functioning of ecosystems become increasingly apparent (Hooper et al. 2012). Climate is considered one of the most important large-scale abiotic factors controlling the distribution of organisms and community structure across multiple spatiotemporal scales (McGill 2010, Midgley 2012). Strong climate-richness relationships have been reported for numerous taxa along latitudinal gradients (e.g., Francis and Currie 2003, H-Acevedo and Currie 2003, Hawkins et al. 2003, Currie et al. 2004, Ricklefs 2004). Species-energy theory, originally extended from species-area theory, proposes that areas of high available energy can support more individuals, enabling species to maintain higher populations and reduce extinction risks, thus promoting regional and local species richness (Wright 1983). The theory has been extended with multiple mechanisms to explain the strong explanatory power of energy, water, and water-energy balance related variables for predicting species richness (Currie et al. 2004, Ricklefs 2004, Evans et al. 2005). Since higher diversity in low latitudes could be attributed to the reduced extinction risk due to long occupancy of species and/or long time for speciation (Kozak and Wiens 2012), it remains debated whether the positive latitudinal climate-richness relationships are a result of energy variation.

The role of site local conditions has been less frequently included in previous analyses of climate-species relationships (Ricklefs 2004). However, site conditions affect local
resource availability and can be an important driver of plant species diversity (Roberts and Gilliam 1995, Chipman and Johnson 2002, Bartels and Chen 2010). Soil fertility is thought to influence the relationship between plant species diversity and climate in forest ecosystems (Holdridge 1971). Theoretically, plant species diversity is expected to peak at intermediate levels of environmental stress, i.e., the humped-back pattern of species richness along gradients of productivity (Grime 1973). The humped-back pattern of species richness along gradients of soil fertility and related net primary productivity (NPP) has been widely recognized (Huston 1980, Fridley et al. 2012), but the role of site productivity on patterns of community species diversity is still deeply debated (Adler et al. 2011, Fridley et al. 2012). Similarly, as predicted by the intermediate disturbance hypothesis (IDH) (Connell 1978), plant diversity peaks at intermediate frequencies of disturbance through succession because high frequencies of disturbance lead to dominance of disturbance-adapted pioneer species, low frequencies of disturbance lead to low-diversity communities of competitive-dominant species, whereas intermediate disturbances result in co-existence of both species groups, thus high species diversity (Connell and Slatyer 1977, Taylor and Chen 2011, Chen and Taylor 2012).

Richness in the canopy tree, shrub, and herbaceous layers of forest stands has been found to respond to actual evapotranspiration at large scales differently, attributable to local biotic interactions through the shading of canopy tree layers (Oberle et al. 2009). However, extensive tests on the relative importance of climate and local site conditions on species richness in local communities are rare, especially over large geographic areas. Many hypothesized mechanisms for species richness and co-existence remain empirically untested across global ecosystems (Gaston 2000, Dawson et al. 2011). Furthermore, little is known about the influence of environmental factors on species dominance (an inverse measure of evenness) in terms of relative abundance among constituent species in a community.
Here, we used nationwide forest inventory data to quantify the impacts of climate and local soil conditions on plant diversity across Canada's forests. Specifically, we predict that (1) plant species diversity increases with available energy, quantified by growing degree days within the growing season, as predicted by species-energy theory (Wright 1983); (2) species diversity decreases with aridity because the ability of plants to utilize available energy is limited by water availability (Evans et al. 2005); (3) species diversity changes along gradients of productivity, determined by soil fertility (Grime 1973, Huston 1980, Fridley et al. 2012), (4) species diversity is regulated by disturbance frequency and succession as predicted by the IDH (Connell 1978, Taylor and Chen 2011, Chen and Taylor 2012, Dornelas et al. 2013); (5) the relative importance of climate and local soil conditions on species diversity may vary between canopy layer and understorey layers because local site conditions may have stronger influences on understorey plant species (Chipman and Johnson 2002, Oberle et al. 2009) and canopy layer exerts strong controls on understorey layers (Bartels and Chen 2013). We use species richness as a primary measure of species diversity, but we also test how canopy tree species dominance responds to climate and local site conditions, which is rarely considered in previous similar studies.

**Materials and Methods**

**Study area and dataset**

We used Canada’s national forest inventory (NFI) dataset to study the entire range of forested ecosystems in Canada (Supplementary Fig. S1). The study area is situated between 53°25’ W and 134°46’ W longitude and 42°37’ N and 68°14’ N latitude. Elevation ranges from 4 to 2170 m above sea level with mean annual precipitation between ~200 mm and ~3100 mm; the mean annual temperature was between -11.2°C and 9.3°C. The current NFI dataset contains one measurement of 988 permanent sample ground plots measured during the period of 2000-2006 by all provincial agencies (Canadian Forest Inventory Committee 2004). The
Ground plots were randomly selected from 20 × 20 km grid photo plots taken across the entirety of Canada’s forests. Ground plots varied in plot size from 125 to 500 m² with the majority of plots ≥ 400 m² (Supplementary Fig. S2). Within each plot, vegetation was assessed based on vertical strata and plant growth forms. The canopy tree layer was defined as all tree stems ≥ 9.0 cm in diameter at breast height (dbh). The understorey was defined as plants < 1.3 or 2.0 m in height depending on provinces. The understorey was further classified into three layers: the shrub layer, which included all woody plants; the herbaceous layer, which included non-woody vascular plants; bryophyte layer, which included all ground-growing non-vascular plants. The canopy tree layer and total plant species were inventoried for the entire plot; the understorey layers, however, were assessed for either the entire plot or in a smaller subplots (Canadian Forest Inventory Committee 2004) (Supplementary Fig. S2), resulting in plot size variation within and among vegetation strata. All species diversity measurements were conducted within their respective plots or subplots.

Species diversity

Canopy tree species richness for each plot was acquired from counts of live trees by species aggregating different varieties within species. For canopy species evenness, we calculated Simpson’s dominance index, (the inverse of Simpson’s evenness index), by using the proportions of basal area for each constituent species. Simpson’s index ranges from 0 (infinite diversity) to 1 (monoculture). Total plant richness included all plant species growing on the ground including trees, shrubs, herbs, bryophytes, and lichens. The majority of plant species were not canopy tree species since canopy tree species accounted for less than 20% of total plant richness in 82.3% of the plots (Supplementary Fig. S3). We also calculated the species richness of understorey vegetation strata. However, plots (n = 170) in Quebec, Prince Edward Island, and New Brunswick were not included from stratum-specific analyses due to code absence for stratification. The shrub layer
definition also differed among provinces. The shrub layer was defined as woody plants that are less than 1.3 m in height in Alberta, Manitoba, Nova Scotia, and Northwest Territories \((n = 274)\). The remaining provinces defined shrubs as woody plants that are less than 2.0 m in height. We included all plots for shrub layers as defined by individual provinces in the analyses to avoid loss of information. The herbaceous layer included forbs, ferns, gramnoids, and saprophytes. Bryophyte layer included all ground-growing non-vascular plants, i.e., mosses, liverworts, and lichens.

**Explanatory variables**

To examine the influence of climate on species diversity, we derived climate data for each NFI plot from BioSIM software that generates long-term (1951-2010) scale-free climate data from geographic coordinates (latitude, longitude, and elevation) (Régnière and Saint-Amant 2008). We used growing degree days (GDD) (base temperature at 5°C) as a measure of overall available energy for plant growth and climate moisture index (CMI = mean annual precipitation-annual potential evapotranspiration) to represent aridity. Higher values of CMI translate to higher water availability for plants (Hogg 1997).

We used soil drainage class (SDC) as a measure of local site conditions and as an integrated measure of the overall site quality. Soil drainage class was determined by field surveys involving soil pit excavations. Similar to soil moisture regime and nutrient regime classifications (Chen et al. 1998, Chen et al. 2002), SDC classification considers multiple factors including: topographic position, organic layer depth, soil permeability, soil texture, soil thickness, and depth of water table (Taylor et al. 2000). Seven classes were used, from A to G, representing very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. We also calculated local site productivity as stand biomass of live trees divided by stand age. This was used as a crude estimate for net primary productivity (NPP), another aspect of site quality. Stand biomass was calculated by summing tree biomass
of stem wood, bark, branches, and foliage, estimated using Canadian tree species biomass
equations (Lambert et al. 2005), and then scaled-up to a per-ha basis.

To account for temporal dynamics of species diversity, stand age (SA) for each plot
was determined according to last stand-replacing fire date or by coring three dominant/co-
dominant trees of each tree species inside or outside the plot at the time of plot establishment.
With coring, SA was the average of ring counts from the tree samples of the species with the
oldest age, used as a conservative estimate of stand age (Senici et al. 2010). Due to variations
in plot size, which is positively associated with species richness (Wright 1983), plot size was
included in all statistical models to account for sampling area dependence of species diversity.
Similarly, because of potential effects of silvicultural activities on species diversity,
presence/absence of any management history was included as a predictor.

Statistical analysis

We used boosted regression tree analysis (BRT) to examine how species diversity is affected
by regional climate, local site conditions, and stand development across the large Canadian
forest biome. BRT resembles an additive regression model in which many simple regression
trees, generated using recursive binary splits based on the performance of a single predictor
variable at each split, are fitted in a stage-wise manner. With the introduction of stochasticity
(termed as bagging) and the division of data for model training and validation (termed as
cross-validation), BRT can achieve greater accuracy in predictions and less bias without
over-fitting. Based on the hierarchical structure of each tree model, interactions are
automatically considered in fitted models because the response to a lower level explanatory
variable depends on values of higher level explanatory variables within a tree model. The
relative influences of explanatory variables represent the percentage of variation of the total
variation accounted for by the BRT model.
BRT can handle the common problematic characteristics associated with observational data such as nonlinear relationships, missing data, multicollinearity among predictors, and violations of parametric assumptions, with desired accuracy in finding relationships between predictors and response variables (De'ath and Fabricius 2000, De'ath 2007, Olden et al. 2008). BRT has also been shown to outperform all conventional statistical methods except Bayesian models for predicting species distributions over broad geographic scales (Elith et al. 2006). Specifically, BRT introduces randomness and multiple model averaging with adjusted weights for each sequential model to reduce sampling bias associated with observational data (De'ath 2007).

We fitted all BRT models using the recommended values for BRT parameters: learning rate (0.005), bag fraction (0.6), and cross-validation (10). Learning rate may have a smaller value than 0.005, depending on the number of observations of the specific models to achieve adequate model fit (Elith et al. 2008). In addition, tree complexity (TC) was chosen as 1 (no interaction) and 2 (two-way interactions among predictors). We opted to report the BRT model with smaller TC if the models with higher TC did not improve prediction error (PE) considerably (i.e., < 1%) to avoid over-fitting. All analyses were performed by using BRT (Elith et al. 2008) with gbm R package (Ridgeway 2010). Because species richness is the count of number of species, a Poisson distribution of errors was used to model species richness response to predictors. Given that SA and NPP are strongly positively skewed, they were transformed by natural logarithm, which resulted in similar or slightly better models. To evaluate spatial structure of the BRT residuals, we applied global Moran’s I statistic, and plotted the correlograms to check the range and type of autocorrelation in various lag classes, using the R package spdep (Bivand 2013).
### Table 1-1 Results from boosted regression tree analysis (BRT) of diversity indices.

Predictors’ relative influences show the relative contributions of predictors to the accounted variation of each BRT model. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, NPP – net primary productivity, approximated by mean annual biomass increment of canopy trees (Mg ha⁻¹ yr⁻¹), SA – stand age (yrs), PS – plot size (m²), MH – management history (managed vs. unmanaged), TC – tree complexity, n – number of sample plots, and PE – model prediction error. Moran's I global tests were conducted on the residual for each fitted model.

