Restoring Grasslands in Southern Ontario Sandpits: Plant and Soil Food Web Responses to Arbuscular Mycorrhizal Fungal Inoculum, Biochar, and Municipal Compost

by

Brian Matthew Ohsowski

B.Sc., Eastern Michigan University, 2003
M.Sc., Eastern Michigan University, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE COLLEGE OF GRADUATE STUDIES

(Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

May 2015

© Brian Matthew Ohsowski, 2015

Abstract

Habitat destruction and land use change are among the anthropogenic influences affecting many ecosystems. Sandpit mining often restricts grassland plant restoration efforts due to the abioticially stressed mine substrate, a lack of viable plant symbionts, and disrupted multi–trophic interactions in soil food webs. Recently excavated sandpits can be ameliorated with soil amendments and arbuscular mycorrhizal fungal inoculum to address these degraded substrate conditions, potentially improving plant performance and accelerating soil food web development. This dissertation describes the results of a multi–year grassland restoration project established in southern Ontario that optimized industrial–scale grassland restoration protocol in post–extraction sandpits. This research tested the effect of soil amendment rate (municipal compost, biochar) and arbuscular mycorrhizal (AM) fungal inoculum (*Rhizophagus irregularis*) in a grassland plant plug trial and a seed application trial. In the plant plug trial, the multi–year effects of the experimental treatments on plant growth, AM fungal colonization of roots, soil microbial biomass (i.e. bacteria and fungi) and soil animal abundance (i.e. nematodes, Collembola, and mites) were explored over two growing seasons. In the plant plug trial, $20\,\mathrm{T\,ha^{-1}}$ (tons hectare^{-1}) of compost mixed with a low rate of biochar ($10\,\mathrm{T\,ha^{-1}}$) yielded the largest positive effect on total plant biomass, microbial community biomass, and soil animal abundance after two growing seasons. AM fungal inoculation did not influence total plant biomass or soil food web development during this trial. In the seed application trial, the multi–year effects of increasing rates of compost and biochar ($0\,\mathrm{T\,ha^{-1}}$ to $40\,\mathrm{T\,ha^{-1}}$) were explored for total plant cover over three growing seasons. AM fungal inoculation combined with high rates of compost ($20\,\mathrm{T\,ha^{-1}}$ and $40\,\mathrm{T\,ha^{-1}}$) and biochar ($20\,\mathrm{T\,ha^{-1}}$ and $40\,\mathrm{T\,ha^{-1}}$) resulted in the highest plant cover over three growing seasons compared to controls. Our results indicate that: (1) co–amending mine substrates with compost, biochar, and AM fungal inoculum are practical land management tools that improve grassland plant growth while increasing soil food web development, (2) AM fungal inoculum increases plant cover when applying seed with high rates of compost + biochar, and (3) amending post–mine substrates with biochar as a solitary amendment may increase biotic stress in the sandpit environment during restoration. I suggest that restoration practitioners emphasize soil community development in tandem with plant community growth when restoring sandpits to maximize restoration success.

Preface

The introduction (Chapter 1) is an adaptation of a published review in the Journal of Applied Soil Ecology (Ohsowski et al. 2012). The published work, *The potential of soil amendments for restoring severely disturbed grasslands*, was wholly drafted by Brian M. Ohsowski with editorial comments by Drs. Miranda Hart, John Klironomos and Kari Dunfield. The research for this study was carried out in a post–extraction sandpit located near Port Rowan, Ontario, Canada (*Chapter 3*–*Chapter 4*). The Ontario Aggregate Research Corporation (TOARC) graded the site and installed a nine–wire fence before plot installation in 2010. I designed and implemented both experiments (i.e. plant plug trial and seed application trial). My field assistant, André Audét, and I manually installed the plots and plant plugs in the summer of 2010. In May of 2011, I applied the seed and AM fungal inoculum to establish the seed application trial. I collected plant biomass data with the help of field assistants in the summer of 2011 and 2012. In *Chapter 2*, I designed and implemented the partial least squares regression experiment. Plant data was collected with my field assistant, Sarah Kruis, in September 2012. I conducted the statistical analysis and composed the manuscript. In *Chapter 3*, I collected plant biomass data, soil cores, and performed root washing/preservation for AM fungal percent colonization with help from André Audét. AM fungal colonization was conducted by the Soil Analysis Laboratory, University of California, Riverside. In *Chapter 4*, I collected soils for the soil food web analysis in the plant plug trial. Soil organisms were measured by the Soil Analysis Laboratory, University of California, Riverside. I performed all linear model analyses for soil organisms. Dr. Anita Antoninka at Northern Arizona University assisted me with the proper implementation and execution of the structural equation models in the statistical program AMOS. I created the *a priori* hypotheses, co–ran the structural equation models, interpreted the data, and composed the manuscript. All thesis chapters were written with the guidance of Drs. Miranda Hart and John Klironomos. Thesis chapters were reviewed by the members of my supervisory committee: Drs. Melanie Jones and David Scott from the University of British Columbia Okanagan Campus and Dr. Kari Dunfield from the University of Guelph in Guelph, Ontario.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgments

I would like to express my gratitude to the faculty, staff, and students at the University of British Columbia and the University of Guelph who helped make this degree possible. In particular, the following people have led to the culmination of this degree:

- I would like thank my supervisors, Drs. John Klironomos and Miranda Hart, for their encouragement, advice, constructive criticism, and friendship throughout my degree. I would not have achieved this milestone and become the person I am today without you.
- To my committee members, Drs. Melanie Jones, David Scott, and Kari Dunfield, your comments, constructive criticisms, and support have been essential to the development and completion of this enormous undertaking.
- To my field assistants, Andre Audet, Jeremy Booth, Nicola Day, Sarah Kruis, your tireless efforts and enduring friendship during long days in the field have been essential to my project's completion. You are as much a part of this project as I am.
- To Josh Stumpf, thank you for your support during the R learning curve. My gratitude for your friendship and patience knows no bounds.
- To the Mary Gartshore and Peter Carson, your knowledge of the restoration projects and local flora has been invaluable when developing my research.
- To the Nature Conservancy of Canada, especially Tom Bradstreet and Wendy Cridland, thank you for donating the land to conduct my research and develop my project.
- To TOARC, especially Danielle Solondz and David Sterrett, thank you for your financial support and editorial comments throughout this project.
- I would like to thank Dr. Kari Dunfield, Mike Mucci, and Tannis Slimmon at the University of Guelph and Dr. Scott Petrie at Bird Studies Canada. You provided me with a scientific home base to conduct my research. I am eternally grateful.
- Thank you to my parents, Christine Sullivan and Larry Ohsowski, for your continued support as I have pursed the life of a perpetual student.

Dedication

To my family, friends, and colleagues...

Chapter 1

Introduction

Human–induced disturbance is pervasive among all ecosystems as the result of waste accumulation, industrial pollution, resource extraction, and urban sprawl (Hannah et al. 1995). Previous land use, ranging from industrial spoils (e.g. mine tailings, contaminated brownfields) to road construction, dictates the approach of a restoration project (Jackson and Hobbs 2009). For my purposes, I define a severely disturbed landscape as an area manipulated in such a way that the pre–existing habitat can no longer be maintained. I will focus on the restoration of grassland vegetation and ecosystem processes after post–mine sand extraction.

The definition of restoration success is largely dependent upon the goals of the restoration practitioner. Goals can range from achieving diversity indices (e.g. organism richness and abundance), vegetative structure (e.g. percent cover, biomass, vegetative profiles), or ecosystem process reestablishment (e.g. nutrient cycling and soil stabilization) (Ruiz-Jaen and Aide 2005). The current paradigm in restoration tends to be phytocentric while under–emphasizing belowground food webs (van der Heijden et al. 2008, Kardol and Wardle 2010) and soil ecological knowledge (Callaham Jr. et al. 2008). Furthermore, restoration projects tend to evaluate short–term outcomes for vegetative and microbial production, as well as soil processes.

Practitioners need viable techniques that influence the recovery of the entire ecosystem. After severe disturbance (i.e. post–mine areas), edaphic conditions and soil communities may not support diverse plant communities. The addition of inoculum and soil conditioners can address some components of the soil environment. Soil amendments should create more suitable conditions for diverse and productive plant communities. With an ecosystem–level approach to restoration, native plant production is the consequence of the restoration practice, not the focus.

1.1 Landscape restoration and successional theory

1.1.1 Historical context

Historically, ecological succession has been viewed in terms of stable, climax community endpoints (Clements 1916, Odum 1969). Current thought recognizes that community diver-

sity is shaped by environmental fluctuations at large spatial, temporal, and organizational scales (Pimm 1991). Successional pathways can be multi-directional, driven by stochastic processes and disturbance, thus long-term community stability will never be maintained (Glenn-Lewin and van der Maarel 1992). This implies that an ecosystem has multiple, alternative stable states separated by unstable transitions (Scheffer et al. 2001). Alternative stable states depend upon the surrounding biotic community, order of organism arrival, and inherent system randomness. In terms of restoration, degraded systems are often in a persistent stable state (Suding et al. 2004). Plant establishment and soil building in degraded habitats may be slow to recover by natural successional processes without human intervention.

Restoration ecology and successional theory often address similar questions, albeit from different perspectives. Successional pathways are comprised of temporal changes in community assembly, biodiversity, and biogeochemical cycles (Walker et al. 2007). Habitat restoration manipulates these processes to accelerate target community establishment (Harris, 2009). Successional research, generally confined to one ecosystem, addresses time scales related to vascular plant life history (10–200 years). In contrast, landscape restoration operates on broad spatial scales (e.g. altitude gradients, moisture gradients, catchment basins) (del Moral et al. 2007), focusing on the duration of human involvement (120 years).

A practical application of successional mechanisms in restoration has not been broadly developed for practitioners (Walker et al. 2007). Restoration ecologists must acknowledge the potentially persistent stable state of degraded systems. Feedback mechanisms between biotic and abiotic factors in degraded systems may suppress plant establishment and community sustainability. For example, Ash et al. (1994) described abandoned waste areas in northwest England that had reduced plant cover and diversity after a century following disturbance.

1.1.2 Post–mine areas as primary succession models

Abandoned mine lands (e.g. ore extraction, gravel pits) are analogous to natural primary succession events such as volcanic activity or glacial retreat. The extraction process completely removes flora, fauna, and soils of the previous system. Following resource exhaustion, post–mine areas are typically characterized by low soil organic matter (SOM) content, low fertility, and poor physio-chemical and biological properties (Bradshaw 2000). The resulting raw substrate (i.e. subsoils and rock material) is a stark contrast to the abiotic and biotic soil complexity of the original habitat. As a consequence, natural reestablishment of above- and belowground communities in abandoned mine areas is typically slow (Bradshaw 1997).

Suppressed regeneration of biotic communities may be due to reduced biological complexity in post–mine substrates. The deposition and subsequent heterotrophic turnover of organic matter is a critical link for facilitating plant establishment. Restoration projects in degraded soils must include attempts to rehabilitate, at least in part, biological complexity. Biological colonization requires a source of energy and nutrients, which may be initially lacking in post–mine substrates (Frenot et al. 1998). One solution is to add organic detritus containing natural microbial assemblages (e.g. bacteria (Tscherko et al. 2003, Bardgett et al. 2007), cyanobacteria (Nemergut et al. 2007), and fungi (Hodkinson et al. 2002)). These microbes actively turn over organic substrates and prime biogeochemical cycles.

1.2 An ecological context of degraded system restoration

Clearly, soil health is paramount to restoration success in devastated landscapes. Soil microbial communities play a major role in the development and sustainability of soil health (Anderson 2003). Soil health is defined as the capacity of soil to function as a living system, sustaining biotic productivity, and maintaining ecosystem services (Doran and Zeiss 2000). Soil microbial communities are well correlated with plant primary production (Bardgett and Wardle 2003, van der Heijden et al. 2008, Heneghan et al. 2008, Benayas et al. 2009) and integral in the recycling of organic matter and nutrients (Wardle et al. 2002). Decomposers (Harte and Kinzig 1993, Reynolds et al. 2003), mycorrhizal fungi (Klironomos 2002) and nitrogen-fixing bacteria (van der Heijden et al. 2008) are key soil functional groups in the rhizosphere (i.e. soil area directly influenced by plant tissues and secretions). Soil microbial communities are also important for soil stabilization via stable aggregate formation (Rillig 2004, Six et al. 2004). These factors can ultimately mediate successional dynamics and plant community composition (Wardle et al. 2004), thus contributing to the reestablishment of natural systems in severely disturbed landscapes.

Edaphic characteristics, resource availability, and soil microorganisms mediate aboveground biotic responses to include primary productivity (Baer et al. 2004), organic matter decomposition rates (Smith and Bradford 2003), and plant community structure (Baer et al. 2003, Heneghan et al. 2008). CENTURY (Parton et al. 1993), an established ecosystem–level model of plant–soil biogeochemical cycles, models the links among plant productivity, decomposition, climate and land management options. Among its many functions, CENTURY emphasizes the role of carbon management decisions under natural and agricultural scenarios. Restoration projects that appropriately manage soil organic matter dynamics and soil microbial feedbacks may increase production and carbon storage in disturbed habitats (Ojima et al. 1993). Practitioners should emphasize soil carbon cycles and microbial processes in tandem with plant establishment in damaged ecosystems (Cairns 2000).

1.3 Grassland vegetation

Grassland productivity varies with habitat classification, ranging from shortgrass steppe (least productive) to tallgrass prairies (most productive) (Knapp and Smith 2001). Grassland productivity is ultimately dictated by the availability of three limiting resources: light, water, and nitrogen (Baer et al. 2003). Resource availability is determined by patterns in precipitation (Sala et al. 1988), soil characteristics (Briggs and Knapp 1995), herbivory (Knapp et al. 1999), and periodic fires (Knapp and Seastedt 1986). Plant production in grasslands will ultimately depend upon adaptations to spatial and temporal availabilities of these limiting resources.

Grassland restoration in severely degraded habitats must recognize the factors that shape and maintain these communities. Grassland plants are evolutionarily adapted to the mentioned environmental context. Restoration projects incorporating locally adapted plant populations are more likely to improve rates of establishment and persistence (Pywell et al. 2002). Resulting plant communities are expected to more closely resemble natural grassland remnants and encourage the conservation of rare flora and fauna.

Four functional groups composed of herbaceous perennials dominate grassland communities: perennial C_4 grasses, C_3 graminoids (grasses and sedges), nitrogen-fixing species (primarily Fabaceae), and late summer flowering, drought-hardy composites (Asteraceae) (Kindscher and Wells, 1995). Cool season C_3 grasses have traits that provide early season plant cover, nutrient-rich plant tissues beneficial to herbivores, and have decreased light requirements ideal for shady refugia. Compared to cool season grasses, warm season C_4 grasses exhibit higher water-use efficiency, higher plant biomass potential, late season growth, and tolerance of full sun exposure (Tiessen et al. 1993). Composite forbs are integral in rapidly colonizing open soil (especially after grazing or fire disturbances), supporting pollinator populations, and driving overall plant community diversity indices (Pokorny et al. 2004). Forbs in the legume family (Fabaceae) form a symbiotic relationship with nitrogen-fixing bacteria. Nitrogen-fixing bacteria are found within legume root nodules, and convert biologically unavailable atmospheric N_2 gas into forms of nitrogen usable by plants. In exchange for usable nitrogen, the plant delivers a supply of nutrition in the form of carbohydrates. Nitrogen-rich legumes within grasslands can contribute to the total nitrogen pool of soils during growth and after senescence (Oelmann et al. 2007). Soil nitrate and ammonium levels are usually limited within grasslands due to rapid utilization and immobilization by primary producers and microbial decomposers (Risser and Parton 1982). The introduction of N-fixing plants may affect the structure and function of grassland systems.

Restoration projects that incorporate multiple functional groups and high numbers of species are more likely to achieve community sustainability (Piper and Pimm 2002). Long–

term ecosystem stability depends on communities containing species or functional groups that are capable of differential response to disturbance (McCann 2000, Hooper et al. 2005). Studies of grassland ecosystems indicate that increased diversity can be expected, on average, to give rise to resistance and resilience (Tilman et al. 1997, Tilman and Downing 1994). Higher species diversity may also lead to increased plant production due to species complementarity (Cardinale et al. 2007). Restoration projects that maintain high species diversity with varied functional traits could increase the likelihood of achieving long-term community stability.

1.4 Soil food webs

Microbial communities (i.e. bacteria and fungi) play a fundamental role in driving biogeochemical cycles in terrestrial ecosystems. Carbon cycling and plant nutrient availability are dictated by bacterial and fungal communities, subsequently mediating plant productivity and soil development in habitat restoration (Harris 2009). Fungivorous and bacteriovorus soil animals (i.e. grazing nematodes, Oribatid mites and Collembola) directly or indirectly consume microorganisms embedded within organic matter, thus contributing to litter breakdown and soil mixing (Lavelle et al. 2006). Microorganisms associated with litter have high nutritional value compared to detritus and become a critical food source that links fungal and bacterial communities to soil animal abundance (Bardgett and Cook 1998). In conjunction with abiotic conditions, food resource availability determines the population size of soil animals within a soil food web (Ingham et al. 1985).

Restoring a grassland ecosystem can be challenging in post–mine sandpits. Land managers often emphasize aboveground plant biodiversity when rehabilitating sandpits while largely ignoring the contribution of soil biota to plant community productivity. Adopting a holistic restoration strategy that focuses on soil recovery can positively influence the restoration trajectory of a degraded area (Heneghan et al. 2008). Thus, re–establishing a functional detrital food web is an essential component of recovering soils in severely degraded systems.

Mine activities, such as sandpit excavation, disrupts and diminishes the multi–trophic interactions among soil biota, consequently reducing the beneficial ecosystem services associated with soil food webs (de Vries et al. 2012, Araújo et al. 2013, Zhao and Neher 2013). In the case of aggregate extraction sites, substrate conditions are stressful as these systems lack high concentrations of essential nutrients, soil organic matter, and a large water–holding capacity. Thus, the growth of all organisms is often restricted and the ecological connections between soil-plants-microbes are usually severed (Maiti 2013). Successfully restoring soil food webs in conjunction with plant communities depends on alleviating abiotic stress in mine substrate (McKinley et al. 2005).

To address the depauperate conditions of post–mine areas, reclamation tools such as soil amendments (i.e. compost and biochar) and microbial inoculants (i.e arbuscular mycorrhizas) are often required to increase plant production (Refer to Chapters 1 and 3). Incorporating organic soil amendments and arbuscular mycorrhizal inoculants strengthens the feedback links among plants-soils-microbes to ultimately re–establish decomposition cycles and accelerate soil development (Elkins et al. 1984, Ros et al. 2003). To accomplish this, researchers must recognize the links among restoration protocol (i.e. amendment application), soil microorganisms, soil animals, and ecosystem functioning when revegetating severely disturbed areas (Coleman and Whitman 2005). To date, an explicit protocol useful to land managers that will increase soil microbial biomass and soil animal abundance in post–mine grassland restoration does not exist.

1.4.1 Soil microbial communities as indicators for post–mine substrate recovery

Fungi and bacteria are key decomposers in soils that are responsible for nutrient availability, nutrient transformations, and litter breakdown. Therefore, estimating the biomass of these microbial constituents in recovering soils are a proxy to soil function (Visser and Parkinson 1992, Karlen et al. 1997) and can be used to evaluate ecological restoration soils (Harris 2009). Assessing the biomass of microbial assemblages elucidates the stage of soil development and food resource availability for grazing soil animals in the soil food web.

The ratio of fungal and bacterial biomass can be a useful tool to assess soil development status. Severe anthropogenic disturbance, such as mining, often shifts the dominance of soil microbial communities from fungal–dominated to bacterial–dominated assemblages due to poor physiochemical conditions in recently exposed substrates (Frey et al. 1999, Bailey et al. 2002, Mummey et al. 2002). Bacterial–based soil food webs are common in ecosystems with poorly developed, low organic matter soils such as mined or conventional agriculture landscapes (Kardol and Wardle 2010). In contrast to mine systems, natural grassland soils are dominated by fungal decomposers due to the higher volume of complex organic matter from nutrient–rich litter inputs (Bardgett and McAlister 1999, Harris 2009). Reduced disturbance facilitates an extensive hyphal fungal network, allowing fungi to access spatially separated limiting nutrients in the soils via fungal translocation (Beare et al. 1992). Restoration ecologists should target fungal–dominated systems to indicate a successful grassland restoration (Bardgett and McAlister 1999, Smith et al. 2003).

1.4.2 Soil animals as indicators for post–mine substrate recovery

Soil animals (i.e. nematodes, Collembola, mites) are ecosystem engineers that enhance biological, chemical, and physical soil properties that benefits the growth of plants. Lavelle et al. (2006) suggested that soil animals enhance nutrient release in the plant rhizosphere, stimulate mutualistic associations, and positively affect soil physical structure. After a severe disturbance, soil animal communities are removed or severely reduced (Curry and Good 1992), thus the ecosystem services provided by these soil organisms are non-existent. Restoration practitioners should create soil recovery plans that promote high densities of diverse soil animals to facilitate multi–trophic interactions among soil microbial communities and soil animals.

Multi–trophic interactions in soil food webs are based on soil animal feeding preferences. Grazing soil animals (i.e. bacteriovorus and fungivorous nematodes, Collembola, and Oribatid mites) depend upon soil microorganisms as a food resource. Grazing nematodes alter soil nutrient cycles and influence organic matter decomposition by ingesting large quantities of fungal and bacterial communities residing in plant litter (Yeates et al. 1993). The dominant microbial community in soils has been shown to determine the abundance of nematode feeding groups. Greater bacterial production supports mainly bacteriovorus fauna (Hendrix et al. 1986) while fungivorous feeding soil animals are expected to thrive in fungal–rich soil (Beare et al. 1992). Comparatively, Collembola and Oribatid mites feed on soil microbial communities but also ingest litter, influencing microbial populations and litter turnover rates in soil systems (Hättenschwiler et al. 2005, Frouz et al. 2006). Collembola, often considered consumers of fungi and bacteria, are also known to be predatory on rotifers with some species consuming nematodes when available (Wallwork 1976, Lee and Widden 1996). Thus, the activities of grazing soil animals can ultimately influence nutrient cycles and litter retention in soil by regulating microbial decomposition rates in a restoration project.

Predators in the soil food web can have a top–down trophic cascade effect on the microbial production at the base of the soil food web (Lenoir et al. 2007). Predatory nematodes feed upon lower trophic levels such as rotifers, protists, and other soil nematodes. Grazing nematode abundance has been shown to be reduced by the feeding activities of predatory nematodes, ultimately influencing microbial productivity (Laakso and Setälä 1999). Furthermore, fungal activity has also been shown to increase when predatory mites consume Collembola in a tri–trophic interaction study (Hedlund and Öhrn 2000). Overall, the complex multi–trophic interactions in soil food webs can indirectly influence soil function and litter decomposition rates in restoration project via consumption of grazing soil animals.

Soil food webs are important indicators for determining restoration success and soil development trajectories. These organisms are invaluable to a restoration project because of the ecosystem services provided by the activities of soil animals. Therefore, soil animal

abundance has been used as indications of soil quality as they are sensitive to disturbance and land management practices (Wardle et al. 1995, Roubíčková et al. 2013). Nematodes are easily collected, respond rapidly to environmental change, and can be easily sorted into functional feeding groups based on morphology. Nematodes have been successfully used to evaluate mine area recovery in a variety of disturbance scenarios (Biederman et al. 2008, Háněl 2008, Courtney et al. 2011). Soil microarthropods are also useful indicators of successional stage and soil system recovery as population growth is reliant on food resources in their immediate environment (Ferris et al. 2001, Parisi et al. 2005). Thus, soil animal abundance should be incorporated into the restoration of degraded systems to assess the development trajectory of soils.

1.5 Techniques for improving disturbed soils in grassland restoration projects

The positive effects of soil amendments on plant and microbial production within agricultural systems, restoration projects, and greenhouse experiments have been extensively recognized. In the following section, I review three typical amendments that are widely accessible to restoration practitioners. These amendments (i.e. biochar, compost, arbuscular mycorrhizas) have had promising results both in greenhouse and field settings.

1.5.1 Vegetation–derived biochar

Application of black carbon to soils is expected to build soil organic matter, enhance nutrient biogeochemical cycles, lower bulk soil density, increase bio–available water, and reduce nutrient leaching (Shrestha et al. 2010). Black carbon consists of all C rich residues, ranging from partly charred material to graphite and soot particles, resulting from the incomplete combustion of organic materials (Schmidt and Noack 2000). Research has shown that prairie soils contain substantial amounts of black carbon resulting from a 10,000 year legacy of prairie fires (Skjemstad et al. 2002, Brodowski et al. 2005). Laird (2008) estimates that between 5% and 15% of the total organic carbon in natural Midwestern prairie soils is composed of black carbon. Within boreal forests, short–term soil fertility effects have been attributed to increased charcoal fractions in the soil after naturally occurring fires (Wardle et al. 1998).

Historically, human agricultural practices (i.e. terra preta soils in the central Amazon) have long recognized plant growth benefits of black carbon soil supplements (Glaser et al. 2002) . Terra preta literally translates to black earth in Portuguese. These ancient soils (500–7,000 YBP) have been anthropogenically amended with black carbon, bones, and ma-

nure. Compared to adjacent infertile soils (terra comum or common soils), the concentration of black carbon in terra preta soils is seventy times greater. Furthermore, these soils still exhibit three times more soil organic matter, nitrogen, and phosphorus in comparison to neighboring terra comum soils (Glaser 2007).

One form of refined black carbon being used in environmental management is biochar, or carbon–rich charcoal (Lehmann et al. 2009). To create biochar, organic materials (i.e. feedstocks) are heated to temperatures between 300 ^{o}C and 800 ^{o}C in a low oxygen environment. Anoxic conditions during heating leads to the incomplete combustion of the organic matter, thus producing biochar. Feedstocks may include agricultural wastes, forestry wastes, wood pellets, or manures. The high temperatures used in pyrolysis induce molecule polymerization within feedstocks to produce aromatic and aliphatic compounds (Sohi et al. 2009). This creates a stable product demonstrated to be a long–term carbon storage pool for atmospheric CO_2 in addition to being a beneficial soil amendment (Lehmann et al. 2006). When incorporated into soils, initial degradation of biochar by chemical oxidation and microbial processes has been noted (Bruun et al. 2008, Nguyen et al. 2008). The recalcitrant properties of black carbon stocks eventually stabilize and resist microbial degradation within soils for 100–1000+ years (Glaser et al. 2002).

Amended soils benefit from biochar's large, oxidized surface area and porous structure. Soils amended with biochar have an increased soil charge density (potential cation exchange capacity [CEC] per unit surface area) in comparison to non–amended soils (Liang et al. 2006). Biochar improves: (1) soil nutrient availability and retention (i.e. major cations, phosphorus, total nitrogen) (Lehmann et al. 2003), (2) acidic soil conditions, (3) organic matter adsorption (Shrestha et al. 2010), and (4) soil aeration (Shrestha et al. 2010).

Biochar as a soil amendment has generated promising results within agricultural systems and greenhouse experiments. Recent research has demonstrated that biochar amended soils have greater crop biomass (Rondon et al. 2007, Major et al. 2010) and enhanced biological N–fixation in leguminous crops (Rondon et al. 2007). A meta–analysis by Biederman and Harpole (2012) shows that biochar increases aboveground productivity, crop yield, soil microbial biomass, rhizobia nodulation, and soil nutrients compared to controls. The fertilizer effect induced in plants may be explained by the retention of beneficial nutrients and pH neutralization.

Indirectly, plant growth may be stimulated by increased mycorrhizal associations (Nishio and Okano 1991, Ishii and Kadoya 1994) and soil microorganism activity (Thies and Rillig 2009). Warnock et al. (2007) proposed four mechanisms that may benefit arbuscular mycorrhizal fungi in soils with biochar: (1) positively changing physio–chemical soil properties (i.e. CEC, bioavailability of phosphate [PO_{4-}] in low P soils), (2) promoting beneficial soil organisms (i.e. phosphate solubilizing bacteria, mycorrhization helper bacteria), (3)

adsorbing plant secretions that may alter mycorrhizal root colonization, and (4) providing a grazing refuge in biochar's porous structure. In general, increased soil microbial activity in biochar amended soils may also be attributed these hypothesized mechanisms for AM fungi.

Biochar soil amelioration in severely degraded landscapes has the potential to increase grassland plant production, enrich soil microbial populations, and stimulate arbuscular mycorrhizal persistence. Biochar is hypothesized to reduce nutrient leaching in well–drained soils. Nutrient retention in impoverished post–mine substrates should increase productivity by stimulating biotic–abiotic feedbacks.

1.5.2 Leaf and yard waste compost

Agricultural societies have historically recognized that ameliorating fields with compost results in improved soil conditions. Soil disturbance (e.g. mining or tillage) generally decreases SOM pools due to erosion and disruption of the biogeochemical mechanisms and microbial communities associated with SOM pools (McLauchlan 2006). When added to soils, composted material increases soil fertility by increasing: (1) soil organic carbon (Crecchio et al. 2004, Walter et al. 2006), (2) available soil nitrogen (Eriksen et al. 1999, Gabrielle et al. 2005), phosphorus (Wortmann and Walters 2007), and micronutrients (i.e. iron, copper, zinc)(Hargreaves et al. 2008), (3) water holding capacity (Movahedi-Naeini and Cook 2000), (4) cation exchange capacity (McConnell et al. 1994), (5) soil aggregation (Bresson et al. 2001, Annabi et al. 2007, Abiven et al. 2009), and (6) neutralization of acid soils (Mkhabela and Warman 2005).

Leaf and yard waste (LYW), largely composed of community organic waste, is typically composted at large scales. During aerobic LYW composting, theromphilic microbes assimilate and mineralize complex organic compounds while releasing heat, water vapor, CO_2, and ammonia waste products. The remaining non–mineralized organic material is humified to form the stable end product, compost. Civic and environmental benefits of composting LYW include waste volume reduction, microbial pathogen and weed sterilization (due to high temperatures), and odor suppression (Jakobsen 1995). LYW compost derived from municipal processing facilities is utilized in gardens, organic agriculture, land reclamation, and slope stabilization projects.

Research demonstrates direct increases to crop biomass (Montemurro et al. 2006) and nutritional quality (Allievi et al. 1993) in compost amended soils. Compost addition strongly influences soil microbial communities by increasing microbial biomass, respiration rates, and soil enzyme activity (Albiach et al. 2001). As bacterial and fungal decomposers utilize and sequester carbon in amended soils, concentrations of total nitrogen and phosphorus increase over time (Iglesias-Jiménez 2001, Wolkowski 2003). Long–term ramifications of

microbial community activity (Ros et al. 2006) and soil biochemical characteristics (García-Gil et al. 2004) due to compost ameliorations have been noted. Pascual et al. (1999) found that microbial biomass, soil basal respiration, and dehydrogenase activity recovered to levels similar to adjacent Mediterranean soils eight years after a single compost amendment. Within restoration projects specifically, compost bolstered arbuscular mycorrhizas inoculum persistence, thus benefiting native plant cover (Noyd et al. 1996, Celik et al. 2004). In semi–arid soil restorations, extensive work from Caravaca *et al.* suggest short–term (Caravaca et al. 2002b;a; 2003b) and medium-term (Caravaca et al. 2003a) influences of mycorrhizal inoculations and compost ameliorations. Sharp increases in plant primary production were attributed to the abiotic–biotic link between bioavailable phosphorus supplied by compost residues and AM fungal phosphorus uptake.

In sandy soils with low SOM, compost improves soil structure (Wahba and Darwish 2008), bioavailable nutrients (P, K, Mg) (Weber et al. 2007), total inorganic N (Busby et al. 2007), plant production (Mkhabela and Warman 2005), and soil microbial activity(Ros et al. 2006). Low SOM and poor physio–chemical properties in post–mine substrates are expected to have restricted microbial community activity and depleted nutrients. LYW compost amendments should increase microbial activity (Ros et al. 2003), mycorrhizal persistence (Gaur and Adholeya 2005), and increase plant biomass. To date, few studies have researched the effect of compost application to native plants and mycorrhizal communities in severely degraded post–mine substrates (Busby et al. 2007).

1.5.3 Arbuscular mycorrhizal fungal inoculation of grassland plants

Arbuscular mycorrhizal (AM) fungi are globally distributed soil microorganisms that form symbiotic associations with more than 80% of terrestrial plants (Smith and Read 2008). These obligate biotrophs constitute a major fraction of the plant–associated soil microbial community. In exchange for host plant–derived photosynthate, arbuscular mycorrhizas benefit plants by: (1) increasing soil nutrient acquisition and subsequent assimilation into plant tissues (especially phosphorus), (2) protecting target plant roots from pathogens, (3) enhancing seedling performance, (4) improving plant water relations, and (5) improving soil stabilization. In addition to improved target plant performance, AM fungal communities directly relate to the biodiversity of plant communities (van der Heijden et al. 1998).

Positive plant growth responses to mycorrhizas have stimulated the emergence of biotech companies promoting the use of commercially–produced AM fungal inoculum as a soil enhancement agent. In horticultural systems (Azcón-Aguilar and Barea 1997) and landscape restoration (Miller and Jastrow 1992), mycorrhizal inoculum has been recommended to increase plant growth performance. The intentional movement of mycorrhizal fungal species is growing, but the potential negative ecological ramifications of non–native arbuscular my-

corrhizal invasion are poorly understood (Schwartz et al. 2006). Evidence indicates that symbiotic associations between plants and fungus range from parasitic to beneficial depending on host plant/AM fungal pairings (Klironomos 2003). Depending on the biogeochemical context and AM fungal–plant associations within an ecosystem, AM fungal inoculation may yield positive, neutral, or negative plant growth effects in the field. Furthermore, a recent greenhouse study by Mummey et al. (2009) indicated that plant pre–inoculation with AM fungi may have unintended implications for resident AM fungal communities. AM fungal inoculum may restrict assembly potentials in resident soil AM fungal communities with divergent phylogenies, thus suppressing plant growth and foliar nutrients. As research indicates AM fungal phylogeny diversity in host plant roots directly correlates to increased plant growth responses (Maherali and Klironomos 2007), restricting native soil inoculum potentials could have ramifications to plant production and soil feedback mechanisms in a restoration project.

The ramifications of pre–inoculating native plants with AM fungal inoculum in severely degraded habitats have not been thoroughly addressed. Evidence indicates that after major soil disturbances such as agricultural tilling, native AM fungal associations are fractured and strongly diminished (Jansa et al. 2002; 2003) . The resident AM fungal community in post–mine substrates is expected to be strongly reduced compared to natural grassland soils. To date, the resident AM fungal community soil inoculum potential within post–mine substrates has not yet been identified. The AM fungal inoculum potential in severely disturbed sites should be determined by spore immigration rates, soil nutrient availability (i.e. phosphorus availability), plant identity (i.e. obligate mycorrhizal plants vs. facultative mycorrhizal plants vs. non–mycorrhizal plants) in the degraded area, and time since landscape disturbance (Allen and Allen 1980).

To date, some AM fungal inoculation research has been conducted in non-toxic post–mine reclamation areas. These field studies indicate that AM fungi benefit native plant production and establishment in severely degraded areas (Johnson 1998, Matias et al. 2009). Mycorrhizal inoculum is anticipated to benefit plant production in post–mine substrates due to a lack of an existing AM fungal community.

1.6 Review conclusions

It is imperative that restoration practitioners integrate soil ecological knowledge into the reclamation of degraded habitats. Emphasizing an ecosystem–level approach to grassland restoration in degraded areas should reduce landscape recovery time and reduce plant failure. When used in combination, the addition of mycorrhizal fungi, biochar, and compost approaches the goal of a viable soil environment for sustainable plant growth.

It is clear that soil amendments are necessary to restore severely disturbed landscapes in a reasonable time–frame. A checklist or key could be developed to facilitate identification of factors that are important for determining the most appropriate amendments and practices. Application rates of biochar and compost could be determined experimentally to establish a feasible restoration protocol under a variety of restoration scenarios. Since universal application of soil microorganisms may not always be beneficial, more studies testing the use of locally bolstered inoculum sources should be conducted to eliminate the environmental impact of foreign inoculum.