<table>
<thead>
<tr>
<th>Diversity index</th>
<th>Error</th>
<th>TC</th>
<th>n</th>
<th>PE</th>
<th>$R^2$</th>
<th>Moran's I test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GDD</td>
<td>CMI</td>
<td>SDC</td>
<td>NPP</td>
<td>SA</td>
<td>PS</td>
</tr>
<tr>
<td>Canopy tree richness</td>
<td>43.2</td>
<td>21.9</td>
<td>16.5</td>
<td>-</td>
<td>17.6</td>
<td>0.3</td>
</tr>
<tr>
<td>Simpson's dominance index</td>
<td>36.9</td>
<td>22.7</td>
<td>15.6</td>
<td>-</td>
<td>23.9</td>
<td>0.9</td>
</tr>
<tr>
<td>Total plant richness</td>
<td>12.2</td>
<td>23.8</td>
<td>49.0</td>
<td>-</td>
<td>14.2</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>9.8</td>
<td>22.4</td>
<td>46.5</td>
<td>8.7</td>
<td>11.7</td>
<td>0.3</td>
</tr>
<tr>
<td>Shrub richness</td>
<td>56.6</td>
<td>3.0</td>
<td>6.3</td>
<td>-</td>
<td>5.2</td>
<td>28.9</td>
</tr>
<tr>
<td></td>
<td>55.6</td>
<td>2.2</td>
<td>5.2</td>
<td>6.5</td>
<td>5.6</td>
<td>24.9</td>
</tr>
<tr>
<td>Herbaceous richness</td>
<td>22.0</td>
<td>23.9</td>
<td>17.4</td>
<td>-</td>
<td>24.9</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td>16.8</td>
<td>19.0</td>
<td>16.1</td>
<td>18.3</td>
<td>18.8</td>
<td>8.0</td>
</tr>
<tr>
<td>Bryophyte richness</td>
<td>19.0</td>
<td>26.1</td>
<td>14.2</td>
<td>-</td>
<td>23.4</td>
<td>16.9</td>
</tr>
<tr>
<td></td>
<td>16.1</td>
<td>21.9</td>
<td>10.4</td>
<td>23.4</td>
<td>15.8</td>
<td>12.0</td>
</tr>
</tbody>
</table>
Results

Growing degree days, CMI, SDC, and SA accounted for 47%, 30%, and 49% of variation in
canopy tree species richness, Simpson’s dominant index, and total plant species richness,
respectively (Table 1-1). By contrast, the same set of predictors accounted for substantially
less variation (27% to 34%) in species richness for the shrub layer, herbaceous plants, and
bryophytes and lichens. Plot size contributed little to canopy diversity and total plant
diversity since there was little variation in plot size (Table 1-1), but plot size accounted for
large variation, ranging from 9% to 29%, in understorey stratum-specific richness due to
substantial plot size variation. Management history had minimal effects on most diversity
indices except some effects on herbaceous richness (Table 1-1). In all cases, we found no
significant spatial structure in the residuals (Table 1-1).

Canopy tree species richness and Simpson’s dominance index

For canopy richness, GDD and CMI were the strongest predictors, followed by SA and SDC
with 43%, 22%, 18%, and 17% relative influences, respectively (Table 1-1). Canopy richness
increased with GDD monotonically (Fig. 1-1a). Richness increased with CMI similar, but to a
lesser extent, to GDD (Fig. 1-1a). Species richness increased and then decreased with reduced
soil drainage. Species richness increased with SA, reached a peak at an intermediate SA
(approximately 70 years old), declined, but then increased in the oldest stands (Fig. 1-1a). By
contrast, Simpson’s dominance index decreased with increasing GDD and CMI, was highest
at the low and high drainage classes, and was lowest at about 70 years of SA (Table 1-1, Fig.
1b). A correlation analysis indicated that canopy tree species richness and Simpson’s
dominance index are strongly negatively correlated ($r = -0.76$, $P <0.001$).
**Figure 1-1** Observed and predicted responses of diversity indices to growing degree days, climate moisture index, soil drainage class, and natural logarithm of stand age on the canopy tree layer. (a) Canopy tree species richness. (b) Simpson’s dominance index. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables.

**Total plant species richness and understory stratum-specific richness**

For total plant richness, SDC was the strongest driver, followed by CMI, GDD, and SA with 49%, 24%, 14%, and 12% relative influences, respectively (Table 1-1). Total plant species richness increased with drainage up to drainage class E (imperfectly drained), and then slightly decreased (Fig. 1-1a).
Figure 1-2 Observed and predicted responses of diversity indices to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), and natural logarithm of stand age on species richness. (a) Total plant species richness. (b) Shrub layer richness. (c) Herbaceous layer richness. (d) Bryophyte richness. Soil drainage class from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables.

The relative influences of predictors (Table 1-1) clearly showed that richness of understorey layers responded to environmental factors differently compared to canopy tree species richness. The relative influences of GDD diminished from shrub, herbaceous, to
bryophyte layers as 56%, 22%, and 19%, respectively, whereas relative influences of
drainage and SA increased. Shrub richness responded to climate and site conditions similarly
as canopy richness did, but to a lesser extent (Fig. 1-2b). The richness of the herbaceous and
bryophyte layers showed distinct patterns compared to canopy richness (Fig. 1-2c-d).

Herbaceous richness correlated negatively with GDD, CMI, and SA, and peaked at an
intermediate SDC (Fig. 1-2c). Bryophyte richness decreased with GGD, increased with CMI
and SA, and was higher in both perfectly drained and very poorly drained than in other
drainage classes (Fig. 1-2d). The BRT models showed similar trends of richness of total plant,
shrub, herbaceous, and bryophyte layers, predicted by GDD, CMI, SDC, and SA with and
without NPP as a predictor (Fig. 1-2 and Fig. S1-4).

Discussion

Using forest inventory data that covers a wide geographical area across Canada’s forest
ecosystems, we show that plant species richness is controlled by climate, local site conditions,
and stand age; however, climatic influences have stronger effects on canopy tree species
richness than total plant species richness. Specifically, canopy tree species richness appears
predominantly controlled by energy and climatic water availability, whereas total plant
species richness, in which understorey plants account for 80-90% of all species in northern
forest ecosystems (Gilliam 2007), is controlled more by soil conditions.

Our analysis of canopy tree species richness provides support for the species-energy
theory along a wide longitudinal gradient. Unlike the positive species-energy relationships
found along global latitudinal gradients (Francis and Currie 2003, H-Acevedo and Currie
2003, Hawkins et al. 2003, Currie et al. 2004), where tropical ecosystems have a longer
evolutionary history than northern ecosystems (Benn and Evans 2010), the positive species-
energy relationship of our ecosystems resulted from longitudinal climatic variation, where all
studied ecosystems experienced the last glaciation and, thus, are of similar ages. This
suggests that reduced extinction risk due to long occupancy of species or long time for speciation (Kozak and Wiens 2012) is not an important mechanism for our observed broad-scale diversity patterns across Canada’s forests. Additionally, the magnitude of positive species-energy relationships was reduced when CMI fell into negative values, i.e., severe water-limitations on productivity (O’Brien 1998), providing support for aridity constraints on the realization of available energy for tree species diversity (Hawkins et al. 2003, Piedallu et al. 2013). Further, consistent with the understanding of environmental stress effects on plant diversity (Grime 1973), canopy tree species richness was higher in intermediate soil drainage classes where local water availability is neither in deficit nor too excessive.

Canopy tree species richness peaked at an intermediate stand age, approximately 70 years old, and there was an increase in oldest stands. These results are consistent with the prediction of IDH, i.e., both intermediate disturbance frequency and intensity promote species diversity (Connell and Slatyer 1977, Connell 1978, Svensson et al. 2012). The peak at intermediate stand age is attributable to canopy co-dominance of early- and late-successional species, qualified as the canopy transition stage of stand development following a stand replacing disturbance (Chen and Popadiouk 2002). The high species diversity in the oldest stands is likely a result of disturbances of intermediate intensity such as outbreaks of spruce budworm (Choristoneura fumiferana), forest tent caterpillar (Malacosoma disstria), and windthrow that create large canopy gaps and allow co-existence of both early- and late-successional species at late-successional stages as evidenced in our previous studies (Taylor and Chen 2011, Chen and Taylor 2012).

Environmental factors influenced evenness (inverse of Simpson’s dominance index) similarly to, but at a lesser extent than canopy tree species richness. The strong negative richness-dominance relationship suggests that species richness may be the most predominant aspect of biodiversity, although the underlying processes for the observed richness-
dominance relationship are not clear. We speculate that local negative plant-soil feedback through pathogens and herbivores (Mangan et al. 2010), and synchrony in density dependence via niche differentiation caused by environmental variability (Thuiller et al. 2007) may be responsible.

The strong positive species-energy relationship observed for canopy tree species was not found for total plant species richness, nor for richness of the herbaceous and bryophyte layers. Rather, local conditions had stronger influences than climate for understorey richness. The greater influence of local site conditions than regional climate on total plant species richness suggests that previous studies (e.g., Hawkins et al. 2003) may have overestimated the importance of climate on total species richness. Our results are consistent with previous studies in temperate (Oberle et al. 2009, Wang et al. 2009) and subantarctic and Patagonian forests (Speziale et al. 2010). The differences in relative influences between regional climate and local conditions for canopy tree species versus understorey plant species richness are attributable to differences in plant life forms and their associated life history strategy in forests (Chipman and Johnson 2002, Oberle et al. 2009, Speziale et al. 2010). The diversity of understorey plants, which are typically shade tolerant, a trait that limits direct impact of solar radiation on performance (Roux et al. 2012), is driven by available local water and nutrients and substrate heterogeneity (Bartels and Chen 2010).

Temporal changes in richness of total plant species and understory layers were substantially different from changes in canopy tree species. This inconsistency is not uncommon (Mackey and Currie 2001) because plant succession and coexistence cannot be attributed to a single mechanism, but rather the amalgamation of various interacting mechanisms (Shea et al. 2004, Dornelas et al. 2013). For instance, temporal change in plant diversity can be influenced by feedback from diversity effects on disturbance severity and frequency that tend to increase mean values of richness (Randall Hughes et al. 2007).
Alternatively, this inconsistency can be attributed to differences in resources requirements.

For example, dependence on light availability differs between canopy trees and understorey plants, such that the negative impact of canopy closure through time may be less apparent on understorey plant than canopy tree species richness (Chipman and Johnson 2002, Bartels and Chen 2010, 2013, Halpern and Lutz 2013)

Our results demonstrate that species diversity is regulated by multiple drivers in forest ecosystems. Canopy tree species richness increases with available energy in the absence of aridity, and is also affected by local site conditions and stand age. However, the influence of climate and local conditions on richness was found to differ between canopy trees and understorey plants in both direction and magnitude. This is possibly the result of differences in tolerance to stress and resource requirements between canopy trees and understorey plants, resulting from their different life history strategies, as well as feedbacks among vegetation layers. The strong correlation between canopy tree richness and Simpson’s dominance index and different responses to environmental factors among forest strata suggest that the specific role of climate and local site conditions on plant species diversity may vary depending on choice of diversity attributes and vegetation strata. Therefore, current ecological theory which focuses on presumed univariate relationships is often controversial and highly context-dependent in empirical studies because of the multivariate nature of ecosystems (Grace et al. 2012a).
CHAPTER 2: INDIVIDUAL SIZE INEQUALITY LINKS FOREST DIVERSITY AND PRODUCTIVITY

Running title: Individual size inequality and DPR

Introduction

Positive ecosystem function and species diversity relationships, particularly positive diversity-productivity relationships (DPRs), have been widely reported for experimental systems, leading to conclusions about profound negative impacts of biodiversity loss on ecosystem functions (Isbell et al. 2011, Zhang et al. 2012). Despite the critical importance of external validity for practical solutions to mitigate and adapt to diversity loss in natural systems (Naeem et al. 2012), diversity effects on productivity in natural systems remain unclear and hotly debated (Grime 1973, Adler et al. 2011, Fridley et al. 2012). Furthermore, the applicability of findings from controlled experiments remains controversial as it may fail to account for the complexity of natural ecosystems (Duffy 2009). Much research is needed to improve our understanding about the patterns and causes of observed DPRs in natural ecosystems.