The list of amendments discussed is by no means complete. Other amendments may include inoculation (e.g. nitrogen–fixing bacteria, earthworms), and organic materials (e.g. biosolids, hydrogels, paper mill sludge). Further research into the integration of these amendments into severely degraded landscapes during restoration projects needs to be conducted. As we make advancements in biotechnology and soil conditioners, we can reduce long–term maintenance costs and create a foundation for sustainable above– and below-ground communities.

1.7 Research objectives

The research conducted for my Ph.D. dissertation tests the efficacy of industrially feasible rates of soil amendments (i.e. compost, biochar) and a commercial AM fungal inoculum (*R. irregularis*) on grassland plant restoration and soil food web development in a post–extraction sandpit. To accomplish this, I installed a large–scale restoration research site near Port Rowan, Ontario, Canada in the summer of 2010. Two planting strategies were implemented: greenhouse grown plant plugs and direct seeding. In September of each year, plant response data was collected for two growing seasons in the plant plug trial (2011–2012) and three growing seasons in the seed application trial (2011–2013). A soil food web analysis was conducted for data collected from the plant plug experiment in September 2012. The results of this dissertation will directly inform land management protocol when restoring grassland plants in post–mine sandpits. The three main research objectives of this Ph.D. dissertation are as follows:

Objective #1 *Develop a minimally destructive statistical method to increase measurement accuracy and reduce data collection time when estimating aboveground plant biomass.* I hypothesized that plant biomass predictive models using multiple morphological plant traits would be more accurate and robust compared to single plant trait model estimates. The rationale for this hypothesis that increased information would be acquired on plant morphology in the field, thus leading to higher prediction accuracy for each plant species.

Objective #2 *Determine the multi–year plant response of both planting strategies to soil amendments and the commercial AM fungal isolate in a post-mine sandpit.* In both experiments, I hypothesized that: (a) AM fungal inoculation, compost, and biochar addition would increase total plant community mass in the plug trial and vegetative cover in the direct seeding trial compared to non–amended controls, and (b) plots with the highest rates of compost + biochar and inoculated with the AM fungal isolate will yield the highest plant responses. The rationale for these hypotheses was that water and nutrient stress of the plants would be ameliorated when the mine substrate is ameliorated. When all amendments are combined, maximal plant response was anticipated due to higher nutrient inputs and retention from the soil amendments and increased nutrient acquisition due to the plant–fungal symbiosis.

Objective #3 *Determine the soil food web response to the addition of soil amendments and a AM fungal isolate in sandpit substrate.* I hypothesized that: (a) adding compost, biochar, and AM fungal inoculum singly would increase soil microbial biomass (i.e. bacteria and fungi) and soil animal abundance (i.e. nematodes, Collembola, and mites), and (b) co–amending sandpits with compost, biochar, and AM fungal inoculum will show the greatest response in soil biota. The rationale for these hypotheses was that nutrient stress would be alleviated due to increased belowground production and rhizosphere activity, and organic matter in sandpit substrate.

Chapter 2

Improving Plant Biomass Estimation

2.1 Background

Aboveground plant biomass is an important measurement relevant across multiple disciplines. Plant biomass is often considered a good approximation of productivity, especially in grassland communities (Hector et al. 1999). Directly measuring plant biomass, however, requires destructive sampling, thus severely disrupting the plant community of interest and burdens the researcher with a large labor cost. Harvesting, drying, and weighing a large volume of plants restricts a researcher's ability to minimize plant community destruction, rapidly collect a high volume of data, and track plant production of an individual at multiple time points (See Table 2.1 for several detailed scenarios). Therefore, minimally destructive estimation methods are useful when data collection is time sensitive, labor force constraints exist, large–scale plant harvesting is impractical, and sampling design requires repeated measures (Catchpole and Wheeler 1992).

Minimally destructive measures have been suggested as alternatives to harvesting plant biomass, with varying success. Techniques include measuring community attributes per unit area (e.g. percent cover, point intercept transects, photographic image analysis (Byrne et al. 2011)) and individual estimates of plant mass (e.g. simple linear regression estimates and allometric equations). However, such approaches tend to have high variability and inherent subjectivity, reducing the predictive power of the models. For example, several authors have found good correlations between biomass and point–intercept methods but high variability still exists depending upon sampling intensity (Jonasson 1988, Glatzle et al. 1993, Vittoz and Guisan 2007). This experimental error due to the sampling method represents potentially important biotic variation that is unaccounted for in a study.

Table 2.1: Experimental or observational situations to employ non–destructive biomass estimation.

Annual Net Primary Production (ANPP): Plant species respond differently to seasonal and inter-annual environmental variation resulting in different biomass maxima throughout the growing season (Briggs and Knapp 1995, Fay et al. 2003, Moya-Larao and Corcobado 2008). Measuring ANPP from one or two time points, such as mid/ late season community biomass, underestimates ANPP due to fluctuations in biomass by different species throughout the season. With minimally destructive methods, researchers track plant species biomass throughout the year to accurately determine ANPP of individual plant species without severely disrupting the plant community.
Repeated Measures / Time Series Experiments: Researchers are interested in spatial and temporal aspects of plant growth in a time–series. Phenological responses of individual plants can be tracked over the duration of an experiment with minimally destructive prediction methods. The establishment of a standard curve to predict biomass at a plant's life stage would allow for high–throughput sample replication while being minimally invasive.
Sampling Accuracy: Increased accuracy and precision in biomass estimates leads to higher confidence in results. Increased measurement accuracy reduces variability in dataset predictions, thus reducing statistical noise due to prediction error. Error reduction enhances the ability to detect discrete differences among experimental treatments and controls.

2.1.1 Techniques to predict plant biomass

Two main methods have been used to predict biomass of a plant individual using minimally destructive techniques. The first approach uses published allometric equations developed mainly to estimate tree biomass. The second approach creates a predictive standard curve from the relationship between a measured plant trait such as plant height and plant mass. These predictive methods have been shown to be unreliable in their ability to consistently predict plant biomass. Thus, I introduce a new multivariate statistical approach, partial least squares (PLS) regression, to increase predictive accuracy of estimating plant mass in the field.

Historically, allometric equations have been used to estimate plant biomass in the field. Allometric equations are mathematical functions published in the literature that are often limited to the estimation of woody plant biomass. These equations are constructed using easily measured predictor variables such as diameter at breast height and total plant height to estimate biomass (Picard et al. 2012). The advantage of allometric equations is that a

researcher can apply these equations to estimate tree biomass without creating a standard curve. However, allometric equations developed to measure herbaceous plants are relatively scarce and can be highly variable in prediction accuracy when available (e.g. (Elliott and Clinton 1993)).

When using published allometric equations, a researcher must be aware of variability in plant morphology due to experimental treatments and local environmental conditions that will reduce prediction performance. In addition, allometric equations may not measure predictor variables that optimally estimate plant biomass (Chave et al. 2004). Thus, published allometric equations are not readily usable when measuring plants under unique experimental conditions.

Alternatives to allometric equations have been used by a measuring predictor variable regressed against plant mass (Catchpole and Wheeler 1992). Plant traits that describe a strong linear relationship between plant biomass and a predictor variable are measured to create a standard curve. Once a standard curve is established, rapid measurement of similar vegetation is straightforward. When using this method, the researcher must balance the precision of the mathematical relationship, conform to statistical assumptions, and weigh the costs associated with direct versus indirect measurements.

When establishing a standard curve for predicting plant biomass, the strength of the relationship between a plant trait and harvested biomass determines the predictive performance of the model. To achieve the most accurate prediction, the standard curve must meet the following assumptions: a linear relationship, equal residual variance (i.e. homoscedasticity), normal distribution of residuals, no highly influential outliers, and no strong multicollinearity among predictor variables (Zuur et al. 2007). Multiple regression models would increase plant estimation accuracy by taking into account several plant traits but must be approached with caution. When predicting plant biomass, morphological plant traits are often strongly correlated, violating the assumption of predictor independence and potentially reducing predictive power of the external samples (Graham 2003).

Partial least squares regression (PLS) is a statistical method commonly used in computational chemistry that predicts a response variable from multiple, collinear predictor variables (Wold et al. 2001). PLS is a robust generalization of multiple linear regression and principle component regression that extracts orthogonal factors (i.e. latent variables) from predictors while taking into account the response variable (Abdi 2010). PLS is becoming increasingly popular in ecological data analysis (Carrascal et al. 2009). PLS is valuable when two conditions exist: (1) the dataset has a high number of predictor variables relative to the number of samples and/or (2) high collinearity amongst predictor variables exists such as the case for most plant biomass estimations. Continuous and categorical data can

be used simultaneously in PLS, an essential feature when measuring morphological aspects of plants.

PLS is a powerful statistical method that will maximize the predictive accuracy and precision of plant mass estimation in the field. PLS should be used in combination with variable reduction techniques such as Bayesian Information Criterion (BIC) model selection to optimize the predictive model and reduce field measurements (Mehmood et al. 2011). Compared to destructive plant harvesting, minimally destructive PLS plant biomass estimation will increase sampling volume, reduce data collection time, and minimize labor. The statistical assumptions associated with PLS make it well–suited to estimating aboveground plant biomass. Thus, using a multivariate plant biomass prediction approach with PLS ultimately increases measurement accuracy and precision and will outperform other methods in the field.

In this chapter, I propose a highly accurate, customizable approach to estimating plant biomass with minimal plant destruction in the field using PLS. My method achieves this by collecting a set of simple measures from the plant population under study. I propose that PLS is an accessible and powerful technique for the estimation of plant biomass compared to other plant estimation alternatives. Thus, I hypothesize that partial least squares regression will increase plant biomass prediction performance in three distinct plant growth forms compared to simple linear regression models using a single predictor variable, plant height.

2.2 Methods

2.2.1 Species selection and data collection

I selected three plant species representing extreme differences in morphology/growth habit in order to determine the robustness of my approach. I tested a small shrub (*Cornus racemosa* Lam., grey dogwood), a tussock grass (*Sporobolus cryptandrus* Torr., sand dropseed), and a fern with radial rhizomes (*Osmunda claytoniana* L., interrupted fern). In September 2012, 41 individuals from each plant species were selected from southern Ontario's hardwood forest near Simcoe, Ontario, Canada along 50 m transects. Plants along each transect were selected to capture the range of sizes and shapes present in the population to establish a standard curve.

2.2.2 Measured plant traits

Easily measured plant traits that had a potential to estimate plant biomass were customized for the growth form of each species. These selected traits were based on measurement variations stemming from plant height and circumference, structural counts (i.e.

number of sand dropseed seed heads), and a weighted plate estimator to approximate plant density (Rayburn and Rayburn 1998). A weighted plate apparatus was constructed using 40.0 cm length x 40.0 cm width x 3.2 mm depth acrylic plexiglass sheet and a 122.0 cm length x 1.9 cm width wooden dowel. A large hole was drilled into the center of the plexiglass plate to insert the wooden dowel. In addition, four small holes were drilled into each corner of the acrylic sheet to attach strings to raise and lower the plate. When taking measurements, the wooden dowel was set–up near the center of the each plant and the weighted plate was lowered until four leaves touched the plate. Plant height was recorded by measuring height of the plate from the soil surface. After measuring plant height, the weighted plate was then lowered to rest upon the plant. The resting plate height gives an estimate of plant density and plate height at rest was recorded. Circumference measures were made by gathering and compressing the plant material and measuring the circumference of the vegetation using a cloth measuring tape. Circumference measurements were also collected in a similar manner at half of the plant height and at 30 cm from the soil surface. This measurement was considered as the plant's basal circumference may be highly variable. A list of measurements collected for each species is given in Figure 2.1. After measuring each individual, aboveground biomass was clipped at the soil surface, dried in a forced air oven at 60 °C for approximately 3–6 days, and weighed. Plants were considered to be dry when the biomass weights stabilized within ± 0.3 g.

In the case of the interrupted fern, plant morphology was distinctly different from the sand dropseed and grey dogwood growth forms. Interrupted ferns have radial rhizomes with multiple fronds growing in a circular cluster. Plant measurements were taken per frond and subsequently averaged and summed per individual. The average and sum measurements were the variables used to create the statistical models.

2.2.3 Model creation

Two datasets, a training dataset and a test dataset, were created from the 41 individuals measured and weighed for each plant species. The training dataset was built by randomly selecting thirty–five plants from each plant population and subsequently used to calculate the standard curve. The remaining six plants were used as external data points to establish a test dataset. Samples from the test dataset were not included in the creation of the predictive model to remove any potential influence when generating the standard curve. The same training and test datasets were used when creating the partial least squares regression (PLS) and linear regression (LR) prediction models.

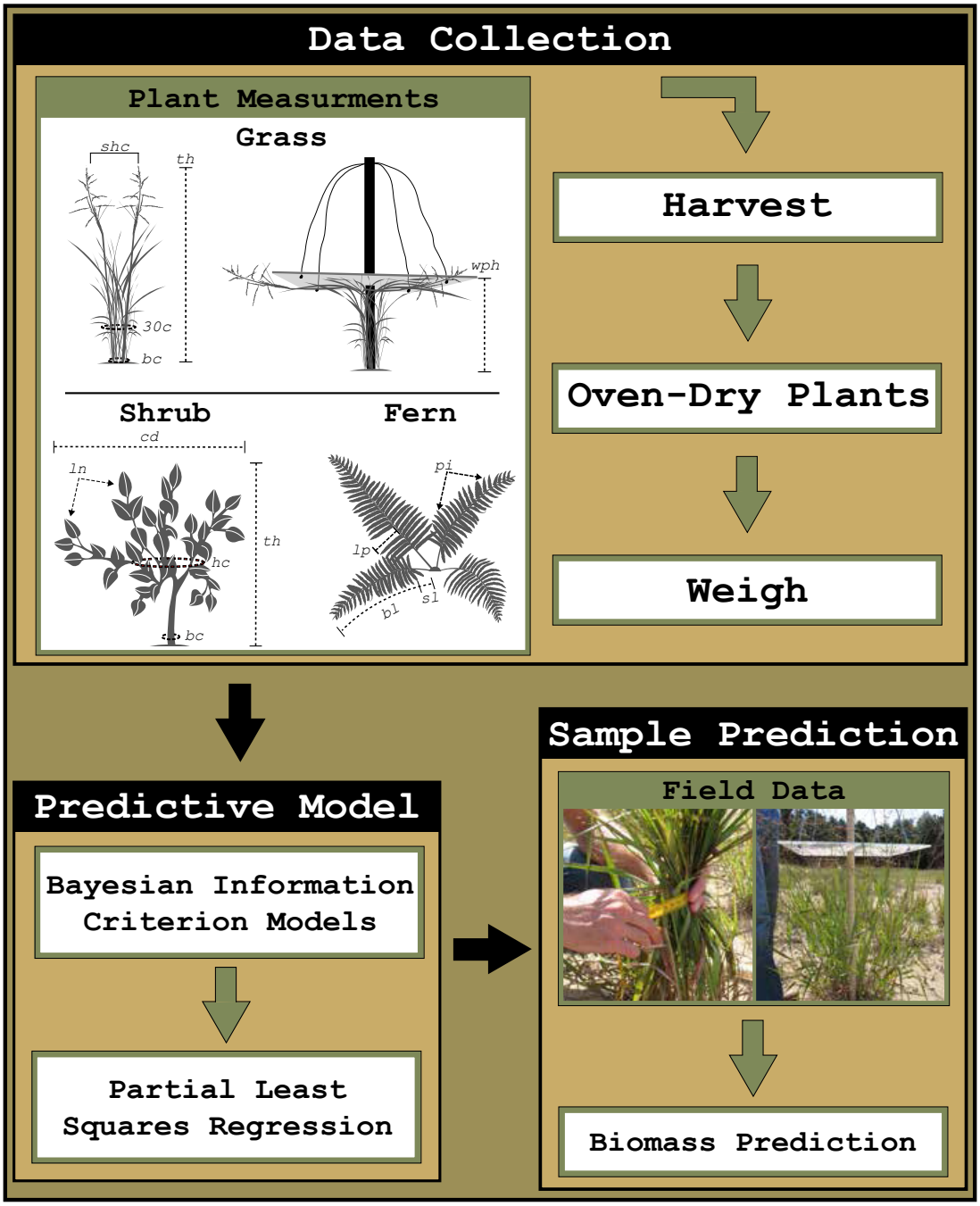


Figure 2.1: Workflow for predicting plant biomass with partial least squares regression. Plant measurement abbreviations: **30c** = circumference at height of 30 cm; **bc** = basal circumference; **bl** = fern blade length; **cd** = maximum canopy diameter; **fl** = frond length (blade length + stipe length); **hc** = circumference at half plant height; **ln** = leaf number; **lp** = longest pinna per blade; **pi** = pinnae number per blade; **shc** = seed head count; **sl** = stipe length; **th** = total plant height; **wph** = resting height of falling plate meter.

2.2.4 Data transformation, auto–scaling, and polynomial terms

Each variable was transformed to approximate normality to maximize the statistical performance of each predictive model (Table 2.2). Data were normalized (mean centered and auto–scaling) as PLS is sensitive to fluctuations in scale and variance among predictor variables. Diagnostic plots indicated potential curvilinear relationships after auto–scaling between predictor and response variables. Polynomial terms (2^{nd} and 3^{rd} orders) were calculated for each response variable after auto–scaling and included in the variable selection calculations (Schielzeth 2010).

Table 2.2: Measured plant traits included in the LR and optimized partial least squares regression models. Partial least squares regression component selection based on lowest root mean squared error from cross–validation (RMSECV) using 10–fold cross–validation. Plant measurement abbreviations: **30c** = circumference at height of 30cm; **bc** = basal circumference; **cd** = maximum canopy diameter; **fc** = frond count; **fl** = frond length (blade length + stipe length); **hc** = circumference at half plant height; **ln** = number of leaves; **sl** = stipe length; **th** = total plant height; **wph** = resting height of falling plate meter.