Species complementarity, interpreted as a result of niche differentiation and facilitation, is regarded as the mechanism for the observed positive DPRs in experimental systems (Loreau et al. 2001). Complementarity effects are also important to maintain species diversity (Levine and HilleRisLambers 2009). It is theoretically plausible that niche differentiation and facilitation are the central processes for maintenance of species diversity and the positive DPRs. However, empirical evidence for niche differentiation and facilitation is scare due to the fact that niche differentiation and facilitation are multifaceted and operate at the individual level, and interspecific interaction strength change spatially and temporally (Clark 2010). As a result of
diversity among and within species and/or growth plasticity of individuals expressed through their interactions with immediate neighbors (Potvin and Dutilleul 2009, Brassard et al. 2013, Mueller et al. 2013), tree size inequality reflects resource partitioning and use efficiency within a site in natural forests (Man and Lieffers 1999, Chesson 2000, Yachi and Loreau 2007, Coomes et al. 2009, Clark 2010). We thus hypothesize that tree size inequality among all individuals is the mechanism responsible for the maintenance of species diversity and positive DPRs in natural forests.

The controversy between experimental and observational studies appears to arise from our limited understanding of the multiple mechanisms that simultaneously affect diversity and ecosystem functioning in natural ecosystems (Grace et al. 2012a, Tilman et al. 2012). For example, the nutrient regime of a habitat can strongly affect DPR, predicted by the multivariate productivity-diversity hypothesis (MPD) (Cardinale et al. 2009). Studies in natural forests also reveal that site conditions and stand age (Oberle et al. 2009) should be included when testing the proposed multivariate and potentially indirect relationships between diversity and productivity in forest communities (Tilman et al. 2012). It is not uncommon in previous DPR studies that well-documented coupling factors were neglected, leading to inconclusive results, especially in forest ecosystems. Furthermore, it is problematic to use only species richness to represent diversity because of the multifaceted nature of biodiversity (Purvis and Hector 2000) and high contributions of species evenness to productivity (Zhang et al. 2012).

Here we aim to examine the multiple relationships between aboveground biomass, species diversity, tree size inequality, stand age, and soil nutrient regime in the boreal forest by using structural equation models (SEMs) (Grace et al. 2012b). Specifically, we test the following paths: (1) positive DPRs are mediated through tree size inequality within each stand; (2) nutrient
regime influences the aboveground biomass, species diversity, and tree size inequality; and (3) stand age affects aboveground biomass, species diversity, and tree size inequality (Brassard et al. 2008).

**Materials and Methods**

**Study area and forest inventory data**

The data used in this study were from permanent sample plots in Saskatchewan, Canada, collected under stratified random sampling scheme to cover as extensive a range of stand ages and site types as possible. Timberline Natural Resource Group provided the data under a user agreement. We included (1) plots originating from wildfire, and not managed; (2) plots located at least 50 m from edges and 100 m from any highway; and (3) plots having an extensive soil survey to determine soil nutrient regime. The resulting dataset consists of 448 plots, mostly measured during the period of 1992-1999, and ranging from 52°30’ – 55°24’ N latitude and from 102°36’ – 108° W longitude (Supplementary Fig. S1). Plots varied in size, 0.06 (n = 78) and 0.08 ha (n = 370). Common tree species in this region included *Pinus banksiana* Lamb., *Picea mariana* (Mill.) B.S.P., *Picea glauca* (Moench) Voss, *Abies balsamea* (L.) Mill., *Larix laricina* (Du Roi) K. Koch, *Populus tremuloides* Michx., *Populus balsamifera* L., and *Betula papyrifera* Marsh.

**Variables used in analyses**

We calculated aboveground biomass (AGB, Mg ha⁻¹) as a surrogate of stand productivity (Supplementary Table S2-1). We estimated aboveground biomass of live trees for each plot using species-specific allometric equations based on diameter at breast height (DBH) developed specifically for Canadian boreal tree species (Lambert et al. 2005).
We chose Shannon’s index as a measure of species diversity to account for species richness and evenness, two of the important aspects of diversity in DPR studies (Maestre et al. 2012). We consider Shannon’s index to be the realistic measure of species diversity in species poor ecosystems such as boreal forests; 85% of 448 plots consisted of less than 5 tree species, with large variability in species evenness. Alternatively, we would choose species richness and evenness as two separate predictors, but no current diversity index can adequately separate them (Smith and Wilson 1996). The often recommended Simpson evenness index may still be highly correlated with richness (Barrufol et al. 2013).

Furthermore, we adopted the life-history trait-based grouping of species (Hector et al. 1999, Fornara and Tilman 2008) since life-history traits, as another aspect of biodiversity, are important for understanding DPR and species coexistence in forests (Lusk and Smith 1998, Verheyen et al. 2003). Differences in leaf habit and shade tolerance can influence spatial niche occupancy, light use efficiency, and nutrient cycling within a site (Man and Lieffers 1999, Yachi and Loreau 2007, Coomes et al. 2009), thereby playing an important role in shaping DPRs. Therefore, we used characteristics of shade tolerance and leaf habit to classify tree species into 4 groups: intolerant deciduous, intolerant evergreen, tolerant deciduous and tolerant evergreen.

Life-history trait variation was defined as the number of groups in each plot. We graded the shade tolerance ratings for the studied species following Spurr and Barnes (1980).

Stand age (SA, years) for each plot was determined according to last stand-replacing fire date or by coring three dominant/co-dominant trees of each tree species inside or outside the plot at the time of plot establishment. With coring, stand age was the average of ring counts from the tree samples of the species with the oldest age, used as a conservative estimate of stand age (Senici et al. 2010). The soil nutrient regime of each site (NR), as indicator of the nutrient supply
and vegetation-soil relationship, was determined based on several factors including slope, texture, depth of horizons, bedrocks, pH and carbon/nitrogen ratio in humus forms, cation exchange capacity, and total nitrogen content in the rooting layer (Courtin et al. 1988). As such, soil nutrient regime was classified as very poor, poor, medium, and rich (1 – 4) in conformity with Saskatchewan’s ecosystem classification guidelines (Beckingham et al. 1996).

We used DBH variation among individual trees within each plot as surrogate of tree size inequity (Potvin and Dutilleul 2009), because the overall DBH variation can be considered as a measure to represent the degree of the realized niche differentiation via positive plant interactions (Yachi and Loreau 2007, Chu et al. 2009). DBH variation was calculated as the coefficient of variation (CV) (Brassard et al. 2008), which is the ratio of the standard deviation of all DBH measurements to the mean DBH within each plot (Supplementary Table S2-1).

Data analysis

To aid in construction of structural equation models (SEMs) and interpretation of results (Grace et al. 2012b), we first examined the bivariate relationships between each hypothesized causal paths according to our hypotheses. We fit each pair of variables using simple linear regression and multiple linear regressions by adding quadratic and cubic polynomial terms. We reported the significant relationships as linear or polynomial (if quadratic term and/or cubic term were significant). Normality was tested for all variables based on a Shapiro-Wilk goodness-of-fit test. As recommended (Grace et al. 2010) and common practices in SEMs (e.g., Oberle et al. 2009, Spasojevic et al. 2014), non-normal continuous variables including Shannon’s index, DBH variation, and stand age were natural-logarithm transformed to mitigate departure from normality and linearity.
As recommended (Grace et al. 2012b), we first specified a metamodel based on the known theoretical construct including the hypothesized multiple paths predicted by the multivariate productivity-diversity hypothesis (Cardinale et al. 2009) with the addition of stand age in the SEM. Then, we fit more complex models including DBH variation as the link between productivity and diversity. An alternative model with opposite direction of the direct path between diversity and DBH variation was also fit. Furthermore, we added the direct causal path from diversity to productivity in all three SEM models above to test whether reciprocal direct causal effects exist between diversity and productivity (Grace et al. 2007).

We used latent variable by incorporating two observable variables, Shannon’s index and life history diversity, to represent species diversity (Grace et al. 2010). Similarly, polynomial terms were incorporated to account for nonlinear effects of predictors on responses by using a zero-error composite variable (Grace and Bollen 2008) for the response of aboveground biomass to stand age. Nutrient regime as an ordinal categorical variable was coded as 1, 2, 3, and 4, being treated as a regular numeric covariate as recommended, provided that NR was strictly endogenous (dependent) variable in our SEMs (Rosseel 2012).

No excessive multivariate skewness and kurtosis were found in our data using Mardia’s multivariate tests ($P = 0.85$ and 1.0, respectively), indicating that the maximum likelihood estimation for SEM was valid. To address the potential issues from nonlinear and remaining univariate non-normality after transformations, we used the nonparametric Bollen-Stine bootstrapping estimations for improved robustness of our SEMs. We chose recommended chi-square tests, root mean square error of approximation (RMSEA), and goodness-of-fit index (GFI) to evaluate the model fit of all SEMs (Kline 2010). A Chi-square with a $P$ value > 0.05 indicates that the observed and expected covariance matrices are not statistically different; RMSEA and
GFI values ranging < 0.05 and > 0.95, respectively, suggest a good model fit (Rosseel 2012). The significant path coefficient for directional paths (single-headed arrows) indicates that the represented causal relationship is statistically significant. Furthermore, the path coefficient, standardized for comparison between pathways, can be a measure for the sensitivity of dependent variable to the predictor (Grace and Bollen 2005). To facilitate the interpretation of our SEM results, the total effects of a given exogenous variable on aboveground biomass was estimated by adding the direct standardized effect and the indirect standardized effect (Grace and Bollen 2005). The SEM was implemented using the lavaan package (Rosseel 2012) in R 3.0.2 (R Development Core Team 2013).

Assessing possible methodological problems

Heterogeneity in sampling plot sizes can influence species diversity estimates as the number of species increases with plot size (Rosenzweig 1995). Similarly, plot size may affect aboveground biomass estimates (Lewis et al. 2009). We tested the associations between Shannon’s index, richness and plot size, and between AGB and plot size by Spearman’s rho correlation. The Spearman’s rho showed no correlation between Shannon’s index and richness to plot size ($P = 0.54$ and 0.83, respectively). However, AGB decreased with plot size ($P < 0.001$). To examine the magnitude of plot size effect on AGB, we used a boosted regression trees (BRT) model to quantify the relative influence of plot size, compared with those from DBH variation, stand age, nutrient regime, and species diversity, on AGB (De'ath 2007, Elith et al. 2008). Plot size accounted for <2% of the relative influence on AGB (Supplementary Fig. S2-2). These analyses show that plot size variation from 0.06 to 0.08 ha had negligible effects on species diversity and AGB estimates, indicating that plot size heterogeneity has minimum effects on results of our SEMs. Furthermore, spatial autocorrelation in the residual of aboveground biomass obtained in
the BRT was checked using Moran’s $I$ global test to identify any potential spatial structure. Moran’s $I$ test showed no significant spatial structure in the BRT residuals (Moran’s $I = 0.07, P = 0.27$).

**Table 2-1** Direct, indirect, and total standardized effects on aboveground biomass of canopy trees based on structural equation models (SEMs). Significant effects are at $P < 0.05$ (*), $< 0.01$ (**), and $< 0.001$ (**).