Species	**Model**	**Comp**	**RMSECV**	**Predictors**($mass \sim x_1 + x_2 + ... x_n$)
grey dogwood	PLS	3	37.4g	$sqrt(\text{mass}) \sim \text{bc} + \text{bc}^2 + \text{cd} + \text{cd}^2 + \text{hc} + sqrt(\text{ln}) + sqrt(\text{ln})^2 + \text{th}$
	LR	NA	55.7g	$sqrt(\text{mass}) \sim \text{th}$
interrupted fern	PLS	3	13.8g	$log(\text{mass}) \sim sqrt^{-1}(\sum(\text{sl})) + log(\text{fc}) + sqrt^{-1}(\bar{x}(\text{bl})) + log(\sum(\text{fl}))$
	LR	NA	16.8g	$log(\text{mass}) \sim log(\bar{x}(\text{fl}))$
sand dropseed	PLS	4	30.9g	$sqrt(\text{mass}) \sim \text{bc} + \text{bc}^2 + square(30c) + cube(\text{th}) + cube(\text{th})^2 + sqrt(\text{wph})$
	LR	NA	42.3g	$sqrt(\text{mass}) \sim cube(\text{th})$

2.2.5 Variable reduction and model averaging

Using the training dataset (n = 35), all possible combinations of transformed variables and associated (2^{nd} and 3^{rd} order polynomials were scored using Bayesian Information Criterion (BIC) model selection in the *MuMin* package (Bartoń 2013) in *R* (R-Core-Team 2013). *MuMin*'s dredge function was used to force the condition that polynomial regression coefficients must be evaluated in conjunction with 1^{st} order regression coefficients to ensure proper fitting of the model (Schielzeth 2010, Symonds and Pither 2012). Models with BIC

≤ 2 are considered to be equivalent. Therefore, BIC models ≤ 2 were averaged and variable importance values for the predictor variables were extracted. The plant measurements selected by BIC model selection represent the optimized variables that will best predict plant mass in the dataset (Johnson and Omland 2004, Grueber et al. 2011).

2.2.6 Partial least squares regression and linear regression models

PLS regression models were created with the *pls* package (Mevik and Wehrens 2007) in *R*. The optimized predictor variables and transformations used to calculate each plant's PLS model are given in Table 2.2. Simple linear regression (LR) models were calculated with the plant height predictor variable to establish a standard curve in *R*'s *lm* function. Predicted biomass from allometric equations were not compared to PLS and LR models because published equations did not exist for the three measured plant species in this study.

The number of orthogonal components (i.e. latent variables) extracted from the PLS models was determined to evaluate each component's contribution to overall predictive fit. The number of components to retain in each model was determined using the root mean squared error of cross–validation (RMSECV) calculated from 10–fold cross–validation. RMSECV is a diagnostic metric used to test each component's contribution to the overall predictive fit of the model. The latent variable with the lowest average RMSECV indicates the number of components to retain in the PLS model, thus maximizing each model's predictive performance.

After generating standard curves from the PLS and LR models, plant mass was predicted and back transformed (P_{mass}) for the training and test datasets. P_{mass} was subtracted from the corresponding reference plant mass (R_{mass}) weighed in the laboratory to determine how well the model predicted each data point. A perfect model prediction for a sample is equal to zero (i.e. P_{mass} - $R_{mass} = 0$). Root mean squared error (RMSE) and R–squared estimates for the linear relationship between P_{mass} versus R_{mass} were calculated to determine the training and test dataset's actual predictive performance. The regression slope used to calculate RMSE and R–squared for the P_{mass} versus R_{mass} training dataset is equal to a slope = 1 with an intercept = 0. Mean and standard deviations of P_{mass} - R_{mass} were calculated for the training and test datasets.

2.3 Results

2.3.1 Variable selection in partial least squares regression models

BIC model selection reduced the number of variables from the full model in the interrupted fern (5 → 4) and sand dropseed (5 → 4) datasets. Comparatively, grey dogwood models included all five field measured variables in the optimized PLS models (Table 2.2). Field measured variables retained in the PLS models for all three species accounted for aspects of plant diameter, height, and structural counts. The most descriptive variables, scaled from 0.00 to 1.00, were determined by relative variable importance measures. The following are variables with high influence in each plant's PLS model: grey dogwood (basal circumference (1.00), canopy diameter (1.00), plant height (1.00), leaf number (1.00), circumference at half height (0.95)), interrupted fern (Σ frond length (0.72), Σ stipe length (0.59), Σ blade length (0.46), Σ frond count (0.43)), and sand dropseed (circumference at 30 cm (1.00), weighted plate height (1.00), basal circumference (0.98), plant height (0.90)). Weighted plate measurements were only relevant when predicting the biomass of sand dropseed. PLS models used for grey dogwood and sand dropseed corrected for curvilinear relationships between plant biomass and several field measured variables (i.e. 2^{nd} order polynomial terms) (Table 2.2).

2.3.2 Comparing models for predicting plant biomass in the training dataset

LR models using plant height predicted plant biomass well in the field but optimized PLS models consistently performed better in prediction diagnostics for all three plant species (Tables 2.2 & 2.3). PLS RMSECV calculated in each training dataset was 37.4 g (grey dogwood), 13.8 g (interrupted fern), and 30.9 g (sand dropseed). Comparatively, training dataset RMSECV predication accuracy in LR models was reduced for all three plant species using plant height as the predictor variable (55.7 g (grey dogwood), 16.8 g (interrupted fern), and 42.3 g (sand dropseed)).

The RMSECV model performance indicators translated to higher prediction accuracy when evaluating P_{mass} versus R_{mass} in each training dataset. All optimized PLS models had R-squared values ranging between 0.985 – 0.995. Predicted P_{mass} versus R_{mass} regression diagnostics were more variable in LR models with R-squared values ranging from 0.784 – 0.945. In all LR models, the lower R-squared values are a result of reduced model performance when predicting heavier plants in the population (Figure 2.2). Therefore, linear regression models introduced higher variability when predicting heavier plants. In comparison, PLS models accurately predicted training dataset plant mass across all plant weights resulting in higher R-squared values (Figure 2.2).

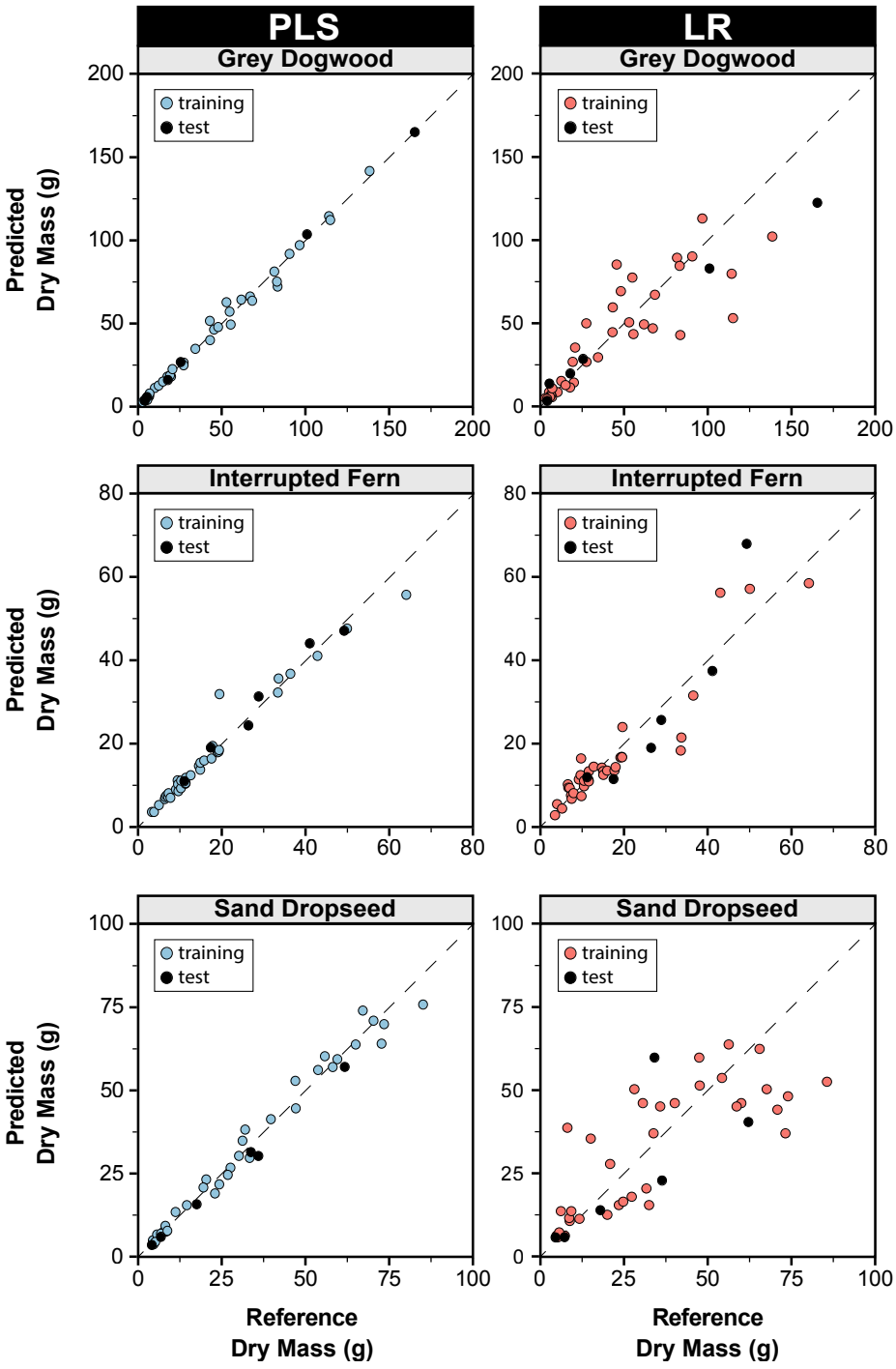


Figure 2.2: Graphs of predicted (P_{mass}) vs. reference (R_{mass}) plant biomass using the optimized PLS models and the LR model for the three plant species. The blue(PLS) and red(LR) points represent internally predicted data used to train each model (n = 35). Black points represent external data predictions from the test dataset using only predictor variables (n = 6). Each dashed line indicates a perfect prediction ($P_{mass} = R_{mass}$) with a slope = 1 and intercept = 0.

Table 2.3: Summary statistics for PLS and LR model training datasets. R-squared (R^2) and root mean squared error (RMSE) values are based on P_{mass} versus R_{mass} estimates where slope = 1 and intercept = 0.

Species	**mass ± 1 SD**	**height ± 1 SD**	**Model**	**RMSE**	**R^2**
grey dogwood	42.9g ± 41.3g	77.7cm ± 32.9cm	PLS	4.6g	0.995
			LR	25.2g	0.872
interrupted fern	17.1g ± 12.9g	81.0cm ± 18.1cm	PLS	2.8g	0.986
			LR	5.1g	0.945
sand dropseed	32.9g ± 23.6g	88.3cm ± 19.0cm	PLS	3.4g	0.994
			LR	14.2g	0.794

Model prediction performance in the training dataset, best determined by RMSE values, indicated that all PLS models consistently outperformed LR models when comparing P_{mass} - R_{mass} data. LR model RMSE increased as average plant mass increased (Table 2.3). This indicates that LR prediction accuracy is reduced when measuring plants with higher biomass in the field thus introducing higher prediction variability (RMSE: 25.2 g grey dogwood, 5.1 g interrupted fern, 14.2 g sand dropseed). In comparison, PLS models prediction accuracy was consistent across all plant growth forms thus reducing variability when predicting plant biomass (RMSE: 4.6 g grey dogwood, 2.8 g interrupted fern, 3.4 g sand dropseed).

2.3.3 Comparing models for predicting plant biomass in the test dataset

Similar to models using the training set data, external data points predicted by PLS models consistently outperformed LR models. Higher variability in LR models was shown in all diagnostic tests compared to the PLS models. R-squared of P_{mass} versus R_{mass} PLS test data points ranged between (0.995 – 0.995) while LR R-squared data ranged between (0.755 – 0.943). RMSE of PLS model P_{mass} versus R_{mass} was consistently lower when using LR models to predict species biomass. Average P_{mass} - R_{mass} ± 1 SD for externally predicted data using PLS was highest in grey dogwood (4.4 g ± 5.8 g) and lowest in interrupted fern (0.4 g ± 2.2 g) (Table 2.4). In comparison, average P_{mass} - R_{mass} ± 1 SD using LR had reduced predictive model performance with highest prediction variability in grey dogwood (−11.1 g ± 21.6 g) and lowest variability in interrupted fern (−1.5 g ± 9.6 g). Using the external dataset, PLS models had superior prediction performance with lower variability in all diagnostic tests.

Table 2.4: Summary statistics for PLS and LR model externally predicted data. R-squared (R^2) and root mean squared error (RMSE) values are based on P_{mass} verses R_{mass} estimates where slope = 1 and intercept = 0.

Species	**Model**	$\bar{x}$**(P_{mass}-R_{mass}) ± 1 SD**	**RMSE**	$\mathbf{R^2}$
grey dogwood	PLS	4.4g ± 5.8g	5.7g	0.995
	LR	-11.1g ± 21.6g	15.5g	0.943
interrupted fern	PLS	0.4g ± 2.2g	2.8g	0.995
	LR	-1.5g ± 9.6g	12.5g	0.926
sand dropseed	PLS	-2.5g ± 2.3g	3.3g	0.996
	LR	-2.5g ± 19.0g	17.4g	0.755

2.4 Discussion

Partial least squares regression was a superior predictive methodology compared to simple linear regression in the three plant species selected in this study. PLS predicted plant biomass had high accuracy and precision in datasets across distinct plant growth forms. This indicates a distinct advantage of using a multivariate approach to predict plant biomass in the field since growth form did not strongly influence the predictive performance of PLS.

2.4.1 Variable selection

Response variable selection is the most crucial step when creating a predictive standard curve to estimate plant mass. A statistical model is only as good as the response variables included in the analyses. For example, if all input variables are weakly correlated to plant mass, the best models chosen by model selection techniques will still result in poorly performing models in the field. Therefore, model selection techniques are used to identify the best model selected from a complete set of variable combinations based on statistical support (Johnson and Omland 2004). BIC model selection optimizes the most descriptive combination of variables that fit the data and improves predictive performance to build a representative model. BIC model selection is more statistically conservative than Akaike Information Criterion (AIC) calculations (Burnham and Anderson 2004). In the case of this study, this statistical property is advantageous as BIC model selection will identify fewer plant response variables to input into PLS analyzes, ultimately reducing field measurement

time. In my methodology, I employed BIC model selection due to its more conservative approach when selecting PLS variables.

In my study, the variables selected in grey dogwood and sand dropseed were strongly correlated to plant mass as shown by variable importance indicators. On the other hand, variables to predict interrupted fern biomass were not as strong. This may be an artifact of using the sum and average frond measurements in the field. This approach may have masked information in the individual frond measurements and reduced the predictive performance of each variable. Despite the weak correlation of the individual variables to plant mass in the interrupted fern population, the multivariate PLS approach still resulted in an accurate model with good predictive performance and highlights the usefulness of the technique in the field.

Plant height was reasonably correlated with plant biomass in the LR models. Thus, it has been used as an easily measured surrogate for plant biomass in the field (Singh et al. 1975, Catchpole and Wheeler 1992). In the field, simple linear regression models using one measurement variable have been shown to have low to moderate statistical accuracy depending on ecosystem type and vegetative structure (Catchpole and Wheeler 1992). To address reduced sampling accuracy, Bonham (1982) noted the need for the development of a multivariate sampling method to predict plant biomass. Models were developed to estimate sampling cost for increased variable collection but no predictive model was proposed. Thus, variable selection methods in this paper did not address the statistical assumption violations associated with collinear predictors in multiple regression. To account for this, the multivariate PLS approach in my study shows the importance of incorporating several predictor variables to estimate plant mass for these three species while accounting for variable collinearity. Incorporating variables that estimated plant density, circumference, and structural counts highlighted the morphological variation in the field as it pertains to correlating plant biomass.

An approach that employs only one predictor variable largely ignores the fact that two plants with identical heights may have distinctly different plant volumes in the field. This ultimately leads to higher variance in prediction accuracy and less reliable results. This effect is shown in my study when LR models of P_{mass} versus R_{mass} become increasingly unreliable as R_{mass} increases (Figure 2.2). Predicting the biomass of larger individuals with LR reduces measurement accuracy and introduces uncertainty into the model. Increased variability in prediction performance is indicated by lower R–squared and higher RMSE estimates in the training (Table 2.3) and external (Table 2.4) datasets compared to PLS.

The best way to avoid selecting weakly correlated response variables is to measure a large number of estimators in the field and iteratively reduce the number of predictor variables post–hoc using model selection techniques (Johnson and Omland 2004). Creative

measurements that are non–traditional in ecology (i.e. weighted plate measurements, plant circumference at a height of 30 cm) may yield surprisingly strong correlations when estimating the plant biomass of a target plant (Rayburn and Rayburn 1998, Rayburn and Lozier 2003). Once a suite of measurements have been collected, the number of variables in full models should be reduced to obtain a more parsimonious prediction model, remove irrelevant variables, and maximize measurement efficiency (Andersen and Bro 2010). This ultimately leads to highly accurate prediction models balanced against the cost of labor in the field.

2.4.2 The statistical advantage of using partial least squares regression when prediction plant biomass

In comparison to LR, PLS improves the precision and accuracy of estimating biomass across the herbaceous and shrub phenotypes by incorporating a multivariate estimation approach (Figure 2.2). PLS models have no statistical restrictions when variables exhibit multicollinearity (Wold et al. 2001), thus allowing for the incorporation of all variables that adequately describe aboveground plant architecture and morphological variation. This feature of PLS results in exemplary predictive performance in the field compared to all tested LR models. For example, the optimized PLS model for grey dogwood includes five collinear measurements (basal circumference, canopy diameter, circumference at half plant height, leaf number, and plant height). These response variables would violate the assumptions of traditional multiple regression methods even though the several estimators would be advantageous when predicting plant mass (Graham 2003). As shown in all PLS models, the multivariate approach led to more robust and accurate statistical models using both the training and test datasets compared to LR.