<table>
<thead>
<tr>
<th>SEM model</th>
<th>Predictor</th>
<th>Pathway to aboveground biomass</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong>, model in Fig. 2A</td>
<td>Diversity</td>
<td>Direct</td>
<td>0.14**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total effect</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Stand age</td>
<td>Direct</td>
<td>0.49***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect through diversity</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total effect</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Nutrient regime</td>
<td>Direct</td>
<td>0.23***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect through diversity</td>
<td>0.02*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total effect</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>B</strong>, model in Fig. 2B</td>
<td>DBH variation</td>
<td>Direct</td>
<td>0.15*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total effect</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Diversity</td>
<td>Direct</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect through DBH variation</td>
<td>0.07***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total effect</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Stand age</td>
<td>Direct</td>
<td>0.45***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect through DBH variation</td>
<td>0.04***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect through diversity</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total effect</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Nutrient regime</td>
<td>Direct</td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect through DBH variation</td>
<td>0.03***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indirect through diversity</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total effect</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Fig. 2-1 Univariate relationships between endogenous (dependent) and exogenous (independent) variables (n = 448). All variables were natural log-transformed except aboveground biomass. Significant regression lines were plotted using linear regression. Parametric assumptions were checked. (a) $y = 101.2 + 153.1x - 150.2x^2$, $R^2 = 0.07$, (b) $y = 199.8 + 59.2x$, $R^2 = 0.13$, (c) $y = -1.42 + 0.612x$, $R^2 = 0.24$, (d) $y = -1343.7 + 608.4x -61.8x^2$, $R^2 = 0.24$; (e) $y = -2.4 + 0.281x$, $R^2=0.1$; and (f) not significant. All fitted regressions are significant at $P < 0.001$. The assumptions of normality and homogeneous variance were validated for all fitted regressions ($P > 0.05$) with an exception of marginal violation of normality of the fitted regression in c ($P = 0.04$).

Results

The overall relationship between canopy tree Shannon’s index and aboveground biomass was quadratic; biomass increased with diversity, then decreased afterward (Fig. 2-1A). The aboveground biomass increased with DBH variation (Figs. 2-1B). Aboveground biomass increased with stand age, and then decreased (Fig. 2-1C). DBH variations increased with both Shannon’s index and stand age (Figs. 2-1D, E), but there was no clear relationship between Shannon’s index and stand age (Fig. 2-1F).
Fig. 2-2 Structural equation models linking aboveground biomass (AGB) and species diversity. (A) Effects of species diversity, soil nutrient regime, and stand age on AGB. (B) The model with tree size inequality, represented by DBH variation, as the linking mechanism. The coefficients are standardized prediction coefficients for each causal path. Solid lines represent significant paths ($P \leq 0.05$) and dash lines for non-significant paths ($P > 0.05$). The path coefficient marked with ‘±’ indicates a nonlinear (quadratic) relationship.

The model without DBH variation as a predictor had a good fit to the data ($\chi^2 = 7.52$, d.f. $= 5$, $P = 0.18$; RMSEA = 0.03; GFI = 0.994) (Fig. 2-2A). While AGB increased with stand age and local nutrient availability, diversity had a positive direct effect on AGB (Table 2-1, Fig. 2-
Similar to the univariate relationship, diversity was not significantly related to stand age (Fig. 2-2A).

An alternative model for the effects of tree size inequality represented by DBH variation on both aboveground biomass and species diversity. The full model including DBH variation as a predictor had a good fit to the data ($\chi^2 = 11.41$, d.f. = 7, $P = 0.13$; RMSEA = 0.04; GFI = 0.99, Fig. 2B), similar to the above simpler model, but yielded additional information with slightly reduced AIC by 0.6 %, suggesting a better model based on parsimony principle. DBH variation had a positive direct effect on AGB. The direct path between diversity and AGB became insignificant, but instead, positive diversity effects on AGB were indirect through increasing DBH variation (Table 2-1, Fig. 2-2B). While the positive direct effects of stand age and nutrient regime on AGB remained, some of the direct effects found in the initial model (Fig. 2-2A) were realized indirectly via increasing DBH variation (Table 2-1, Fig. 2-2B). In both models (Figs. 2-2A and 2-2B), the covariance between life-history trait groups and Shannon’s index were not significant, indicating that they are independent aspects of species diversity.
An alternative model with altered direction for the path between diversity and DBH variation had a slightly better fit to the data ($\chi^2 = 7.51$, d.f. = 7, $P = 0.4$; RMSEA = 0.01; GFI = 1.00) (Fig. 2-3). Similar to the model in Fig. 2-2B, nutrient regime and stand age had strong positive effects on DBH variation. DBH variation had positive effects on not only aboveground but also diversity (Fig. 2-3). The models assuming reciprocal direct causal effects between diversity and productivity showed that the two-way direct causal paths insignificant (see Fig. S2-3).

**Discussion**

Consistent with the pattern reported for species diversity of herbaceous plants along a gradient of resource supply (Grime 1973), we found a significant overall humped-back relationship between Shannon’s index and aboveground biomass in natural forests with a wide range of variation in local soil resource availability and stand age. This finding contrasts with weak and variable relationships found for natural grasslands where the variability of local soil resources was limited for each relationship (Adler et al. 2011). These results indicate that the DPRs are strongly dependent on how environmental variations are controlled in sampling natural systems (Fridley et al. 2012) and whether variations in climate and soils are properly accounted for when net diversity effects are examined (Tilman et al. 2012). As previously hypothesized (Cardinale et al. 2009), we show that diversity has a positive effect on aboveground biomass while both diversity and aboveground biomass are positively influenced by resource supply, represented by soil nutrient regime.

We found that positive diversity effects are mediated via increasing tree size inequality. This finding suggests that positive diversity effects result from high vertical occupation of available space within a canopy, which can increase resource acquisition and utilization (Man...
and Lieffers 1999, Chesson 2000, Yachi and Loreau 2007, Coomes et al. 2009, Clark 2010). This result complements the evidence that diversity increases space utilization belowground (Brassard et al. 2013, Mueller et al. 2013). The underlying biological processes appear to be dependent on individual plant responses to resource availability and neighbors (Cahill et al. 2010).

Additionally, we found that soil resource availability and stand age are positively correlated with tree size inequality, consistent with the idea that the complementarity effects increase with local resource availability (Coomes et al. 2009) and time (Reich et al. 2012, Zhang et al. 2012).

It is valuable, and often necessary to evaluate alternative SEMs based on observational data (Grace et al. 2012b). Our alternative SEM model (Fig. 2-3) shows significant positive feedback of tree size inequality on diversity. Because positive interactions among individual plants can increase size inequality (Chu et al. 2009), we speculate that tree size inequality determines both species diversity and stand productivity. This provides a potential solution to resolve the inconsistent patterns and directions in previous findings of DPRs by linking studies of DPRs and species coexistence influenced by productivity that have been studies in two separate lines (Grime 1973, Loreau et al. 2001, Adler et al. 2011).

The insignificant two-way paths assuming reciprocal direct causal effects between diversity and productivity are consistent with previous findings that diversity effects become weak or insignificant when direct effects go both ways between diversity and productivity (Grace et al. 2007). However, instead of concluding that diversity effects are weak in natural ecosystems (Grace et al. 2007), our results suggest that the reciprocal effects between diversity and productivity are likely indirect in complex ecosystems.
In summary, we show a positive DPR in natural forests of varying stand ages and local nutrient availability. Positive diversity effects on aboveground biomass are mediated by increasing tree size inequality. While both soil nutrient availability and stand age have strong direct effects on aboveground biomass, some positive effects of soil nutrient availability and stand age on aboveground biomass are also indirectly achieved via increasing tree size inequality. Tree size inequality appears to be the central regulating mechanism the positive diversity effects on productivity, and potentially as the shared mechanism that simultaneously regulates productivity and diversity likely via increased resource acquisition and utilization as well as facilitation among individuals.
CHAPTER 3: DOES DIVERSITY OF CANOPY TREES MATTER THE MOST IN FOREST ECOSYSTEM?

Running title: Effects of canopy tree diversity on biomass across forest strata and soil carbon storage

Introduction

Concerns over the continuing loss of species and consequent deterioration of ecosystem functioning have become central foci in ecology over the last three decades. Many studies across taxa and habitats, mostly in grasslands, have observed a positive relationship between biodiversity and ecosystem functioning, i.e., the biodiversity-ecosystem function relationship (BEF) (e.g., Tilman et al. 2001, Isbell et al. 2009). This positive association is often attributed to complementarity effects from interspecific facilitation and/or niche complementarity (Hooper et al. 2005, Cardinale et al. 2007, Fargione et al. 2007). However, more empirical studies are needed to strengthen our understanding of BEF in natural ecosystems due to contradicting observed BEF patterns and lack of the consensus between results from natural ecosystems and controlled experiments (Grace et al. 2007, Adler et al. 2011). This is especially true in natural forest ecosystems due to their complex temporal and spatial dynamics (Chen and Popadiouk 2002, Wardle et al. 2004). Lack of BEF knowledge in forest systems has potentially resulted in underestimation of the importance of diversity on forest functioning and services. For example, global estimates of forest carbon sequestration (Pan et al. 2011) are generally attributed to environmental factors such as temperature and precipitation and stand development such as stand age (Liu et al. 2014), but the role of biodiversity is often overlooked. Thus, the role of canopy tree diversity on ecosystem functions other than canopy tree productivity is not clear, and often contradicting (e.g., Cavard et al. 2011, Gamfeldt et al. 2013).
Whether or not species richness is a reliable predictor of ecosystem functioning is still debated, especially in natural environments (Hillebrand & Matthiessen 2009; Reiss et al. 2009). Largely because the impact of many aspects of biodiversity, as a complex concept ranging from genetics to ecosystem levels, is largely unknown (Purvis & Hector 2000). The identity or specific traits of concurrent species may be the real drivers for productivity (Leps 2004). Therefore, species richness may not be the consistent predictor to estimate the consequences of species loss on productivity in natural communities. Rather, productivity may be more closely linked to species identity, trait dissimilarity between coexisting species (Hillebrand & Matthiessen 2009), instead of number of concurrent species (Nadrowski et al. 2010).

Although plant functional diversity, a measure of divergence on core traits affecting plant performances, has been considered the key to explaining the so-called complementarity effects (Hillebrand & Matthiessen 2009), it remains unclear whether the trait divergence among all constituent species or species richness contributes more to productivity, hence the debate between singular hypothesis and functional redundancy hypothesis (Loreau 2004). For example, the singular hypothesis (Naeem et al. 2002) suggests that each plant species contributes to ecosystem functioning uniquely (e.g., Meinen et al. 2009; Eisenhauer et al. 2010). In contrast, the functional redundancy hypothesis proposes that the high degree of redundancy in plant functional traits enable functions of one species replaced by similar but not identical species so that the functions of ecosystem can be maintained with limit species richness (Thibault et al. 2010). Therefore, it is beneficial to know whether one aspect of species diversity is significant for observed aboveground biomass and soil carbon storage, or all of them are equally important in order to tailor more accurate conservation goals for specific management targets.
Here, we conducted an analysis to examine how canopy tree species diversity affects multiple ecosystem functions including biomass accumulation across canopy strata and total aboveground biomass, and total soil carbon storage. Specifically, we hypothesize: (1) higher canopy tree species diversity, including, richness, evenness, and life-history trait diversity index, will have positive effects on multiple ecosystem functions in forest ecosystems (Gamfeldt et al. 2013); (2) tree species richness is the most important aspect of tree species diversity to influence ecosystem functions; (3) ecosystem functions will be enhanced with stand age because the complementary effects of tree species diversity may increase with time (Cardinale et al. 2007, Weis et al. 2007) despite of the possible decline in forest productivity with aging (Wardle et al. 2004); and (4) climate and local site conditions will affect ecosystem functions through the effects on species diversity (Zhang et al. 2014). We aim to improve understanding of how the climate and soil condition affects diversity effects (of canopy trees) on overall carbon storage (including belowground portion), which is crucial to predict how the boreal carbon stock responds to changes of future environmental and forest managements.

Material and methods

Study area and data

We conducted our analyses using the National Forest Inventory (NFI) database, consisting of 991 permanent sample plots (PSP) systematically allocated across Canada’s forests, covering a wide climatic and geographical gradient, ranging between 53° 25’W and 134° 46’W longitude and 42° 37’N and 68° 14’N latitude. The NFI dataset currently consist of only a single measurement of all plots, conducted during the period of 2000-2006. Measurements were carried-out primarily by provincial agencies following the same ground sampling guidelines (Canadian Forest Inventory Committee, 2004) to ensure precision and accuracy of the data.
Details about sampling and compilation procedures of the NFI dataset can be found in (Zhang et al. 2014).

Response variables: biomass and carbon storage

We include biomass for tree, shrub and herb, and bryophyte layers to represent the overall productivity along the vertical strata of Canadian forests. Furthermore, total aboveground biomass, total aboveground live and dead biomass, and total soil carbon storage were calculated. For each tree (taller than 1.3 m) within plots, the stem wood, stem bark, and branches and foliage biomass were estimated using individual tree biomass equations, and summed to total tree biomass at the plot level. All included allometric equations for individual tree biomass in biomass compilation were carefully checked to ensure accuracy. The biomass equations at national scale in Canada (Lambert et al. 2005) was preferred, and provincial equations where the national equation was absent was used for a given species. Total aboveground biomass of shrubs and herbs, bryophytes, and woody debris at the plot level were estimated based on the weighted oven-dried samples collected within the plots. Aboveground biomass of different layers was summed and scaled-up to total biomass per unit area (Mg/ha) for tree, shrub and herb, and bryophyte layers, respectively. Total aboveground live biomass was the summed biomass of all three live vegetation layers (tree, shrub and herb, and bryophyte); and the aboveground dead biomass was the biomass of dead trees, stumps, and downed woody debris. Total aboveground biomass was the sum of all live and dead biomass.

Counting the complete carbon stock of forest ecosystems, the organic carbon content of forest floor (≤ 8 mm fraction of the forest floor), between 0 to 15cm from mineral soil surface, between 15 to 35cm below surface, and between 35 to 55cm below surface were measured and
summed to the plot level. Then, the carbon content was scaled up to total soil carbon content per unit area (Mg C ha\(^{-1}\)).

**Explanatory variables**

We used species richness, evenness, and a life-history trait index of canopy tree species as measures of diversity. Canopy trees were defined as all stems in each plot with diameter at breast height ≥ 9.0 cm. Species richness (S) was calculated as the count of all live canopy tree species in each plot. Tree species evenness was calculated using Pielou's evenness index (\(J'\)) weighted by the basal area of constituent tree species in each plot (Pielou 1969).

We calculated the life-history traits index as a continuous numeric index (FDis, functional dispersion). We extracted data of 32 selected life-history traits or characteristics (See appendix Table S3-1) from the USDA PLANTS database (USDA and NRCS 2013) for all canopy tree species in our plots, and quantified plant trait divergence between tree species with the R package FD, based on a distance-based framework (Laliberte and Legendre 2010). Using trait-based approaches, functional trait dissimilarity between species is considered the underlying cause of the observed complementarity effects in local competitive communities. However, uncertainty remains on how to choose the best measure of functional diversity defined by traits of individual species affecting performance (Hillebrand & Matthiessen 2009), e.g., arbitrary grouping (number of non-objectively classified functional groups) (e.g., Fornara & Tilman 2008), presence/absence of selected traits (e.g., Cadotte et al. 2009), and a distance-based measure of functional diversity (e.g., Paquette & Messier 2011). We preferred the FDis index based on the well-documented life-history traits of tree species for the following reasons: 1) it accommodates trait types as quantitative, semi-quantitative, and qualitative in a multidimensional trait space; 2) missing values for any given trait is tolerated; and 3) trait divergence between species is...
weighted by relative abundance (basal area) of species within a plot to enable FDis independent of species richness, and to avoid overestimating trait values of rare species. We chose life-history traits to evaluate functional diversity instead of the commonly used phenotypic traits, such as specific leaf area (e.g., Schöb et al. 2013), because life-history traits, as outcomes of long-term evolutionary history, are more reliable predictors than the latter, which are subject to the phenotypic plasticity and heterogeneity of habitats, especially at the plot scale (Cordlandwehr et al. 2013). In addition, the life-history traits for common tree species can be easily acquired from authoritative databases, where actual monitoring of traits of tree species in the field is neither economical nor possible.

We included exogenous factors including climate, soil drainage class, and stand age as covariates in our analyses. Climate data for each plot were estimated using BioSIM software that generated long-term (1951-2010) scale-free climate data from geographic coordinates (latitude, longitude, and elevation) (Régnière and Saint-Amant 2008). Specifically, we used the growing degree days (GDD) (base temperature 5°C) as a measure of the overall available energy for plant growth, and climate moisture index (CMI = mean annual precipitation-annual potential evapotranspiration) to represent drought severity, in which a higher value of CMI means higher water availability for plants (Hogg 1997). We used soil drainage class (SDC) as an integrated measure of the overall site quality, taking into consideration topography, soil texture, and soil nutrient status (Taylor et al. 2000). Soil drainage class was classified into 7 classes, representing very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. We considered the last recorded time since stand replacing disturbance as stand age or estimated stand age as the mean age of the dominant and co-dominant trees of the major species within the stand if time since disturbance was unknown.
All numerical variables were tested for normality using the Shapiro-Wilk test, and were found to be significantly non-normal. Therefore, all numerical variables were natural-log transformed to address departure from normality and potential non-linearity. Specifically, numerical values of 21, 1, and 1 were added to CMI, FDis, and $J'$, respectively, before the natural-log transformation to offset zero and negative values.

We chose boosted regression tree analysis (BRT) for our analyses to accommodate violations of the assumptions of conventional statistics, which are common for censored observational data from natural ecosystems. Specifically, the problems of missing data and partial incompatibility of definitions and classifications between provincial agencies, e.g., heterogeneity introduced by variability in sampling plot sizes or schemes between provincial agencies (See Zhang et al. 2014), prevent conventional parametric statistics from being effective. Boosted regression tree analysis, is an increasingly recognized statistical method that combines the advantages of regression trees through recursive binary splits and adaptive model averaging. It is particularly suitable for analyzing ecological data from natural ecosystems (Elith et al. 2008).

For example, BRT has many desirable traits, including high predictive strength, tolerance of missing values in predictors, invariant to monotonic transformations, and no requirement for those prior assumptions required in conventional parametric statistics (De'ath 2007, Elith et al. 2008).

There are four main parameters required before fitting a BRT model: tree complexity (TC), learning rate (LR), bag fraction (BG), and folds of cross-validation (CV). Tree complexity is the restrain on the complexity of the individual trees in BRT model fitting, i.e., two terminal nodes with a single binary split by the predictor when tc = 1; increasing the numbers of splits
within each the individual when trees tc ≥ 2 (De'ath 2007, Elith et al. 2008). The degree of one predictor influencing other predictors in determination of the response is automatically accounted for via the hierarchical structure of each tree model. The learning rate (LR) determines the contribution of each consequent tree in a shrunk rate. The bag fraction (BG) and cross-validation (CV) introduce stochasticity and the division of data for model training and validation respectively, to account for uncertainties in observational data from natural ecosystems.

We fitted the same set of explanatory variables to each response variable using the recommended values for BRT parameters: LR (0.005), BG (0.6), and CV (10) (Elith et al. 2008). Learning rate was reduced to 0.001 when the minimum number of trees (1000 trees) was not reached to ensure a reliable model-fit (Elith et al. 2008). Each BRT was fitted with TC from 1 to 4 with the above parameter values. We opted to report the simpler BRT model, i.e., the model with a smaller TC, when prediction error (PE) is similar for an alternative, more complex model, to avoid over-fitting. All analyses were performed with the gbm R package (Ridgeway 2010). Because biomass and soil carbon storage are numerical, continuous variables, a Gaussian distribution of errors was used for all BRT fittings. Further, because of stochasticity in our observational data, especially when compiled from multiple sources, we were cautious to avoid over-interpretation of the BRT results. As such, only clear trends and relatively influential predictors were emphasized in the result section.

Results

The set of predictors: GDD, CMI, SDC, SA, S, $J'$, and FDis accounted for 76%, 71%, and 65% of variation in tree layer biomass, total aboveground live biomass, and total aboveground biomass (Table 3-1). By contrast, the same set of predictors accounted for substantially less
variations: 32%, 53%, 42%, and 33% in biomass of understory vegetation (shrub, herb, and bryophyte layers), aboveground dead biomass, and soil carbon storage, respectively (Table 3-1).
Table 3-1 Results from boosted regression tree analysis (BRT) to test the effects of canopy tree species diversity on aboveground biomass and soil carbon storage. Predictors’ relative influences represent the percentage contributions of predictors in the accounted variation of each BRT model. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, SA – stand age (yrs), S – canopy tree richness, $J'$ – canopy tree evenness index, FDis – canopy tree life-history trait index, TC – tree complexity, n – number of sample plots, and PE – model prediction error. All numerical variables were natural log transformed except for SDC.

<table>
<thead>
<tr>
<th>Function</th>
<th>Predictors’ relative influences (%)</th>
<th>TC</th>
<th>n</th>
<th>PE</th>
<th>Trees</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GDD</td>
<td>CMI</td>
<td>SDC</td>
<td>SA</td>
<td>S</td>
<td>$J'$</td>
</tr>
<tr>
<td>Tree aboveground biomass</td>
<td>8.91</td>
<td>11.22</td>
<td>14.76</td>
<td>17.81</td>
<td>33.25</td>
<td>3.28</td>
</tr>
<tr>
<td>Shrub and herb biomass</td>
<td>37.87</td>
<td>10.69</td>
<td>16.84</td>
<td>3.29</td>
<td>5.32</td>
<td>7.31</td>
</tr>
<tr>
<td>Bryophyte biomass</td>
<td>34.71</td>
<td>10.62</td>
<td>16.6</td>
<td>9.12</td>
<td>9.39</td>
<td>3.44</td>
</tr>
<tr>
<td>Total aboveground biomass</td>
<td>7.22</td>
<td>10.71</td>
<td>16.41</td>
<td>24.14</td>
<td>28.14</td>
<td>3.28</td>
</tr>
<tr>
<td>Aboveground dead mass</td>
<td>6.28</td>
<td>21.57</td>
<td>35.27</td>
<td>16.79</td>
<td>9.44</td>
<td>4.07</td>
</tr>
<tr>
<td>Aboveground live mass</td>
<td>6.17</td>
<td>17.88</td>
<td>19.98</td>
<td>24.54</td>
<td>21.49</td>
<td>2.72</td>
</tr>
<tr>
<td>Soil carbon storage</td>
<td>9.76</td>
<td>27.83</td>
<td>34.19</td>
<td>9.31</td>
<td>5.93</td>
<td>6.61</td>
</tr>
</tbody>
</table>
a)  
Tree aboveground biomass (Mg ha⁻¹)  

b)  
Shrub and herb biomass (Mg ha⁻¹)  

c)  
Bryophyte biomass (Mg ha⁻¹)  

d)  
Total aboveground biomass (Mg ha⁻¹)  

e)  
Relative influences (%)  

$R^2 = 0.76$  

$R^2 = 0.32$  

$R^2 = 0.53$  

$R^2 = 0.65$  

$R^2 = 0.33$  

Sal carbon storage (Mg ha⁻¹)  

Canopy tree richness  
Canopy tree evenness  
Canopy tree functional diversity
Figure 3-1 Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to canopy tree diversity indices: canopy tree richness, canopy tree evenness index, and canopy tree life-history trait index. (a) Tree aboveground biomass. (b) Shrub and herb biomass. (c) Bryophyte biomass. (d) Total aboveground biomass. (e) Soil carbon storage. Scatter points are observed values plotted by the respective response and explanatory variables. Abbreviations are GDD − growing degree days (number of days above 5 °C), CMI − climate moisture index (cm), SDC − soil drainage class, SA – stand age (yrs), S – canopy tree richness, $J'$ – canopy tree evenness index, and FDis – canopy tree life-history trait index. All numerical variables were natural log transformed except for SDC.