Variance in externally and internally predicted data tended to increase with higher average plant mass in the field for the three plant species (Figure 2.2). Larger prediction error can reasonably be expected due to the higher variability associated with plant growth rate response during competition for water, light and nutrients (Poorter and Nagel 2000). Hence, I show that samples predicted in the grey dogwood model (mass: 42.9 g $\pm$ 41.3 g) has the highest RMSE values while interrupted fern models (17.1 g $\pm$ 12.9 g) had the lowest RMSE in the regression diagnostics for PLS and LR. Thus, higher variation would be expected in plant populations with a larger range of plant mass in the field. But in all cases, LR was more sensitive to plants with higher average biomass and standard deviations resulting in less robust biomass prediction models using plant height as the sole predictor variable. PLS models, on the other hand, exhibited high predictive performance across all plants, regardless of plant biomass ranges in the field. The robustness of the multivariate

approach largely accounts for more morphological variability thus increasing the reliability of the models.

LR models were also less reliable when measuring larger plants within each plant population. LR models exhibited higher variation in biomass predictive performance near the upper end of plant mass in all species. This effect was most pronounced when evaluating predicted mass versus reference mass in the sand dropseed population, but was present in all evaluated plant species (Figure 2.2). This means that the plant height response variable in the LR models is less descriptive when predicting plants with larger biomass in the field, most likely due to increased morphological variation at larger sizes in the plant population due to variability in field response (Poorter and Nagel 2000). The use of a single response variable in all species highlights the need for multivariate measurements since prediction accuracy is not uniformly reliable across the entire plant population (Gholz et al. 1979).

Comparatively, all PLS models performed equally well and had relatively uniform prediction accuracy regardless of mass in all species (Figure 2.2). In all PLS models, predicted mass linearly increases near the perfect prediction slope indicating excellent predictive performance in the training and test datasets. RMSE diagnostics show that 66% of the predicted data in the externally predicted test dataset will fall within 2.8 g (interrupted fern), 3.3 g (sand dropseed), 5.7 g (grey dogwood) of the reference mass compared to the larger RMSE prediction errors associated with LR models. Therefore, I show that more uncertainty is introduced when predicting plant biomass using only one response variable. Thus, PLS is shown to be a more robust statistical technique that increases prediction confidence in experimental scenarios.

2.4.3 Practical applications of partial least squares regression

This paper follows a workflow (Figure 2.1) that integrates common statistical techniques (i.e. BIC model selection, data–transformations) with PLS. The intensity of labor cost and time necessary to create a PLS model in an experimental setting was similar to collecting a single response variable using LR. As shown by (Bonham 1982), optimizing the allocation of response variables for model input will maximize labor efficiencies and reduce data collection costs. Thus, implementing PLS into data collection schemes will increase plant prediction accuracy without introducing significantly higher opportunity costs such as increased sampling time. This approach is easily adapted to a variety of field and greenhouse situations, thus increasing sample replication, work efficiency, and prediction accuracy.

When creating a PLS model, choose a suite of morphological traits that are measured quickly and accurately under greenhouse or field conditions. As PLS evaluates categorical and continuous variables, morphological measurements should be tailored to the plant(s) of interest. This approach incorporates the flexibility to choose the number of variables

to be used in the predictive model. BIC model averaging extracts predictor variables with the highest parsimony. If a large suite of variables was measured, a complex model with 5 or more measurement variables could potentially be extracted. In this case, sampling volume during data collection may be a higher priority when weighed against final prediction accuracy. As BIC models ≤ 2 are considered equivalent, each equivalent model should be calculated using the PLS algorithm. RMSECV results using 10–fold cross–validation can be subsequently compared to determine the predictive capabilities within each model. PLS models should be optimized to create the best model for external predicting data while also considering sampling efficiency.

Compared to published allometric equations, PLS is customizable to a researcher's study system or greenhouse experiment. Allometric equations for predicting biomass have been shown to differ as a function of morphologic features and environmental conditions (Niklas and Enquist 2002). Thus, reliance upon published equations is not necessary for herbaceous plants and shrubs when utilizing PLS. Unlike published allometric equations, a drawback of this technique is the creation of the biomass standard curve. The destruction of a small subset of plants is inevitably required for all non–destructive biomass prediction analyses. Destructive harvesting can be accounted for when designing an experiment. A researcher can adjust the experimental design of a project by increasing sample replication with the intent of destructive harvesting, creating a preliminary experiment under the same environmental conditions and harvesting its biomass, or choosing a representative population in the field similar to the population of interest.

Several statistical methodologies have been proposed to predict multivariate, collinear datasets. The main alternatives to PLS regression are principal component regression (PCR), ridge regression (RR), and artificial neural networks (ANN) (Hastie et al. 2001). Compared to PLS, PCR does not account for variance associated with response variables and resulting models tend to be less parsimonious with higher variability. Studies evaluating PLS performance compared to RR show similar (Frank and Friedman 1993) or marginally better (Yeniay and Goktas 2002) predictive performance using external datasets while both outperform PCR. In general, PLS models are more parsimonious, easier to interpret, and more user friendly than RR. Alternatively, advances in computational statistics and machine learning suggest that ANN will create better predictive models than all of the preceding linear regression techniques. Currently, ANN methodologies are not widely used in ecology. ANN computations have a steep learning curve as the underlying statistics do not use common statistical methods. Thus, ANN methods are less accessible and more difficult to implement compared to PLS regression techniques.

2.5 Summary

In this chapter, I describe a double sampling method for accurately estimating individual herbaceous plant and small shrub biomass in the field. Partial least squares regression is a robust statistical technique that should be employed to accurately predict plant biomass in ecological experiments. In comparison to liner regression using a sole predictor variable, partial least squares regression increases prediction confidence and reliability in ecological experiments.

This chapter is intended to be a simple, customizable guide for ecologists and land managers. My approach maximizes aboveground biomass prediction accuracy with high measurement efficiency using simple statistical methods and inexpensive tools in the field. The customizable nature of this technique makes PLS a powerful statistical tool for researchers in ecological and environmental science.

Chapter 3

The Restoration of Grassland Vegetation in Post–Extraction Sandpits

3.1 Background

Ontario's sand plain prairies support a high biodiversity of regionally unique plants, insects, and animals (Gartshore et al. 1987). Surveys indicate that approximately 22% of Ontario's rare plant species are found in these prairie ecosystems (Ontario-Biodiversity-Council 2010). Many of these species have been elevated to endangered status due to habitat loss from land–use change, invasive species colonization, and fire suppression. It is estimated that Ontario's prairies occupy less than three percent of their original coverage (Rodger 1998). Increasing patch size on marginal lands through prairie restoration will facilitate the survival of sensitive habitat in addition to supporting species at risk. Excavated sandpits are candidate areas to restore prairie plant species but edaphic conditions limit the spontaneous development of high diversity plant communities (Wali 1999, Prach and Hobbs 2008). If post–mine substrate is left unassisted, plant communities can take decades to recover, if ever (Bradshaw 1997).

3.1.1 Biochar and compost as sandpit amendments

When added to soils, researchers suggest that biochar alters physiochemical soil properties by directly releasing nutrients or indirectly altering plant available nutrient concentrations (Chan and Xu 2009). Several studies, limited to agricultural systems, indicate that plant nutrient bioavailability of macro- (P,K) and micro–nutrients (Ca,Mg) have increased in response to charcoal application (Lehmann et al. 2003, Major et al. 2010, Rondon et al. 2007). Meta–analysis shows that biochar significantly translated to increased agricultural crop biomass and plant tissue macro–nutrients across all soil types and climates (Biederman and Harpole 2012). Biochar's largest positive influence on agricultural plant production has been shown in acidic, nutrient poor soils (Jeffery et al. 2011).

To date, there is little information on the effect of biochar on native plant growth in a restoration setting. Biochar amendments influenced grassland plant biomass inconsistently (Adams et al. 2013) and has the potential to cause shifts in species composition within managed grasslands (Schimmelpfennig et al. 2014). To date, no field study has investigated biochar as a soil conditioner when restoring grassland plants in degraded landscapes. The large–scale implication of biochar as a land management tool to grow native grassland plants still remains unexplored.

As a solitary soil amendment, compost has demonstrated ameliorative effects on soils in agricultural and mine restoration settings (Shiralipour et al. 1992, Giusquiani et al. 1995, Ouédraogo et al. 2001). Compost amendment increases organic matter content, water holding capacity, and soil nutrients, thus improving soil quality in degraded systems (Termorshuizen et al. 2004). Soil organic matter is a major component of soil quality because it directly or indirectly contributes to physical, chemical, and biological properties of functioning soils (Lal 2009). Compost strongly influences soil microbial communities by increasing in microbial biomass, respiration rates, and soil enzyme activity (Allievi et al. 1993). Microbial activity and soil fertility are generally related as compost is mineralized by microorganisms, thus releasing important elements (C, N, P and S) to the soil solution (Frankenberger and Dick 1983). Thus, increasing soil organic matter in soil is essential to restoring degraded landscapes by alleviating infertile conditions through the reestablishment decomposition cycles. As a land management tool, compost application to severely degraded landscapes increases grassland plant survivorship and primary production thus influencing restoration success (Hortenstine and Rothwell 1972, Norland and Veith 1995, Noyd et al. 1996).

3.1.2 Arbuscular mycorrhizal fungi as inoculum

In comparison to natural systems, post–mine areas have reduced arbuscular mycorrhizal diversity and abundance in addition to low nutrients and organic material (Stahl et al. 1988, Ganesan et al. 1991, Diaz and Honrubia 1994). This compounds the nutrient stress of these environments because some plants may be unable to establish and persist simply because they lack important microbial symbionts. Even if pre–mine area topsoil is stockpiled and retained, mining activities have been shown to degrade the efficacy of pre–mine populations of arbuscular mycorrhizas (Stahl et al. 1988). Thus, target plants can be inoculated with fungal propagules to facilitate plant production in disturbed mine areas (Bi et al. 2003, Taheri and Bever 2010).

Arbuscular mycorrhizal inoculum has been used in the restoration of mine areas for more than thirty years because of the ability to enhance plant establishment and survival. Plants inoculated with AM fungi in post–mine areas show positive growth responses (Khan 1981,

Johnson 1998, Enkhtuya et al. 2005, Rydlová et al. 2008). Several studies highlight the need to screen AM fungal isolates in order to determine their efficacy in the abiotically stressed edaphic conditions (Taheri and Bever 2010, Püschel et al. 2011). Not all commercial AM fungal isolates will be adapted to the harsh abiotic conditions present in post–mine areas.

3.1.3 Synergisms among biochar, compost, and arbuscular mycorrhizas

In degraded mine systems, the use of compost to restore plant communities is effective but its land management potential may be underestimated. Co–amending soils with biochar and compost may be synergistic as biochar's high cation exchange capacity and large surface area has the potential retain nutrients released from mineralized compost (Fischer and Glaser 2012). Initial studies show mixed results when co–amending soils with biochar and compost in terms of plant growth. Research on cultivar production ranged from a neutral (*Vitis vinifera* L., Schmidt et al. 2014) to positive (*Avena sativa* L., Schulz and Glaser 2012; *Samanea saman* F.Muell. and *Suregada multiflora* (A.Juss.) Baill., Ghosh et al. 2014) impact on the growth and quality of plants in soils co–amended with biochar compared to compost only treatments. To date, no field studies have investigated the impact of biochar co–amended with compost on native grassland plants.

Combining AM fungal inoculum with biochar and compost is anticipated to promote larger gains in plant community biomass compared to adding soil amendments alone. The application of organic amendments have a positive effect on the proliferation of natural AM fungi in agricultural systems (Harinikumar et al. 1990). AM fungi are able to exploit nutrients released by mineralization of organic matter due to the activities of soil microorganisms (Hodge et al. 2010). The combination of organic amendments in degraded systems and AM fungal inoculum has been shown to produce larger plants compared to either treatment alone when reclaiming desertified areas with shrubs (Caravaca et al. 2003b) and coal-mine spoil banks with biofuel cultivars (Püschel et al. 2011). Studies that have investigated the effect of biochar and AM fungal inoculation on plant biomass show inconclusive results. Both positive (Warnock et al. 2007) and negative (Birk et al. 2009, Warnock et al. 2010) effects on plant biomass are dependent upon pyrolysis temperature and quality of the biochar produced. Therefore, determining the optimized combination of compost, biochar, and AM fungal inoculum to increase plant response is essential when restoring grasslands in post–mine systems.

Considering the goal of grassland plant community restoration, the effect of AM fungal inoculation, municipal compost, and biochar has never been tested in degraded systems. In this multi–year study, two large–scale experiments using common grassland plants were sown in a recently excavated sandpit in southern Ontario, Canada. In the first experiment, a fully factorial combination of compost [CP], biochar [BC], and AM fungal inoculum were

applied to greenhouse grown plant plugs in a post–mine sandpit for two growing seasons. In a second experiment, a direct seeding experiment measured total plant cover along gradients of industrially feasible rates of compost and biochar with and without AM fungal inoculum. I hypothesized that AM fungal inoculation, compost, and biochar addition would individually increase plant dry mass and cover compared to non–amended controls. The rationale for this hypothesis was that additions were expected to alleviate water and nutrient stress in post–mine sandpit substrates. I further hypothesized that plots with the highest rates of compost + biochar and AM fungal inoculum would yield the highest plant dry mass and total cover. The rationale for this hypothesis was that the largest nutrient input, retention, and acquisition was expected through the fungal symbiosis. This will be evident by increased plant growth and total cover when compared to non–amended control plots. The overall aim of this study was to prescribe industrially feasible abiotic and biotic soil amendments to facilitate the long–term growth of a grassland plant community in post–mine sandpits while understanding the role of biochar, compost, and AM fungi in vegetative restoration.

3.2 Methods

3.2.1 Research site establishment

My research site was established on a recently active sand extraction area (0.5 hectares(ha)) near Port Rowan, Ontario, Canada (42 °40'17"N, 80 °28'46"W, elevation 211 m). The mine area, surrounded by Carolinian forest on three sides, is dominated by black oak (*Quercus velutina* Lam.), sassafras (*Sassafras albidum* (Nutt.) Nees.), and tulip tree (*Liriodendron tulipifera* L.) with interspersed exotic Scots pine (*Pinus sylvestris* L.). In 2010, the north side of the research site had a cover crop of soybean (*Glycine max* (L.) Merr.) followed by seeding with endemic grassland vegetation in 2011.

In the summer of 2010, the mine area was graded flat by an earthmover and a nine–wire fence was installed on the research site perimeter before the experimental plots were established to minimize deer browsing. The mine area substrate was poorly developed and composed of unconsolidated mineral substrate with no evidence of coarse soil organic material. The exposed sand substrate was easily eroded by wind which created a slight berm at the field site after one growing season. In 2011, relative height of each plot was measured where difference between the highest to lowest plots in the plant plug trial and seed application trial was 0.76 meters(m) and 1.02 m respectively.

3.2.2 Experimental design

I tested the effects of soil amendments (biochar, compost and AM fungal inoculation) on the establishment and growth of endemic grassland plants in a post-mining sandpit using two planting approaches: plant plugs, whereby plants were established in greenhouse and planted as plugs, and direct seeding at the site. Two different plant response measurement methods were used in the field: predicted plant biomass estimation in the plant plug trial, and vegetative cover estimation for the seed application trial. In the plant plug trial, a minimally destructive biomass estimation methodology was used because each plant plug location was known and could be tracked over multiple years. This methodology allows for the precise estimation of individual plant response to the experimental treatments over time. In the seed application trial, vegetative cover was estimated because tracking the growth of plant individuals was less feasible in the field. Therefore, vegetative cover was used as a non–destructive proxy for plant response, thus estimating plant germination and establishment rates in the field.

Biochar

I used biochar created from wood pellet feed stock that was pyrolyzed at 500 °C in an industrial scale non–oxygenated vacuum reactor. My biochar was supplied by the large-scale biochar producing facility, New Earth Renewable Energy Inc., based in Quebec, Canada. As biochar is a relatively unknown commodity as a soil amendment, I tested two industrially feasible biochar rates ($5\,\mathrm{T\,ha^{-1}}$ and $10\,\mathrm{T\,ha^{-1}}$) in the plant plug experiment. These rates balance amendment cost against the potential plant growth benefit of biochar relevant to the industrial–scale restoration of sandpit areas. In the direct seeding experiment, I tested six rates of biochar ranging from no biochar to rates at the upper end of cost feasibility in sandpit restoration ($0\,\mathrm{T\,ha^{-1}}$ to $40\,\mathrm{T\,ha^{-1}}$) (Table 3.1).

Compost

I used compost derived from municipal lawn and leaf urban waste streams distributed by Try Recycling in London, Ontario, Canada. I tested one industrially feasible compost rate ($20\,\mathrm{T\,ha^{-1}}$) in the plant plug experiment. As compost is relatively less expensive compared to biochar, compost can be applied at a higher rate when budgeting for an industrial–scale grassland restoration project. In the direct seeding experiment, I tested six rates of compost ranging from $0\,\mathrm{T\,ha^{-1}}$ to $40\,\mathrm{T\,ha^{-1}}$(Table 3.1).

Table 3.1: Experimental treatments for the seed application experiment. All treatment levels are fully factorial. Each treatment combination was applied to one plot only. Total number of plots was 72.

Biochar Level	**Compost Level**	**AM Level**
$0.0\,\mathrm{T\,ha^{-1}}$	$0.0\,\mathrm{T\,ha^{-1}}$	No inoculum
$2.5\,\mathrm{T\,ha^{-1}}$	$2.5\,\mathrm{T\,ha^{-1}}$	*Rhizophagus irregularis*
$5.0\,\mathrm{T\,ha^{-1}}$	$5.0\,\mathrm{T\,ha^{-1}}$	
$10.0\,\mathrm{T\,ha^{-1}}$	$10.0\,\mathrm{T\,ha^{-1}}$	
$20.0\,\mathrm{T\,ha^{-1}}$	$20.0\,\mathrm{T\,ha^{-1}}$	
$40.0\,\mathrm{T\,ha^{-1}}$	$40.0\,\mathrm{T\,ha^{-1}}$	
Factorial = biochar level × compost level × AM fungal inoculum level		

Arbuscular mycorrhizal fungal inoculum

I used a commercial arbuscular mycorrhizal fungal inoculant, *Rhizophagus irregularis* (Blaszk., Wubet, Renker & Buscot) C. Walker & A. Schüßler (2010), supplied by Mikro–Tek located in Timmins, Ontario, Canada. In the plant plug experiment, each plug container was inoculated with 20 spores contained in a proprietary powdered medium. The spore medium was added just below the soil surface of each plant plug container during seed sowing in April 2010. In the seed application experiment, each plot received 2 liters of water with suspended with spores. To mix the solution, a proprietary water–soluble powdered medium containing the spores was added to a watering can and applied evenly over the plot following seed compaction. Spores were applied at Mikro–Tek's recommended rate of $1000\,\mathrm{spores/m^2}$.