Tree layer biomass and total aboveground biomass

Diversity of the tree layers ($S, J'$, and FDis) accounted for 47.3% (33.3%, 3.28%, and 10.76%) of the relative influence on tree layer biomass, while the environmental predictors (SDC, CMI, and GDD) accounted for 34.9%, and stand age, 17.8% (Table 3-1). Tree layer biomass increased with canopy tree species richness monotonically, while trends in response to $J'$ and FDis were not as clear, consistent with the relative influences of those predictors (Fig. 3-1a). Similarly, for total aboveground biomass, SA and S were the strongest predictors, followed by SDC, CMI, FDis, GDD, and $J'$ with 24.5%, 21.5%, 20.0%, 17.9%, 7.2%, 6.2%, and 2.7% relative influences, (Table 3-1). Total aboveground biomass increased with canopy tree species richness monotonically, while the trends in response to $J'$ and FDis were not as clear, consistent with their relative influences of those predictors (Fig. 3-1d). Tree layer biomass increased with CMI monotonically, and increased with GDD similarly, but to a lesser extent (Fig. 3-2a). The biomass increased and then decreased with reduced soil drainage, peaked at an intermediate SDC (Fig. 3-2a). As expected, tree layer biomass increased with SA monotonically (Fig. 3-2a). Reponses of total aboveground biomass (Fig. 3-2d) were similar with those of tree layer biomass.
Figure 3-2 Results of BRT models showing the observed and predicted responses to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), and stand age (SA). (a) Tree aboveground biomass. (b) Shrub and herb biomass. (c) Bryophyte biomass. (d) Total aboveground biomass. (e) Soil carbon storage. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical variables were natural log transformed except for SDC.

Figure 3-3 Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to canopy tree diversity indices: canopy tree richness, canopy tree evenness index, and canopy tree life-history trait index. (a) Aboveground live biomass. (b) Aboveground dead biomass. Scatter points are observed values plotted by the respective response and explanatory variables. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, SA – stand age (yrs), S – canopy tree richness, $J'$ – canopy tree evenness index, and FDis – canopy tree life-history trait index. All numerical variables were natural log transformed except for SDC.
Figure 3-4 Results of BRT models showing the observed and predicted responses to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), and stand age (SA). (a) Aboveground live biomass. (b) Aboveground dead biomass. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical variables were natural log transformed except for SDC.

The relative influences of predictors and responses of aboveground live biomass were identical to those of total aboveground biomass (Fig. 3-3a & 3-4a). In contrast, the influence of canopy tree diversity was less important on aboveground dead biomass (Fig. 3-3b).

Aboveground dead biomass increased with CMI, but decreased with GDD (Fig. 3-4b) and reduced soil drainage (Fig. 3-4b). Aboveground dead biomass decreased with SA, then increased, following a u-shape pattern (Fig. 3-4b).
**Biomass of shrub, herb, and bryophyte layers**

Understory, stratum-specific, biomass responded to canopy tree species richness, GDD, and SA differently compared to tree layer biomass. The relative influence of GDD was 37.9% and 34.7% for the shrub and herb layer (combined) and the bryophyte layer, respectively; whereas the relative influence of SA was 3.3% and 9.1%, respectively. The biomass of the shrub and herb, and bryophyte layers showed distinct patterns compared to tree biomass (Fig. 3-1b-c).

Shrub and herb biomass and bryophyte biomass correlated negatively with canopy tree species richness and life-history trait index (Fig. 3-1b-c). Understory biomass increased with CMI and SDC, but decreased with GDD (Fig. 3-2b-c). The bryophyte biomass increased with SA (Fig. 3-2c), but no clear trend in shrub and herb biomass (Fig. 3-2b).

**Soil carbon storage**

Soil drainage class and CMI were the strongest predictors of soil C storage, followed by GDD, SA, J’, FDis, and S with 34.2%, 27.8%, 9.8%, 9.3%, 6.6%, 6.4%, and 5.9% relative influences, respectively (Table 3-1). Soil carbon storage increased with canopy tree species richness but to a much lesser extent comparing with the trends found for tree biomass and aboveground biomass, while trends in response to J’ and FDis were not as clear (Fig. 3-1e). Soil carbon storage increased with CMI and SDC (Fig. 3-2e). The soil carbon storage increased and then decreased with reduced soil drainage, peaked at an imperfectly drainage (Fig. 3-2e). However, soil carbon storage appeared no correlation with SA (Fig. 3-2e).

**Discussion**

To our best knowledge, this is a first study showing that canopy tree diversity was not the main factor in productivity of understorey and soil carbon storage. Contrary to the previous findings that canopy tree diversity has strong positive effects on ecosystem functions and services of other
strata and growth form (Gamfeldt et al. 2013), we showed that understorey biomass was negatively influenced by canopy tree richness and life-history trait index. Although canopy tree species diversity, especially richness, showed strong positive diversity effects on tree layer biomass, consistent with a global meta-analysis in forests (Zhang et al. 2012), diversity effects of canopy tree can be negative, rather than consistent positive diversity effects. For example, high tree species diversity leads to high biomass of tree layers, reducing resources available for understorey, consequently lower understorey vegetation abundance and species diversity as predicted by resource availability hypothesis (Bartels and Chen 2010). Therefore, the assumptions that high value in one aspect of biodiversity within given organisms in a community leads to high ecosystem functions and services (Zavaleta et al. 2010, Isbell et al. 2011) may be misleading.

We found that species richness were more importance predictor on forest functions and services than life-history trait index and evenness of canopy trees. This is consistent with the prediction of the singular hypothesis (Naeem et al. 2002) suggests that each plant species contribute to ecosystem functioning uniquely (e.g., Meinen et al. 2009; Eisenhauer et al. 2010) in contrast to the prediction of functional redundancy hypothesis that high degree of redundancy in plant functions enable important functions of one species replaced by similar but not identical species (Thibault et al. 2010). The lack of clear patterns and the small sizes of the relative importance of canopy tree evenness on those ecosystem functions is not consistent with previous generalization in forests that evenness may be one of the most important aspect of species diversity in shaping the BEFs (Zhang et al. 2012). This suggests that importance of evenness (Kirwan et al. 2007) may be overestimated so that the identity of the dominant plant species is the key predictor for community productivity (Mulder et al. 2004). However, the alternative
explanation may be that the lack of effective diversity index to separate the effects of evenness and richness from each other. Similarly, the life-history trait index showed positive but non-monotonic effects on tree layer biomass and considerable negative impact on bryophyte biomass, indicating that the effects of trait divergence may be overestimated. However, our results cannot rule out that the value of the trait index depends highly on the traits being selected, often arbitrarily by the researchers. In addition, the continuous trait diversity index (FDIs) had integrated the aspects of richness, evenness, and dispersion of selected traits (Spasojevic et al. 2014), may not be independent from richness (Farwig et al. 2013).

The effects of GDD, CMI, and SDC varied, but the influences were substantial among forest functions and services. The important role of climate and local site conditions are consistent with previous findings that moisture and temperate related climatic factors and local soil conditions affect aboveground tree biomass but the effects vary across biomes (Lehmann et al. 2014). For example, GDD positively promoted tree layer biomass but negatively impacted the understorey aboveground biomass. We speculate that the differences in responses of different growth forms and strata can be indirect results through the influences of environmental factors on species diversity of each growth form and stratum directly (Zhang et al. 2014), in addition to the direct effects of environments on forest functions and services (Oberle et al. 2009). The understorey biomass and soil carbon storage favoured soils with poor drainage while the tree layer biomass peaked at sites with the intermediate drainage. We speculate that this may be results of the divergence in life-history traits between tree species and understorey species, for example, shade tolerance, draught tolerance, etc.

The stand age had considerably amount of influences on forest functions and services across growth forms and forest strata. This is consistent with previous knowledge that forest
function as a carbon source increases with stand age (Lewis et al. 2009). However, we showed that the aboveground biomass increased with stand age, much apparently than that of the soil carbon storage. This discrepancy between the age effects on aboveground and belowground component of carbon and biomass is not consistent with previous findings that the soil and total ecosystem carbon increases with the time since fire (Wardle et al. 2003, Wardle et al. 2012).

In this study, we show that canopy tree diversity, especially the richness, are important for tree aboveground biomass, despite the large variations in climate and site conditions in natural Canadian forest ecosystems. Further, the negative effects of canopy tree diversity on understorey plant biomass are against the common presumption about the positive effects of the canopy tree species diversity on full-range of forest functions including the productivity of understorey and soil carbon storage (Gamfeldt et al. 2013). This study is one step further from previous study in boreal forest (Cavard et al. 2011) by directly linking canopy tree diversity to soil carbon storage and understorey dynamics rather than the coarse classification of canopy tree vegetation type with much greater extent of the study area, and thus the increased inference space. Further, the total aboveground biomass and tree layer biomass increased monotonically with stand age, and lack of response to age in soil carbon storage, suggesting that the role of carbon sink of forests may improve through time.
CHAPTER 4: IS POSITIVE BIODIVERSITY-ECOSYSTEM FUNCTIONING RELATIONSHIP CONSISTENT ACROSS FOREST STRATA?

Running title: Patterns and drivers of diversity-productivity relationship in natural forests

Introduction

More than three decades of intensive studies have established the important role of biodiversity on maintaining ecosystem productivity and other functioning and services (Reich et al. 2012, Tilman et al. 2012) (Cardinale et al. 2012). Many experimental results, mainly from grasslands, demonstrate that aboveground productivity is positively correlated with increasing species richness (e.g., Isbell et al. 2009). Although less common, studies of biodiversity effects on ecosystem functioning (BEF) in forests have shown similar trends (e.g., Vilà et al. 2003, Vilà et al. 2007, Morin et al. 2011). However, many BEF studies conducted in forest systems have been criticized because they fail to account for many of the complexities of natural forests (Flombaum and Sala 2008, Duffy 2009, Willig 2011). For instance, many previous BEF studies in forests have focused on the effects of canopy tree species diversity (e.g., richness and evenness) on productivity with less consideration of other vegetation life forms. Lack of consideration of all vegetation stratum in complex structured forests may result in misleading conclusions concerning the magnitude and patterns of BEF in forest ecosystems (Fowler et al. 2012).

In this study, we examine BEF across forest vegetation strata, including the tree layer, shrub and herb layer, and bryophyte layer. Further, we compare the relative importance of total plant richness and tree richness on total aboveground biomass and total soil carbon content as important forest ecosystem functions. Specifically, we hypothesize: (1) the magnitude and direction of diversity-productivity relationships (DPRs) differ because DPR may be highly context dependent, attributable to the complexity of natural ecosystems (Willig 2011, Midgley...
2012); (2) diversity of tree species may contribute more to total aboveground biomass and soil 
carbon storage than those from total plant richness, given the dominant role of canopy trees in 
multiple forest ecosystem functions (Gamfeldt et al. 2013); (3) species diversity effects of plants 
on ecosystem functioning may differ between forest strata and growth form because richness of 
different growth form responds to climatic environmental factors differently (Oberle et al. 2009) 
including GDD, CMI, and SDC (Zhang et al. 2014); and (4) stand age may influence the 
accumulative functions such as plant biomass and soil carbon storage positively due to the 
enhanced diversity effects over succession (Reich et al. 2012).

Material and methods

Study area and available data

We used the National Forest Inventory (NFI) data including 987 permanent sampling plots 
(ground plots). The study area covered by those ground plots are between 53° 25'W and 134° 
46'W longitude and 42° 37'N and 68° 14'N latitude. Since the same dataset was used in previous 
chapters (Ch. 1 and 3) to tackle different aspects of the BEF studies, we do not repeat the 
information about this dataset in this chapter. The details about sampling and compilation 
procedures of this NFI dataset can be found in the published chapter 1 of this thesis (Zhang et al. 
2014).

Response variables: biomass and carbon storage

The response variables in this chapter include biomass of tree, shrub and herb, and bryophyte 
layers, and total aboveground biomass include biomass from dead trees, stumps, and woody 
debris. The total carbon storage in forest floor and mineral layers were also measured and 
compiled. For the details about the compilations of biomass and soil carbon storage, refer to the 
chapter 1 and 3.
Explanatory variables

Similarly as previous chapters, we included GDD, CMI, SDC, and SA as covariates in all models to account for the influences of external factors. The main predictors are species richness for each canopy stratum: tree richness for tree layer, combined richness for shrub and herb layer, and bryophyte richness for bryophyte layer including all other non-vascular plants as well. In addition, total plant richness was included as a predictor in a separate model with total plant richness to replace richness of other layers. To calculate tree species richness, we counted the total number of species found in each plot by pooling individual tree data of large tree plot and small tree plot. Similarly, we counted the total number of unique species for each plot by pooling species data for the overall richness of shrub and herb layers. For other details about the compilations of those predictors, refer to the chapter 1 and 3.

Statistical analyses

Similar to Chapter 3, I used BRT to examine the relationship between biomass and various predictors. The rationale of using BRT and details about BRT parameter settings, and fitting procedures can be found in previous chapter 1 and 3.
Table 4-1 Results from boosted regression tree analysis (BRT) of aboveground biomass and soil carbon storage. (a) models with plant richness of each plant layer. (b) models with tree richness as predictor. (c) models with total plant richness as predictor. Predictors’ relative influences represent the percentage contributions of predictors in the accounted variation of each BRT model. Abbreviations are GDD – growing degree days (number of days above 5 °C), CMI – climate moisture index (cm), SDC – soil drainage class, SA – stand age (yrs), S – plant richness, TC – tree complexity, n – number of sample plots, and PE – model prediction error. All numerical variables were natural log transformed except for SDC.

<table>
<thead>
<tr>
<th>Model</th>
<th>Ecosystem functions</th>
<th>Predictors’ relative influences (%)</th>
<th>GDD</th>
<th>CMI</th>
<th>SDC</th>
<th>SA</th>
<th>S</th>
<th>TC</th>
<th>n</th>
<th>PE</th>
<th>Trees</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Tree aboveground biomass</td>
<td>14.41</td>
<td>16.66</td>
<td>14.25</td>
<td>41.15</td>
<td>13.53</td>
<td>3</td>
<td>969</td>
<td>0.2</td>
<td>1700</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shrub and herb biomass</td>
<td>40.37</td>
<td>7.09</td>
<td>18.4</td>
<td>5.51</td>
<td>28.63</td>
<td>2</td>
<td>606</td>
<td>0.21</td>
<td>4350</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bryophyte biomass</td>
<td>32.88</td>
<td>10.9</td>
<td>17.9</td>
<td>12.83</td>
<td>25.5</td>
<td>2</td>
<td>554</td>
<td>0.32</td>
<td>2150</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>Aboveground live biomass</td>
<td>13.69</td>
<td>16.65</td>
<td>13.2</td>
<td>38.22</td>
<td>18.24</td>
<td>4</td>
<td>977</td>
<td>0.14</td>
<td>1750</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total aboveground biomass</td>
<td>8.53</td>
<td>23.39</td>
<td>22.19</td>
<td>29.16</td>
<td>16.74</td>
<td>2</td>
<td>987</td>
<td>0.1</td>
<td>3200</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soil carbon storage</td>
<td>15.34</td>
<td>30.68</td>
<td>31.97</td>
<td>15.75</td>
<td>6.26</td>
<td>3</td>
<td>720</td>
<td>0.09</td>
<td>2450</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>Aboveground live biomass</td>
<td>18.68</td>
<td>18.79</td>
<td>18.9</td>
<td>40.98</td>
<td>2.65</td>
<td>4</td>
<td>977</td>
<td>0.15</td>
<td>1200</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total aboveground biomass</td>
<td>12.3</td>
<td>24.41</td>
<td>27.06</td>
<td>31.48</td>
<td>4.76</td>
<td>4</td>
<td>987</td>
<td>0.1</td>
<td>1600</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soil carbon storage</td>
<td>11.08</td>
<td>30.82</td>
<td>30.19</td>
<td>10.27</td>
<td>17.64</td>
<td>3</td>
<td>720</td>
<td>0.09</td>
<td>1300</td>
<td>0.35</td>
<td></td>
</tr>
</tbody>
</table>
a) Relative influences (%)

Tree richness | Climate moisture index (cm yr⁻¹) | Growing degree days | Soil drainage class | Stand age (yrs)

b) Relative influences (%)

Shrub and herb richness | Climate moisture index (cm yr⁻¹) | Growing degree days | Soil drainage class | Stand age (yrs)

c) Relative influences (%)

Bryophyte richness | Climate moisture index (cm yr⁻¹) | Growing degree days | Soil drainage class | Stand age (yrs)

R² = 0.65
R² = 0.29
R² = 0.62
**Figure 4-1** Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to growing degree days (GDD), climate moisture index (CMI), soil drainage class (SDC), stand age (SA), and species richness of forest strata (S). (a) Tree aboveground biomass. (b) Shrub and herb biomass. (c) Bryophyte biomass. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical variables were natural log transformed except for SDC.

**Results**

Our combined set of explanatory variables, including GDD, CMI, SDC, SA, and richness of trees, shrub and herb, and bryophyte layers, accounted for 65%, 29%, and 62% of variation in the biomass of the pertinent canopy stratum (Table 4-1a). While accounting for variation from covariates GDD, CMI, SDC, our models that included tree layer richness as an explanatory variable accounted for 66%, 60%, and 41% of variation in aboveground live biomass, total aboveground biomass, and soil carbon storage, respectively (Table 4-1). Similarly, models that included total plant richness, instead of tree richness, accounted for 61%, 61%, and 35% of variation in aboveground live biomass, total aboveground biomass, and soil carbon storage, respectively (Table 4-1).

**Biomass of tree layer**

For tree layer biomass, SA was the strongest predictors, followed by CMI, GDD, SDC, and S with 41.2%, 16.7%, 14.4%, 14.3%, and 13.5% relative influences, respectively (Table 4-1). Tree layer biomass increased with tree species richness monotonically (Fig. 4-1a). Tree layer biomass increased with CMI and GDD (Fig. 4-1a). The biomass slightly increased and then decreased with reduced soil drainage, peaked at an intermediate SDC (Fig. 4-1a). As expected, tree layer biomass increased with SA monotonically (Fig. 4-1a).
Biomass in shrub and herb combined layer

The relative influences of predictors (Table 4-1) clearly showed that the understory stratum-specific biomass responded to species richness of each stratum, GDD, SDC and SA differently compared to tree layer. For biomass of shrub and herb layer, GDD and S were the strongest predictors, followed by SDC, CMI, and SA with 40.4%, 28.6%, 18.4%, 7.1%, and 5.5% relative influences, respectively (Table 4-1). The biomass from the shrub and herb combined layer increased with tree species richness but not strictly monotonic (Fig. 4-1b). The biomass decreased with GDD monotonically (Fig. 4-1b). The biomass increased with reduced soil drainage, peaked at an intermediate SDC (Fig. 4-1b). The biomass appeared not to be affected by CMI and SA (Fig. 4-1b).

Biomass in bryophyte layer

For biomass of bryophyte layer, GDD and S were the strongest predictors, followed by SDC, SA, and CMI with 32.9%, 25.5%, 17.9%, 12.8%, and 10.9% relative influences, respectively (Table 4-1). The biomass from the bryophyte layer increased with tree species richness, but decreased with GDD monotonically (Fig. 4-1c). The biomass increased with reduced soil drainage, peaked at the poorly drained site (Fig. 4-1c). The biomass appeared to be positively correlated with CMI and SA (Fig. 4-1c).
Figure 4-2 Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to growing degree days (GDD), climate moisture index (CMI), soil drainage class (SDC), stand age (SA), and total plant species richness (S). (a) Aboveground live biomass. (b) Total aboveground biomass. (c) Soil carbon storage. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical variables were natural log transformed except for SDC.

**Aboveground live biomass and total aboveground biomass**

For the aboveground live biomass, SA and tree richness were the strongest predictors, followed by CMI, GDD, and SDC with 38.2%, 18.2%, 16.7%, 13.7%, and 13.2% relative influences, relatively (Table 4-1). However, species richness became less important when replaced by total plant richness with 2.7% relative influence, while other predictors remained comparable value of the relative influences (Table 1). For the total aboveground biomass, the model yielded similar results (Table 4-1; Fig. 4-2a & 4-2b).

The aboveground live biomass increased with total plant richness, CMI, GDD, and SA, while the biomass increased and then decreased with reduced soil drainage, peaked at an intermediate SDC (Fig. 4-2a). When total plant richness was replaced by tree layer richness, the partial dependence plots showed similar trends and patterns but noted that tree species richness had more prominent positive trends with much higher relative importance for the total aboveground live biomass (Fig. 4-3a). The total aboveground biomass showed similar trends (Fig. 4-2b & 4-3b).
Figure 4-3 Results of BRT models showing the relative influences of all predictors on responses and the observed and predicted responses to growing degree days (GGD), climate moisture index (CMI), soil drainage class (SDC), stand age (SA), and tree species richness (S). (a) Aboveground lives biomass. (b) Total aboveground biomass. (c) Soil carbon storage. Soil drainage classes from A to G represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective response and explanatory variables. All numerical variables were natural log transformed except for SDC.

Soil carbon storage

For soil carbon storage, SDC and CMI were the strongest predictors, followed by SA, GDD, and tree layer richness with 32.0%, 30.7%, 15.8%, 15.3%, and 6.3% relative influences, respectively (Table 4-1). In contrast, the total plant richness had 17.6% relative influences in model that included total plant richness instead of the tree layer richness (Table 4-1). The soil carbon storage increased with total plant richness monotonically (Fig. 4-2c), while the trends in response to tree layer richness were not clear (Fig. 4-3c). Tree layer biomass showed complex response to CMI and GDD with a general positive correlation (Fig. 4-2c & 4-3c). The biomass increased and then decreased with reduced soil drainage, peaked at an extremely poor drainage (Fig. 4-2c & 4-3c). However, soil carbon storage appeared no correlation with SA (Fig. 4-2c & 4-3c).

Discussion

This analysis is, to our best knowledge, the first to reveal the positive diversity effects on forest aboveground biomass and soil carbon storage across forest strata/growth forms. Specifically, we demonstrate the positive richness effects of plant species within each stratum or growth form on plant aboveground biomass and the positive effects of the total plant richness on the total soil carbon storage. To some extent, the positive richness effects are partially consistent with the previous findings that positive DPR is ubiquitous in forest ecosystems (Zhang et al. 2012), when considering within the specific forest strata and/or growth form. The implication of this general positive diversity effects within strata/growth
form is consistent with previous conclusions in BEF studies that more species are required to
maintain multiple ecosystem functions (Zavaleta et al. 2010, Isbell et al. 2011). However, our
findings in chapter 3 are contradicting to the previous assumptions/believes that species
diversity of the dominant species within a community, for example the canopy tree species
(Gamfeldt et al. 2013), have positive effects on multiple ecosystem functions across different
strata and/or growth forms. The limitation of strata or growth on the positive diversity effects
suggests that it is less meaningful or even invalid in some cases to presume the positive
species diversity effects on forest functions out of context, e.g., assuming positive diversity
effects of canopy trees on the understorey plant biomass.

In contrast to previously found positive canopy tree diversity effects on soil carbon
storage in forests (Gamfeldt et al. 2013), we found that total plant richness had stronger
positive influences than those of canopy tree species richness on soil carbon storage.
Conversely, total plant richness has fewer influences on the aboveground biomass than those
of canopy tree species richness. The varied relative influences between total plant richness
and canopy tree species richness on varied ecosystem functions demonstrate that it is
unrealistic to assume positive diversity effects without confining the ecosystem functions and
services to certain contexts. The belief that the increased tree species richness will improve
the full range of the forest services (e.g., Gamfeldt et al. 2013) may be an over-simplification
of the complex natural ecosystem, thus requiring a second thought about the indispensible
role of plant species within other forest strata and growth forms.