3.2.3 Plants used in restoration

The eight grassland plant species selected for this project met the following criteria: plant species that are common in Ontario prairies, tolerant of sandy soils and dry conditions, endemic to the study area, and known to associate with arbusuclar mycorrhizal fungi. Details about these plants are given in Table 3.2.

Table 3.2: The eight grassland plant species used in the plant plug trial and seed application trial. The abbreviation column indicates the plant code scheme associated with Figure 3.1. The final two columns indicate the abundance (i.e. number of plant plugs) of all species in each plot and the core sampling areas for the plant plug trial.

Species	**Common Name**	**Abbreviation**	$\frac{\textbf{plants}}{\textbf{plot}}$	$\frac{\textbf{plants}}{\textbf{core}}$
C_4 Grasses				
Andropogon gerardii Vitman	Big Bluestem	AG	11	5
Panicum virgatum L.	Switchgrass	PV	11	4
C_3 Grasses				
Elymus canadensis L.	Canada Wild Rye	EC	8	3
Bromus kalmii A. Gray	Prairie Brome	BK	8	4
N–Fixing Forbs				
Desmodium canadense L.	Showy Tick–trefoil	DC	11	5
Lespedeza capitata Michx.	Roundhead Bushclover	LZCA	3	3
Composite Forbs				
Liatris cylindracea Michx.	Ontario Blazing Star	LC	10	4
Symphyotrichum laeve(L.) Á. Löve & D. Löve var. laeve	Smooth Blue Aster	SL	10	5
		TOTAL	**72**	**33**

3.2.4 Plant plug trial

In the plant plug trial, plants were grown as plugs for 16 weeks in a commercial greenhouse by Pterophylla / St. Williams Nursery & Ecology Centre in St. Williams, Ontario, Canada from March 1^{st} to June 24^{th}. No supplemental light or heating was used in the greenhouse. Plant plugs were grown in 72 cell Landmark plug trays each filled with 57 cubic centimeters of a proprietary growing medium containing pine bark, sphagnum peat, leaf and yard waste compost and perlite. At the time of plug sowing, half of the plug containers were inoculated with AM fungal spores in the greenhouse. The growing medium used in the plugs was not sterilized to mimic industrial conditions. Background AM fungal communities were anticipated in non–inoculated plant plugs due to potential growing medium contamination in the industrial–scale greenhouse setting. The plant source material was collected from local plant populations by Pterophylla / St. Williams Nursery & Ecology Centre in the vicinity of the restoration project.

Plots (size: $10.2\,m^2$) were established by June 22^{nd}, 2010 using a fully–crossed, randomized factorial design, and monitored for 3 growing seasons (2010–2012). The two factors were: soil amendments (no amendment, $5\,T\,ha^{-1}$ biochar, $10\,T\,ha^{-1}$ biochar, $20\,T\,ha^{-1}$ compost, $5\,T\,ha^{-1}$ biochar + $20\,T\,ha^{-1}$ compost, $10\,T\,ha^{-1}$ biochar + $20\,T\,ha^{-1}$ compost) and *Rhizophagus irregularis* inoculation (±). Each of the 12 factorial combinations was replicated ten times for a total of 120 plots. Compost and biochar were raked into the upper 6 cm of substrate in May 2010. Control plots were not amended and were planted with non–inoculated plant plugs. A one meter buffer zone separated each plot to minimize plant interactions.

Native plant plugs were transplanted to the field between June 24^{th}, 2010 – July 1^{st}, 2010. Seventy–two plant plugs per plot were pre–mapped to have identical positions across all field plots (plug spacing = 33 cm)(Figure 3.1). A hexagonal plug arrangement was chosen to minimize spatial variability. Of the 8,640 plant plug positions, only two plugs were incorrectly planted as noted during vegetative censuses.

AM fungal quantification

AM fungal colonization of roots was quantified for greenhouse grown plant plugs (June 2010) and field plots (September 2011 / 2012) in the plant plug trial. For plugs, ten non–inoculated and ten inoculated plugs from each of the eight plant species were randomly selected in the greenhouse to assess root colonization before adding plugs in the field. In each field plots, sixteen soil cores per plot were collected and pooled near designated plug locations in September 2011 / September 2012 to minimize spatial variability. Plugs and the pooled field soil cores were washed free of soil in a 1 mm sieve to extract the roots.

Roots were removed, cut into 1 cm pieces, and preserved in 50% ethanol until microscopic analysis. Roots were stained with Chlorazol Black E (Brundrett et al. 1984) and then counted systematically under a microscope using the gridline intersect method (McGonigle et al. 1990).

Plant biomass estimation

I used partial least squares (PLS) regression to predict plant biomass using the same model creation methodology described in *Chapter 2*. A subset of randomly selected plots, one from each factor combination, was destructively harvested to create a PLS standard curve for six of eight plant species between September 14^{th}, 2011 and September 16^{th}, 2011. Similarly, a second set of plots was also destructively harvested between August 28^{th}, 2012 and August 31^{th}, 2012 as plants grew larger and morphological predictor characteristics were anticipated to change from first to second growing season. In both years, the C_3 grasses (see above) were not estimated; living plant tissue was not available in September due to early season senescence and poor plant performance, resulting in unreliable partial least squares regression estimates in the field. Since aboveground biomass harvesting may have introduced a plant growth bias in subsequent growing seasons, plots harvested in the first year were excluded from the final analyses. Plots destructively harvested in the second growing season were included in final analyses as they were not disturbed prior to harvesting.

Biomass was estimated for plants in the center of each plot (i.e. the "core area"(Figure 3.1)) in September 2011 and 2012. Core area plants were measured to reduce any confounding edge effects present in the field plots. Thirty–three (33) plant plugs in the core area were measured for each plot for a total of 3,960 plug locations measured per growing season.

I measured morphological plant characters related to height, diameter, and stem counts when appropriate for each plant species for the Fall 2011 and Fall 2012 growing seasons (Table 3.3). The predictor variables were selected via BIC model selection then used to measure the remainder of the 3,960 plant plugs in the field each season (Table 3.4). Partial least squares predicted mass was subtracted from the corresponding reference plant mass (P_{mass}–R_{mass}) $\pm$ 1 standard deviation (SD) to estimate prediction error. A value of zero indicates $P_{mass} = R_{mass}$, hence a perfect prediction. Statistical details of measurement accuracy for each species are given in Table 3.5.

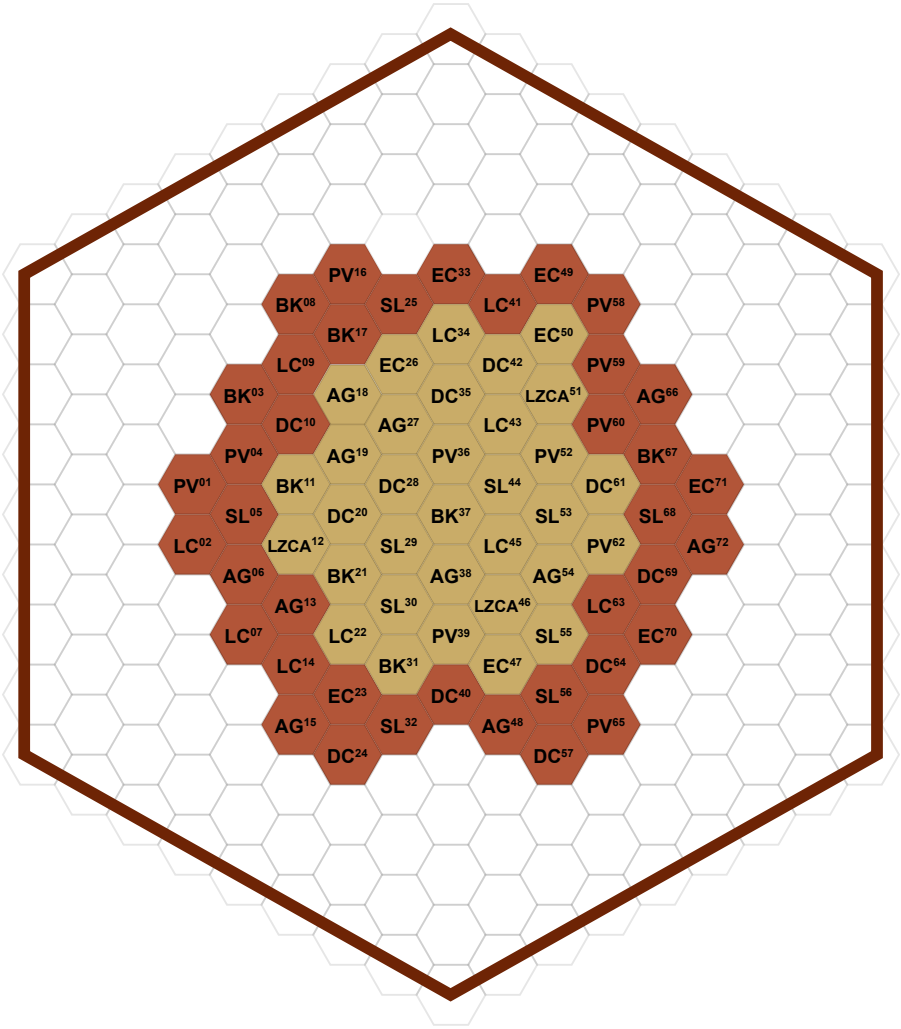


Figure 3.1: Diagram of the plant plug layout with plant positioning. Each hexagonal cell signifies the location and identity of one plant taxa added to the plot as a plant plug. All plots have the same plug configuration to minimize spatial variability. Plug spacing = 33 cm. Plants sampled in the core are indicated in beige. See Table 3.2 for plug abbreviations.

Table 3.3: Morphological characters measured in the field for the six plant species in September 2011 and September 2012. The most parsimonious combination of variables was selected via Bayesian Information Criterion model selection to create the predictive models.

Species	**Year**	**Measured Variables**)
C_4 Grasses		
Andropogon gerardii	2011	height; weigh plate; # of tillers; # of seed heads; basal circumference
	2012	height; weigh plate; # of tillers; # of seed heads; basal circumference; circumference @ 30cm
Panicum virgatum	2011	height; weigh plate; # of tillers; # of seed heads; basal circumference
	2012	height; weigh plate; # of tillers; # of seed heads; basal circumference; circumference @ 30cm
N–Fixing Forbs		
Desmodium canadense	2011	stem length; # of stems; length of stem inflorescence; basal circumference
	2012	stem length; # of stems; length of stem inflorescence; basal circumference; circumference @ 30cm
Lespedeza capitata	2011	stem length; # of stems; # of stems with inflorescence; inflorescence length; basal circumference
	2012	stem length; # of stems; # of stems with inflorescence; inflorescence length; basal circumference
Composite Forbs		
Liatris cylindracea	2011	# of leaves; height; # of inflorescence; length of stems with inflorescence
	2012	# of leaves; height; # of inflorescence; length of stems with inflorescence
Symphyotrichum laeve	2011	stem length; # of stems; # of stems with inflorescence; inflorescence length
	2012	stem length; # of stems; # of stems with inflorescence; inflorescence length

Table 3.4: Morphological characters selected for the six plant species measured in September 2011 and September 2012. The variables given in the table were selected via Bayesian Information Criterion model selection to create the predictive models using partial least square regression.

Species	**Year**	**Comp**	**Predictor Vars**($mass$ ~ x_1 + x_2 + ... x_n)
C_4 Grasses			
Andropogon gerardii	2011	2	(height) + (weigh plate)
	2012	3	(weigh plate) + (circumference @ 30cm) + (circumference @ 30cm)^2 + (# of seed heads) + (# of seed heads)^2
Panicum virgatum	2011	2	(basal circumference) + (weigh plate)
	2012	2	(circumference @ 30cm) + (weigh plate) + (weigh plate)^2 + (# of seed heads)
N–Fixing Forbs			
Desmodium canadense	2011	2	(mean stem length) + (sum stem length) + (sum stem length)^2
	2012	1	(basal circumference) + (circumference @ 30cm)
Lespedeza capitata	2011	2	(sum stem length)+ (mean stem length)
	2012	2	(sum stem length) + (sum stem length)^2
Composite Forbs			
Liatris cylindracea	2011	2	(# of leaves) + (# of leaves)^2 + (# of leaves)^3 + (height) + (height)^2
	2012	1	(# of leaves) + (height)
Symphyotrichum laeve	2011	3	(sum stem length) + (mean stem length) + (mean stem length)^2
	2012	4	(mean inflorescence length) + (mean inflorescence length)^2 + (sum stem length) + (sum stem length)^2

Table 3.5: Partial least squares regression diagnostics for the six plant species measured in September 2011 and September 2012. All prediction data is based on variables selected via Bayesian Information Criterion model selection (Table 3.4). Mass data is given in grams (g) dry weight based on weighed plants used to create the standard curve. For each species, predicted plant mass from the partial least squares regression model was subtracted from the reference plant mass (P_{mass} – R_{mass}) ± 1 standard deviation (SD) to calculate within–model estimates. When $P_{mass} = R_{mass}$, predicted mass is equal to reference mass, thus represents a perfect prediction. R-squared, root mean squared error (RMSE), and p–values were calculated for P_{mass} – R_{mass} using linear regression for each plant species to indicate prediction accuracy. All regression diagnostics are based on a slope = 1 and intercept = 0.

Species	**Year**	**mass ± 1 SD**	**Rep**	**(P_{mass}–R_{mass}) ± 1 SD**	**RMSE**	**R^2**	**p–value**
C_4 Grasses							
Andropogon gerardii	2011	8.5g ± 5.1g	36	0.2g ± 2.4g	9.9g	0.944	<0.001
	2012	30.7g ± 22.9g	41	0.9g ± 5.1g	28.6g	0.983	<0.001
Panicum virgatum	2011	15.7g ± 11.1g	34	-0.7g ± 6.2g	17.6g	0.892	<0.001
	2012	94.1g ± 84.9g	41	-1.2g ± 21.5g	100.0g	0.970	<0.001
N–Fixing Forbs							
Desmodium canadense	2011	33.5g ± 20.1g	36	-0.4g ± 7.6g	36.1g	0.962	<0.001
	2012	29.6g ± 21.2g	41	-1.0g ± 12.9g	38.1g	0.873	<0.001
Lespedeza capitata	2011	2.7g ± 2.2g	25	-0.1g ± 1.0g	2.9g	0.923	<0.001
	2012	2.2g ± 3.3g	25	-0.1g ± 1.1g	1.9g	0.924	<0.001
Composite Forbs							
Liatris cylindracea	2011	2.5g ± 1.3g	35	0.0g ± 0.6g	3.1g	0.957	<0.001
	2012	0.9g ± 0.8g	41	0.9g ± 0.8g	1.0g	0.917	<0.001
Symphyotrichum laeve	2011	7.9g ± 4.1g	36	-0.3g ± 2.8g	10.2g	0.901	<0.001
	2012	7.9g ± 9.7g	41	-0.2g ± 1.8g	6.9g	0.981	<0.001

3.2.5 Seed application trial

The seed application experimental plots were set–up adjacent to the plug experiment using the same plot dimensions and soil amendment incorporation protocol. Plots were established in August of 2010 using a fully–crossed, randomized factorial design and monitored for 3 growing seasons (2011–2013). Factors were: six rates of each amendment given in Table 3.1 and *Rhizophagus irregularis* inoculation (±). Amendment and inoculation combinations were not replicated, for a total of seventy–two plots. To minimize overwinter seed mortality and undesired seed movement via wind scour, seeding and inoculation were not done until May 2011.

In April 2011, seeds were pre–weighed into bags, mixed with moist vermiculite, and stored at 4 °C in the refrigerator for one month (Table 3.6). This process of cold–moist stratification promotes rapid spring germination of dormant plant seeds. In May 2011, cold–moist stratified seeds were hand sown and lightly mixed with a steel rake into the sandpit substrate. A seed roller was used to press the seed into the sandpit floor to ensure soil–seed contact. Mycorrhizal inoculum was added to one set of the amendments via a liquid medium containing spores at the recommended rate. Seeds were applied at double the standard rate for recommended for tallgrass prairie restoration projects to ensure plant establishment in the experiment.

Table 3.6: Seeding rate in grams (g) for the eight grassland plant species used in the seed application trial. Plot size was 10.2 m^2. Seeds were cold–moist stratified at 4 °C for one month until the time of sowing in the field (May 2011).

Species	**Seeding Rate (g)**
Andropogon gerardii	7.0 g
Panicum virgatum	1.5 g
Elymus canadensis	5.5 g
Bromus kalmii	2.5 g
Desmodium canadense	1.5 g
Lespedeza capitata	1.5 g
Liatris cylindracea	4.0 g
Symphyotrichum laeve	6.0 g

Percent cover estimation

A photographic time–series technique was used to estimate the percent cover of plant growth in the seed application trial. This approach tracked the germination and establishment rates of the seeded grassland species across the treatment levels. An angle camera monopod was constructed to take overhead pictures in each plot (Figure 3.2). Photos were cropped to analyze a 2.6 m^2 area. Total plant cover was measured using the SamplePoint software (Booth and Cox 2008). In SamplePoint, a 100 point overlaying grid was used to classify pixels as grass, composite forb, N–fixing forb, soil, or plant litter. Native plant cover was estimated from the summation of grasses, forbs, and N–fixing forbs and subsequently divided by total pixels estimated. Photographs of each plot were taken in September for three growing seasons (2011–2013).