In addition to previous findings (Oberle et al. 2009) that magnitude of diversity
effects varies between plant growth forms, we show that direction of the response in
ecosystem functions can be altered by climatic and local soil conditions, for example, the
negative response in biomass of shrub and herb combined, and bryophyte layers to the GDD.
This is consistent with results from pervious chapter 1 that richness of different growth forms
respond to climatic and local environmental factors differently (Zhang et al. 2014). However, we argue that the altered BEFs among strata and growth forms can be results of the altered pertinent species diversity in addition to the previous concluded dominant role of canopy tree species through canopy density in determinations of the varied BEFs among forest strata and growth forms, especially for the understorey (Oberle et al. 2009). For ecosystem functions (biomass) in each forest stratum or growth form, the environments influence the specific ecosystem function indirectly through the altered species diversity (richness) in pertinent layer or growth form.

Interestingly, we found no clear effects of stand age on soil carbon storage, suggesting that belowground carbon sink most likely will not decline through aging. On the other hand, the total aboveground plant biomass increased with stand age, suggesting that the role of carbon sink in forest ecosystems may be enhanced by stand age, rather than decline to become a carbon source, at least before reaching a really old age.

In this paper, the generalization of the NFI data is consistent with our hypotheses. First, we show that the positive diversity effects are consistent for each forest stratum. Second, the strength of positive diversity effects varies across strata. Further, we show strong resource filtering of tree layer over understorey. Particularly, we highlight that a specific target of forest management can be improved by a specific aspect of plant diversity in forests, rather than looking for an universal diversity index for improving/maintaining the full range of forest services. The underlying mechanisms for those observed patterns could be one of the many potential ecological and biological processes and their interactions, for example, interspecific competition, facilitation, and Janzen-Connell effects (Janzen 1970, Connell 1971, Dohn et al. 2013). However, BRT based on the censored observational data from natural forests may not be the best approach for determination of the mechanisms. The establishment of the causal relationships between the specific aspect of the plant species diversity and any
identified desirable ecosystem services requires purposely-designed experiment with the
adequate controls.
GENERAL CONCLUSION

This study provides a guideline to facilitate further investigations in the patterns of species coexistence and BEF and the underlying mechanisms, namely the complex interactive web intertwined with endogenous biotic factors and the exogenous environmental factors in boreal forest ecosystems. Further, this study will provide new knowledge to prioritize the questions being asked, methods being used, and the utilization of available experimental data in forest ecosystems for future studies. Practically, our efforts to clarify concepts, causes, and consequences of species loss or gain is essential to accomplish better predictions about community dynamics and ecosystem functions in response to the current threats such as climate and land-use changes and invasive species. The outcomes of this study will benefit forest management practices aimed at multidimensional goals such as higher per-unit-area productivity or conservation of biodiversity.

In summary, this study has deepened our understanding to the key features of the natural ecosystems. I highlight that relationships between biodiversity and multiple ecosystem functions and services are multifaceted, subject to several covariates including climate, local site conditions, time since stand-replacing disturbances, and the reciprocal interactions between diversity and ecosystem functions and services. Several underlying mechanisms could be involved in those ecological processes in a multivariate space. A mechanistic understanding of a general framework on the multifaceted relationships between diversity, productivity, species coexistence, and their interactions with multiple exogenous gradients has not been established conclusively. The common belief is that the adequate substantiations to establish causal relationships may still rely on deliberately designed experiments with satisfactory controls over multiple covariates, mostly environmental factors and biotic factors that are not being studied in current study. Therefore, our findings call for more complementary experiments to identify and verify those potential
mechanisms and evaluate the relative importance of those mechanisms. In future experimental studies, the pertinent questions could be (1) how the current resources management can be improved to meet the projected goals in enhancing ecosystem functions and services, and (2) how the improved diversity in one layer of the hierarchical structure, such as canopy strata, growth forms, and trophic levels, influence the species diversity effects in the target ecosystem services?
Figure S1-1 The distributions of 988 national forest inventory ground plots.
Figure S1-2 Distribution of plot sizes of the national forest inventory ground plots. (a) Canopy tree species richness, Simpson’s dominance index, and total plant species richness. (b) Shrub layer species richness. (c) Herbaceous layer species richness. (d) Bryophyte species richness.
Figure S1-3 Frequency distribution of species richness. a) Canopy tree species richness, b) total plant species richness, and c) the ratio of canopy tree species richness to total plant species richness.
Figure S1-4 Observed and predicted responses of diversity indices to growing degree days (GGD), climate moisture index (CMI), natural logarithm of primary productivity (NPP), soil drainage class (SDC), and natural logarithm of stand age on species richness. (a) Total plant species richness. (b) Shrub layer richness. (c) Herbaceous layer richness. (d) Bryophyte richness. Soil drainage class from A to E represent very rapidly, rapidly, well, moderately well, imperfectly, poorly, and very poorly drained, respectively. Scatter points are observed values plotted by the respective dependent variable and predictor.
Figure S1-5 Predicted response of canopy tree species richness to growing degree days and climate moisture index.
### APPENDIX II. SUPPLEMENTAL INFORMATION FOR CHAPTER TWO

**Table S2-1** Endogenous (dependent) and exogenous (independent) variables used in univariate and structural equation models ($n=448$). Nutrient regime is defined as very poor (1), poor (2), medium (3), and rich (4). Values are mean ± SD and range in brackets for continuous variables or median and range for ordinal variables.

<table>
<thead>
<tr>
<th>Conceptual category</th>
<th>Variable</th>
<th>Transformation</th>
<th>Scale (units)</th>
<th>Mean ± SD (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productivity</td>
<td>Aboveground biomass (AGB)</td>
<td>-</td>
<td>Continuous (Mg ha$^{-1}$)</td>
<td>131.2 ± 46.4 (13.7 – 271.2)</td>
</tr>
<tr>
<td>Diversity</td>
<td>Shannon’s index</td>
<td>Natural log</td>
<td>Continuous</td>
<td>0.42 ± 0.23 (0.00 – 0.95)</td>
</tr>
<tr>
<td></td>
<td>Life-history trait variation (LH)</td>
<td>-</td>
<td>Ordinal</td>
<td>2 (1 – 4)</td>
</tr>
<tr>
<td>Size inequality</td>
<td>DBH variation</td>
<td>Natural log</td>
<td>Continuous</td>
<td>-1.16 ± 0.29 (-1.96 – -0.41)</td>
</tr>
<tr>
<td>Stand age</td>
<td>Stand age</td>
<td>Natural log</td>
<td>Continuous (years)</td>
<td>4.41 ± 0.32 (3.5 – 5.24)</td>
</tr>
<tr>
<td>Resources availability</td>
<td>Soil nutrient regime</td>
<td>-</td>
<td>Ordinal</td>
<td>2 (1 – 4)</td>
</tr>
</tbody>
</table>
Fig. S2-1 Locations of the 448 sample plots from Saskatchewan, Canada.
Fig. S2-2 Relative influence of plot size, compared with other predictors, on aboveground biomass. Predictors included in boosted regression models are DBH variation, stand age, soil nutrient regime, Shannon’s index, and plot size (see Supplementary Table S1 for more explanations).
Fig. S2-3 Structural equation models linking aboveground biomass (AGB) and species diversity. Additional direct path from productivity to species diversity is added in addition to the effects of diversity, soil nutrient regime, and stand age on AGB. The coefficients are standardized prediction coefficients for each causal path. Solid lines represent significant paths ($P \leq 0.05$) and dash lines for non-significant paths ($P > 0.05$).
### Table S3-1 Summary of the characteristics included in calculating the numeric life-history diversity index.

Information extracted from USDA plant database. Online available: http://plants.usda.gov/charinfo.html

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active growth period</td>
<td>Seasonal period in which plants have their most active growth</td>
</tr>
<tr>
<td>C:N ratio</td>
<td>Percentage of organic carbon divided by the percentage of total nitrogen in aboveground organic material</td>
</tr>
<tr>
<td>Growth rate</td>
<td>Growth rate after successful establishment relative to other species with the same growth habit</td>
</tr>
<tr>
<td>Known allelopath</td>
<td>Plant species shown to be allelopathic to at least one other species</td>
</tr>
<tr>
<td>Leaf retention</td>
<td>Tree, shrub, or sub-shrub retain its leaves year round</td>
</tr>
<tr>
<td>Lifespan</td>
<td>Expected lifespan (in years) of a perennial plant relative to other species with the same growth habit</td>
</tr>
<tr>
<td>Nitrogen fixation</td>
<td>How much nitrogen is fixed by this plant in monoculture</td>
</tr>
<tr>
<td>Re-sprout ability</td>
<td>Woody perennial re-sprout following top (above ground biomass) removal</td>
</tr>
<tr>
<td>Shape and orientation</td>
<td>Growth form or predominant shape of an individual plant</td>
</tr>
<tr>
<td>Adapted to coarse textured soils</td>
<td>Capability to establish and grow in soil with a coarse textured surface layer</td>
</tr>
<tr>
<td>Adapted to fine textured soils</td>
<td>Capability to establish and grow in soil with a fine textured surface layer</td>
</tr>
<tr>
<td>Adapted to medium textured soils</td>
<td>Capability to establish and grow in soil with a medium textured surface layer</td>
</tr>
<tr>
<td>Anaerobic tolerance</td>
<td>Relative tolerance to anaerobic soil conditions</td>
</tr>
<tr>
<td>CaCO3 tolerance</td>
<td>Relative tolerance to calcareous soil</td>
</tr>
<tr>
<td>Cold stratification required</td>
<td>Cold stratification significantly increase the seed germination percentage of this plant</td>
</tr>
<tr>
<td>Drought tolerance</td>
<td>Relative tolerance of the plant to drought conditions</td>
</tr>
<tr>
<td>Fertility requirement</td>
<td>Relative level of nutrition (N, P, K) required for normal growth and development</td>
</tr>
<tr>
<td>Fire tolerance</td>
<td>Relative ability to resprout, regrow, or reestablish from residual seed after a fire</td>
</tr>
<tr>
<td>Frost free days</td>
<td>The minimum average number of frost-free days within the plant’s known geographical range</td>
</tr>
<tr>
<td>Hedge tolerance</td>
<td>Relative tolerance of woody perennials to hedging (close cropping) by livestock or wildlife</td>
</tr>
<tr>
<td>Moisture use</td>
<td>Ability to use available soil moisture relative to other species in the similar soil moisture regime</td>
</tr>
<tr>
<td>Character</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Root depth minimum inches</td>
<td>The minimum depth of soil (in inches) required for good growth</td>
</tr>
<tr>
<td>Salinity tolerance</td>
<td>Relative tolerance to soil salinity</td>
</tr>
<tr>
<td>Shade tolerance</td>
<td>Relative tolerance to shade conditions</td>
</tr>
<tr>
<td>Temperature minimum (°F)</td>
<td>The lowest tolerable temperature recorded in the plant’s historical range</td>
</tr>
<tr>
<td>Bloom period</td>
<td>Seasonal period in which the plant bloom the most</td>
</tr>
<tr>
<td>Fruit seed abundance</td>
<td>Amount of seed produced by the plant compared to other species with the same growth habit</td>
</tr>
<tr>
<td>Fruit seed persistence</td>
<td>Fruit or seed generally recognized as being persistent on the plant</td>
</tr>
<tr>
<td>Seed spread rate</td>
<td>Capability of the plant to spread through its seed production</td>
</tr>
<tr>
<td>Seedling vigor</td>
<td>Expected seedling survival percentage of the plant compared to other species with the same growth habit</td>
</tr>
<tr>
<td>Vegetative spread rate</td>
<td>Spread rate, which a plant can spread compared to other species with the same growth habit</td>
</tr>
<tr>
<td>Palatable browse animal</td>
<td>Relative palatability of this plant to browsing animals</td>
</tr>
</tbody>
</table>
REFERENCE


USDA, and NRCS. 2013. The PLANTS Database National Plant Data Team, Greensboro, NC 27401-4901 USA.