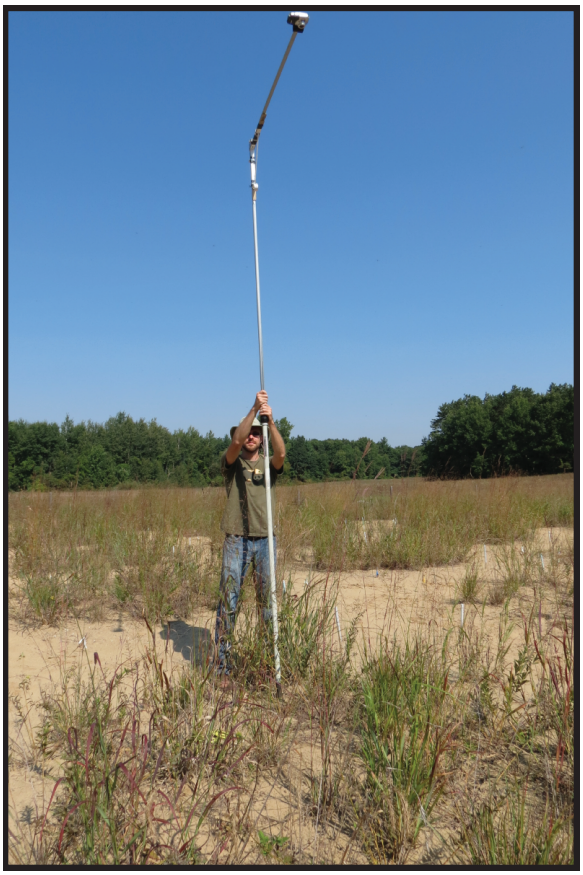


Figure 3.2: Collecting photographic data to analyze percent plant cover. A right–angled monopod was designed to take over–head photographs used to estimate plant cover in the seed application trial. The monopod was raised and leveled with the camera on a delayed setting to capture a picture for cover estimation in the SamplePoint software. (Photo Taken: September 2012)

3.2.6 Statistical analyses

Linear mixed effects models were used to test treatment significance for the plant response estimates and AM colonization of field roots. Linear mixed effects models use multivariate statistical procedures that account for random variability associated with plots at the field site. AM colonization of plant plug roots in inoculated versus non–inoculated treatments were evaluated using a t–test. Data transformations were employed when necessary to approximate a normal distribution of model residuals. Relative plot height was included as a covariate in all linear mixed effects models. Linear mixed effect model selection procedures iteratively removed non–significant variables using Chi–squared tests. This resulted in the most parsimonious models to analyze statistical significance for each response variable. Linear mixed effects models were analyzed using the *lme4* package in R (R-Core-Team 2013, Bates et al. 2014). Significance levels (p–values) derived from Markov Chain Monte Carlo methods, % explained deviance (an R–squared proxy, abbreviated: % Expl. Dev.), and main level post–hoc comparisons were calculated using the R package LMERConvenienceFunctions by Tremblay and Ransijn (2013).

3.3 Results

3.3.1 Plant plug trial

AM fungal establishment in greenhouse plug roots

All eight plant species were colonized by AM fungi in the greenhouse (Figure 3.3). The application of *R. irregularis* inoculum resulted in significant increases in percent colonization in all species compared to non–inoculated plants ($p<0.001$) (mean ranges of AM fungi in inoculated roots: 16.9% (*E. canadensis*) – 30.1% (*Andropogon gerardii*)). As expected in the unsterile greenhouse environment, low levels of AM fungal colonization of roots were detected in non–inoculated plant plugs across all plant species (<5.0% mean AM colonization of roots). These results indicate that the AM fungal inoculum treatment was established at the onset of the plant plug trial.

AM fungal establishment in field roots

R. irregularis inoculum persisted in the field after two growing seasons. Significant increases in AM colonization rates of field roots were detected in inoculated vs. non–inoculated plots ($p<0.001$)(Figure 3.4a). AM fungal inoculated treatments nearly doubled in the rate of root colonization between September 2011 and September 2012 (mean AM fungal colonization of roots: 22.4% (2011) to 45.8% (2012)). Mean AM fungal colonization

rates of roots in the non–inoculated plots tripled from 5.5% (2011) to 15.8% (2012). Roots in inoculated plots experienced a larger relative increase in AM fungal colonization between September 2011 and September 2012 compared to non–inoculated treatments ($p<0.001$).

Plant biomass responses in the plant plug trial

Increases in total plant biomass were influenced by soil amendments ($p=0.056$), growing season ($p<0.001$), and the relative plot height covariate ($p=0.068$) in the plant plug trial. No significant difference in total plant biomass was detected in plots receiving soil amendments compared to non–amended controls ($p>0.05$). No significant interactions among the model terms were detected.

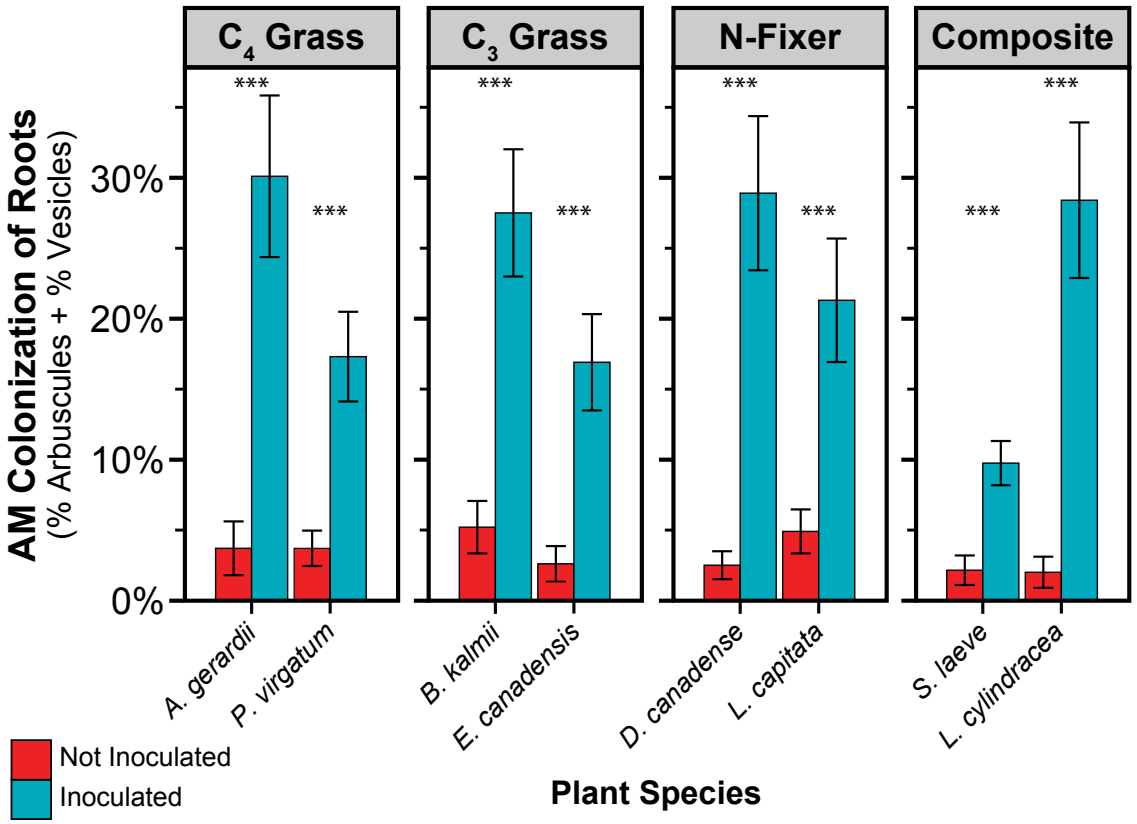
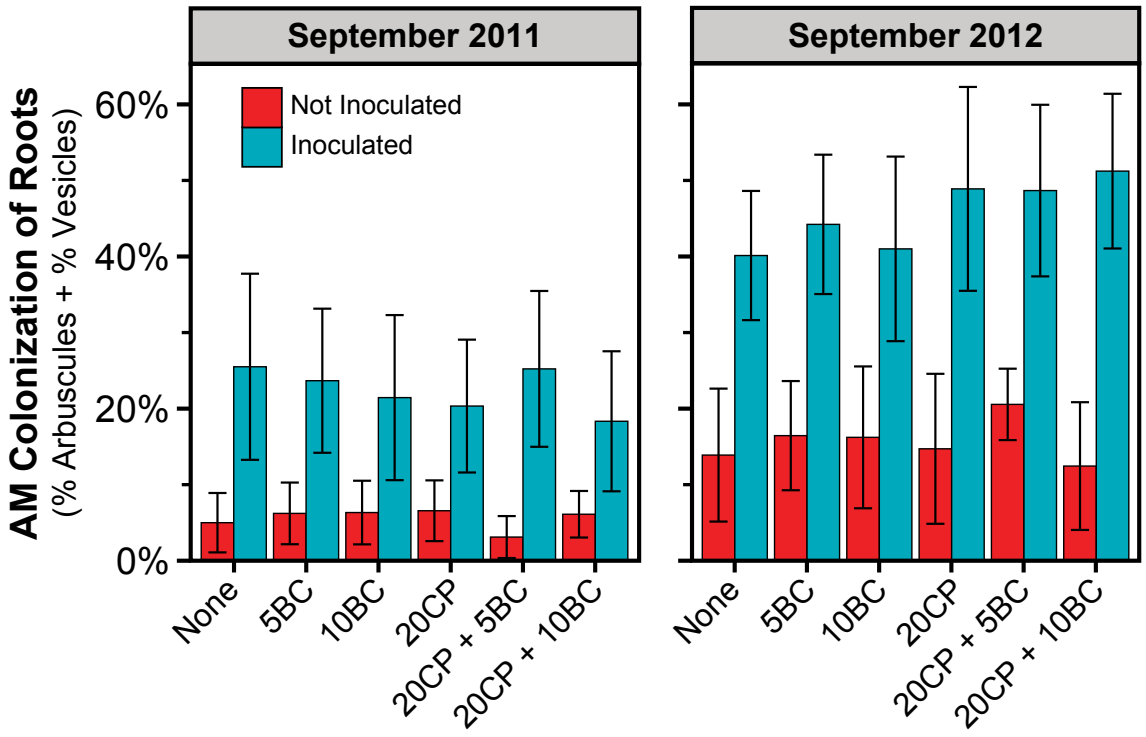


Figure 3.3: Percent AM fungal colonization of greenhouse grown plant plug roots. Plant plugs were randomly selected just prior to sowing plant plugs in the field (June 2010). Ten plant plugs from each treatment level (± *R. irregularis*) of all eight species were analyzed for AM colonization of roots using t-tests comparing inoculated and non–inoculated plants. Raw data ± 1 SD is presented in the graph. Each asterisk represents a p–value (***) $<$ 0.001 for comparisons between inoculation treatment levels. Replication = 10.

(a)



(b)

Model Terms	p–value	sig. level	% Expl. Dev.
AM Inoculation	<0.001	***	33.64%
Year	<0.001	***	24.19%
Plot Height (Dry → Wet)	0.003	(**)	0.94%
Interactions			
Year × Plot Height	0.056	.	0.39%
Year × AM Inoculation	<0.001	***	3.35%
Amendment × AM Inoculation × Year	0.035	*	1.27%

Significance: * ≤ <0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.4: AM fungal colonization of the mixed community of field roots in the plant plug trial. Panel (a) represents the graph of raw data with error bars (± 1 SD) based on the most parsimonious linear mixed effects model. Experimental treatment replication = 9. The left graph panel represents data after one growing season. Labels on the x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Statistical output (b) shows significant main effect terms and interactions. Main effects included in the model were: Amendment, AM inoculation, Plot Height, and Year. % explained deviance is abbreviated as % Expl. Dev. in the output. Model terms with negative regression slopes are indicated in parentheses.

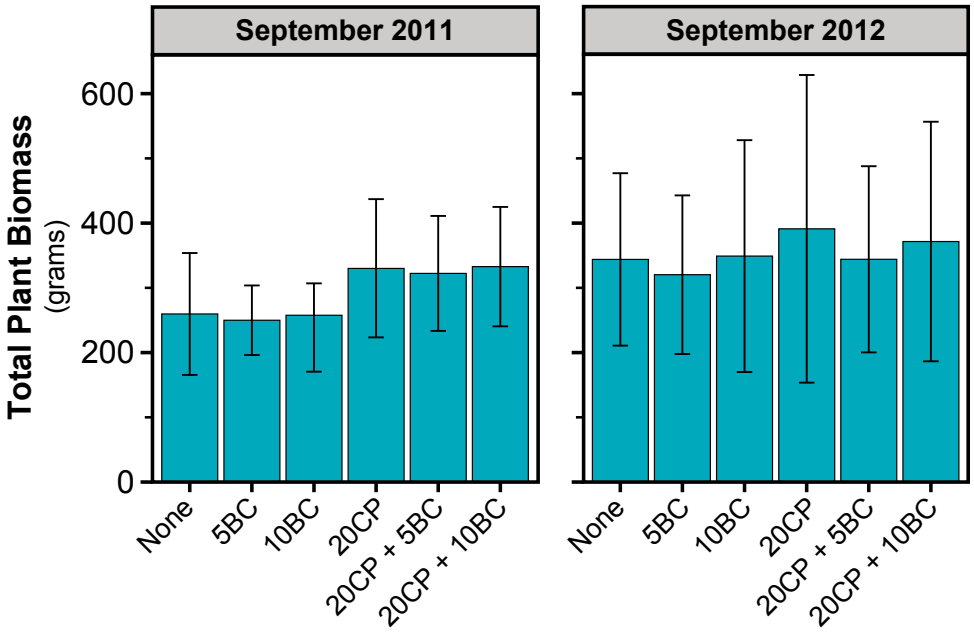
Biochar amendments compared to non–amended control plots, direct effects of biochar were not detected on total plant biomass ($p>0.05$)(Figure 3.5a). Unexpectedly, biochar only amendments significantly reduced total plant biomass compared to the majority of compost and compost + biochar amended plots (Figure 3.5a). Only *Andropogon gerardii* responded negatively to the addition of biochar compared to non–amended control plots (Figure 3.6a). All other measured plant species exhibited no direct response to either biochar rate.

Compost amendment The total plant biomass response to the compost only amendment was positive compared to non-amended controls although not statistically significant ($p=0.125$). *Andropogon gerardii* biomass was reduced in the presence of compost compared to non–amended control plots (Figure 3.6a). *Desmodium canadense* experienced significant increases in plant biomass in the presence of compost amendments. No other direct compost only effect were detected in plant response for the four other plant species in this trial.

AM fungal inoculation AM fungal inoculation did not significantly influence total plant biomass in the plant plug trial (Figure 3.5a) although each species varied in plant biomass when inoculated with *R. irregularis*. The biomass of *Panicum virgatum* ($p<0.001$) and *Lespedeza capitata* ($p=0.021$) responded positively to *R. irregularis* inoculation. In contrast, *Andropogon gerardii* and *Liatris cylindracea* biomass was significantly reduced in AM inoculated plots ($p<0.05$). No inoculation response was detected for *Symphyotrichum laeve* and *Desmodium canadense*. Altogether, interspecies variation in plant response to AM inoculation resulted in a neutral effect on the total biomass response in the community.

Synergistic effects of biochar, compost, and AM fungal inoculation Direct total plant biomass effects of $10\,\mathrm{T\,ha^{-1}}$ of biochar + $20\,\mathrm{T\,ha^{-1}}$ of compost were positive compared to non–amended control plots although not statistically significant ($p=0.117$).(Figure 3.5a). Co–amended plots with $10\,\mathrm{T\,ha^{-1}}$ of biochar + $20\,\mathrm{T\,ha^{-1}}$ of compost significantly increased total plant biomass compared to plots with $5\,\mathrm{T\,ha^{-1}}$ of biochar ($p=0.037$) and $10\,\mathrm{T\,ha^{-1}}$ of biochar ($p=0.008$) amended plots. The interaction of AM fungal inoculum and soil amendments did not significantly influence total plant biomass. No significant differences were detected when comparing total plant biomass in compost and compost + biochar.

(a)



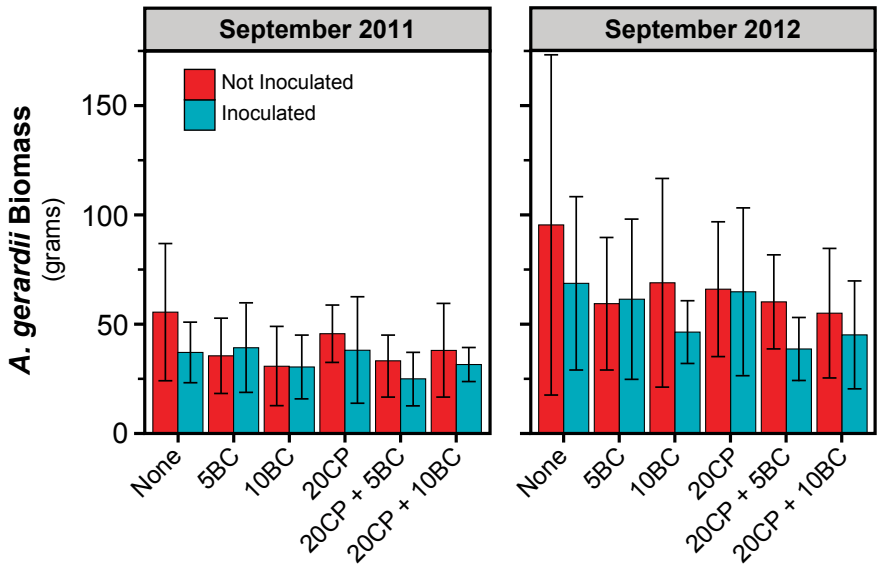
(b)

Model Terms	p–value	sig. level	% Expl. Dev.
Amendment	0.056	.	2.95%
Year	<0.001	***	3.99%
Plot Height (Dry → Wet)	0.068	.	0.90%
Sig. Post–Hoc Comparisons	**p–value**	**sig. level**	
5BC → 20CP	0.039	*	
5BC → 20CP + 10BC	0.037	*	
10BC → 20CP	0.008	**	
10BC → 20CP + 5BC	0.052	.	
10BC → 20CP + 10BC	0.008	**	

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.5: Predicted total plant biomass in the plant plug trial. Panel (a) represents the graph of raw data with error bars (± 1 SD) based on the most parsimonious linear mixed effects model. Experimental treatment replication = 9. The left graph panel represents data after one growing season. Labels on the x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Statistical output (b) shows significant main effect terms and interactions. Main effects included in the model were: Amendment, AM inoculation, Plot Height, and Year. % explained deviance is abbreviated as % Expl. Dev. in the output. Note that model terms with negative regression slopes are indicated in parentheses around the significance levels.

(a)



(b)

Model Terms	p–value	sig. level	% Expl. Dev.
Amendment	0.048	(*)	1.45%
Year	<0.001	***	17.61%
Plot Height (Dry → Wet)	0.061	.	0.45%
AM Inoculation	0.022	(*)	0.67%
Sig. Post–Hoc Comparisons	**p–value**	**sig. level**	
None → 5BC	0.094	(.)	
None → 10BC	0.012	(*)	
None → 20CP + 5BC	0.003	(**)	
None → 20CP + 10BC	0.027	(*)	
10BC → 20CP	0.069	.	
20CP → 20CP + 5BC	0.018	(*)	

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.6: Predicted *Andropogon gerardii* biomass in the plant plug trial. Panel (a) represents the graph of raw data with error bars (± 1 SD) based on the most parsimonious linear mixed effects model. Experimental treatment replication = 9. The left graph panel represents data after one growing season. Labels on the x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Statistical output (b) shows significant main effect terms and interactions. Main effects included in the model were: Amendment, AM inoculation, Plot Height, and Year. % explained deviance is abbreviated as % Expl. Dev. in the output. Note that model terms with negative regression slopes are indicated in parentheses around the significance levels.

Lespedeza capitata was positively influenced by the interaction among the soil amendments and AM fungal inoculum compared to non–amended control (Figure 3.7a). *Desmodium canadense* experienced biomass gains in the presence of all compost + biochar amended treatments compared to biochar only and non–amended plots (Figure 3.8a). *Desmodium canadense*'s large response to compost and compost + biochar amended plots strongly influenced total biomass results in September 2011 (Figure 3.5a). Comparatively, only *Andropogon gerardii* responded negatively to the compost + biochar treatments compared to non–amended control plots (Figure 3.6a). Overall, these results indicate that compost + biochar addition with AM fungal inoculation enhances the plant community response at the species level in post–mine sandpits.

Growing season Interspecies growth response was highly variable after two full growing seasons. In all models, growing season explained the highest amount of variation in the species biomass datasets. The biomass of N-fixing forbs (Figure 3.7a & Figure 3.8a) and composite forbs (Figure 3.10a & Figure 3.11a) was significantly reduced between September 2011 and September 2012. Comparatively, the C_4 grasses (*Andropogon gerardii* and *Panicum virgatum*) experienced biomass gains between September 2011 and September 2012. The C_4 grasses were amongst the largest contributors to total biomass, accounting for total biomass gains from September 2011 to September 2012.

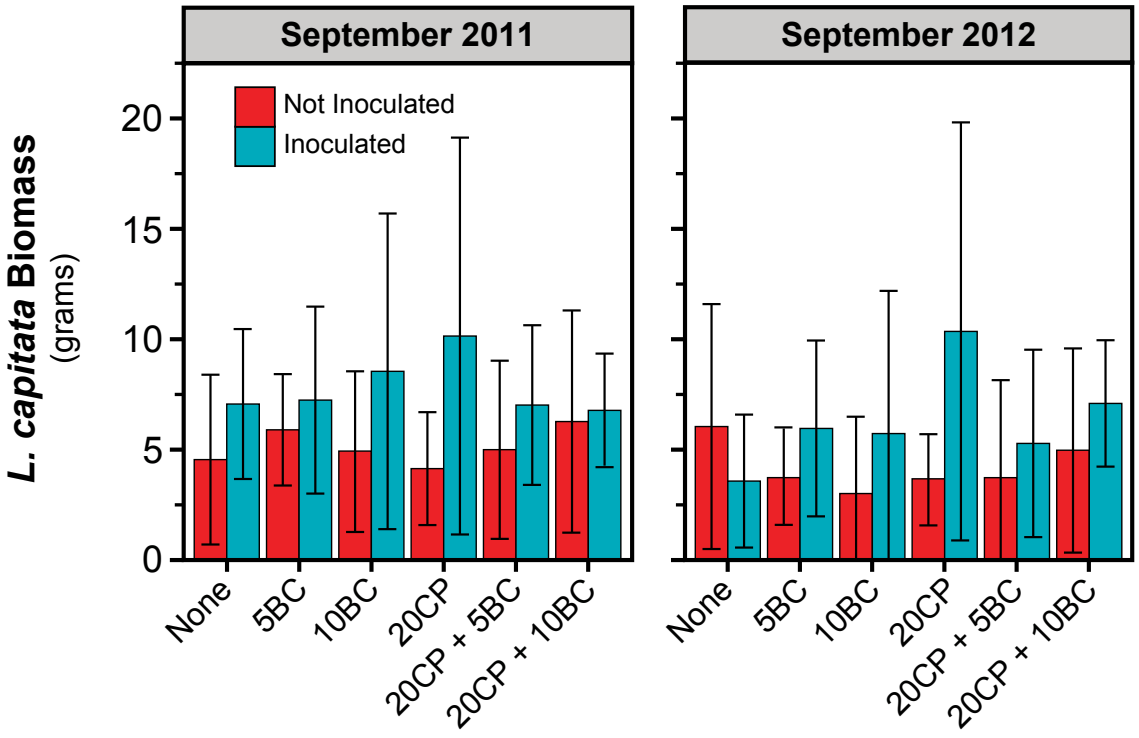
Plot height covariate A trend was detected in the response of total biomass to the plot height covariate ($p=0.068$)(Figure 3.5a). This indicates that plant biomass increased in plots lower on the landscape regardless of treatment. Similarly, the plot height covariate was significant and positive for all measured plant species except for *Desmodium canadense* ($p>0.05$) and *Panicum virgatum* ($p>0.05$). The influence of the plot height covariate in the total biomass analysis may have been suppressed since *Desmodium canadense* and *Panicum virgatum* are among the largest contributors to total plant biomass.

3.3.2 Seed application trial

Note: Two plots in the southeast corner of the seed application trial were removed from the analysis due to close proximity to the research site's water table. These outlier plots, $5\,\mathrm{T\,ha^{-1}}$ biochar + AM fungi (% cover in 2013: 53%) and $5\,\mathrm{T\,ha^{-1}}$ biochar - AM fungi (% cover in 2013: 61%), were nearest to a former wet depression at the field site and exhibited high vegetative density compared to all other plots. Mean total % cover ± 1 SD excluding outlier plots was 21% ± 9% in 2013.

The eight species sown at the mine site accounted for the vast majority of the vegetation in seed application trial. Non–seeded volunteer plant cover was negligible throughout

(a)



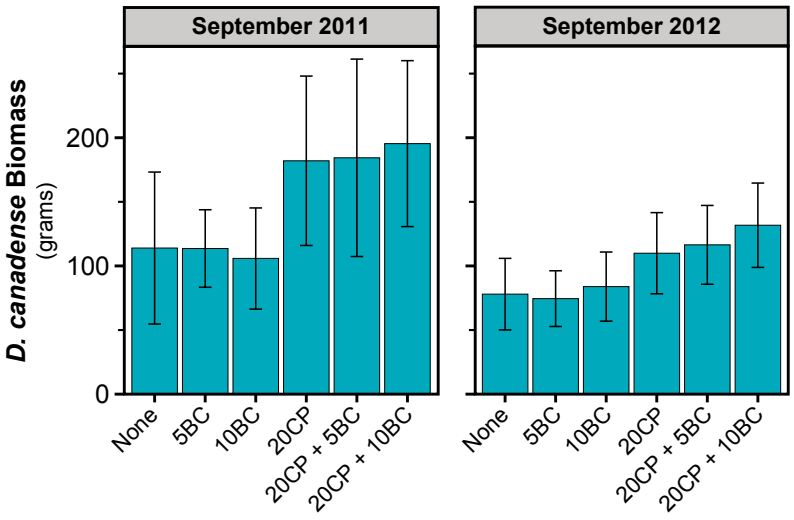
(b)

Model Terms	p–value	sig. level	% Expl. Dev.
Year	<0.001	(***)	4.00%
Plot Height (Dry → Wet)	0.035	*	0.50%
AM Inoculation	0.021	*	0.60%
Interactions			
Amendment × AM Inoculation × Year	0.014	*	1.61%

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.7: Predicted *Lespedeza capitata* biomass in the plant plug trial. Panel (a) represents the graph of raw data with error bars (± 1 SD) based on the most parsimonious linear mixed effects model. Experimental treatment replication = 9. The left graph panel represents data after one growing season. Labels on the x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Statistical output (b) shows significant main effect terms and interactions. Main effects included in the model were: Amendment, AM inoculation, Plot Height, and Year. % explained deviance is abbreviated as % Expl. Dev. in the output. Note that model terms with negative regression slopes are indicated in parentheses around the significance levels.

(a)



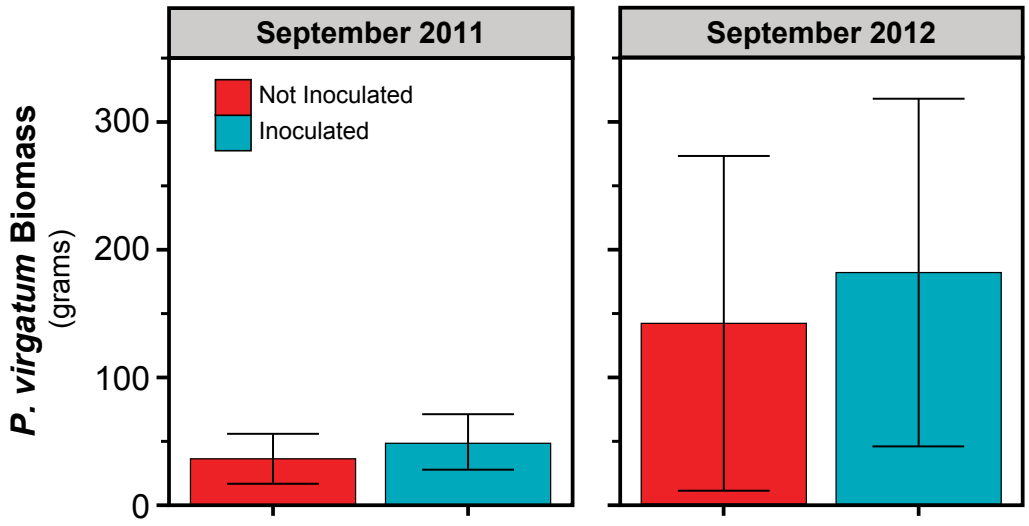
(b)

Model Terms	p–value	sig. level	% Expl. Dev.
Amendment	<0.001	***	9.27%
Year	<0.001	(***)	14.05%
Sig. Post–Hoc Comparisons	**p–value**	**sig. level**	
None → 20CP	<0.001	***	
None → 20CP + 5BC	<0.001	***	
None → 20CP + 10BC	<0.001	***	
5BC → 20CP	<0.001	***	
5BC → 20CP + 5BC	<0.001	***	
5BC → 20CP + 10BC	<0.001	***	
10BC → 20CP	<0.001	***	
10BC → 20CP + 5BC	<0.001	***	
10BC → 20CP + 10BC	<0.001	***	

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.8: Predicted *Desmodium canadense* biomass in the plant plug trial. Panel (a) represents the graph of raw data with error bars (± 1 SD) based on the most parsimonious linear mixed effects model. Experimental treatment replication = 9. The left graph panel represents data after one growing season. Labels on the x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Statistical output (b) shows significant main effect terms and interactions. Main effects included in the model were: Amendment, AM inoculation, Plot Height, and Year. % explained deviance is abbreviated as % Expl. Dev. in the output. Note that model terms with negative regression slopes are indicated in parentheses.

(a)



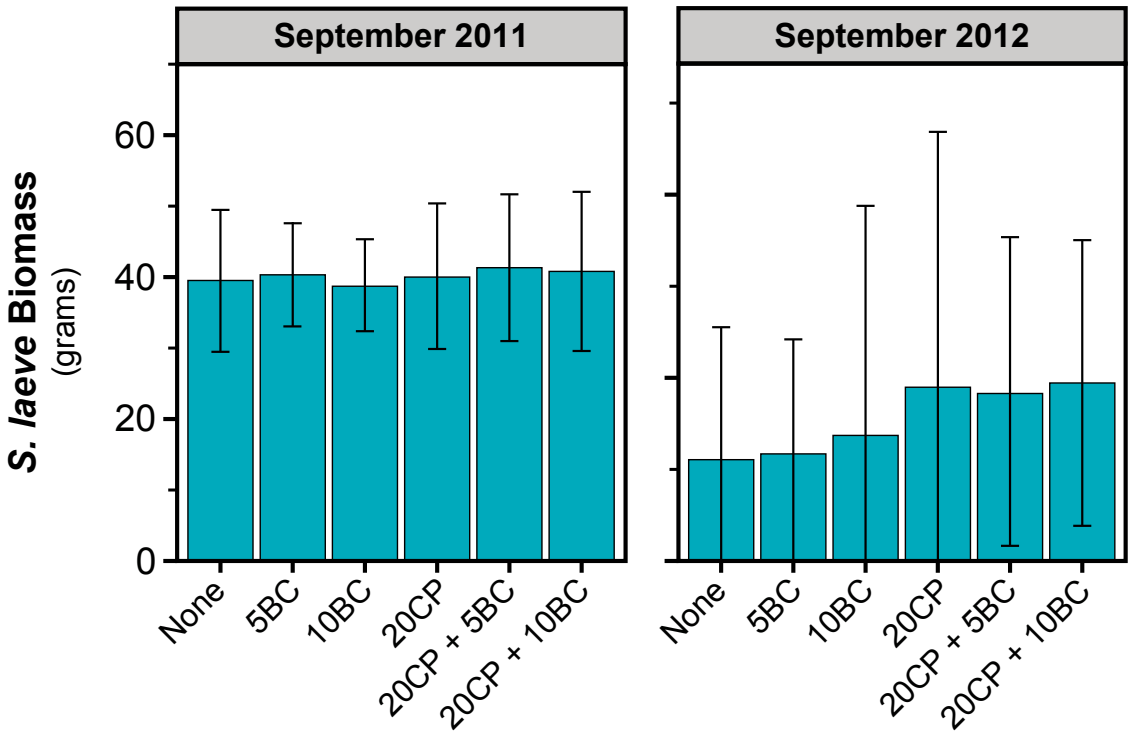
(b)

Model Terms	p–value	sig. level	% Expl. Dev.
Year	<0.001	***	43.34%
AM Inoculation	<0.001	***	0.88%

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.9: Predicted *Panicum virgatum* biomass in the plant plug trial. Panel (a) represents the graph of raw data with error bars (± 1 SD) based on the most parsimonious linear mixed effects model. Experimental treatment replication = 9. The left graph panel represents data after one growing season. Statistical output (b) shows significant main effect terms and interactions. Main effects included in the model were: Amendment, AM inoculation, Plot Height, and Year. % explained deviance is abbreviated as % Expl. Dev. in the output. Note that model terms with negative regression slopes are indicated in parentheses around the significance levels.

(a)



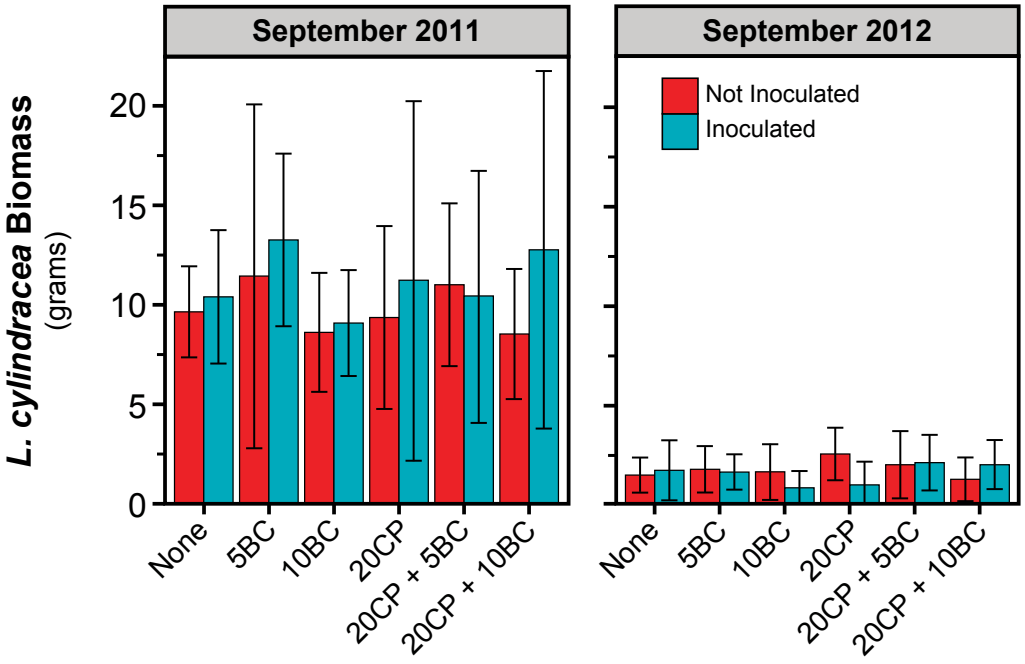
(b)

Model Terms	p–value	sig. level	% Expl. Dev.
Year	<0.001	(***)	34.74%
Plot Height (Dry → Wet)	0.001	**	1.86%
Interactions			
Amendment × Year	0.094	.	1.65%
Year × Plot Height	<0.001	***	2.21%

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.10: Predicted *Symphyotrichum laeve* biomass in the plant plug trial. Panel (a) represents the graph of raw data with error bars (± 1 SD) based on the most parsimonious linear mixed effects model. Experimental treatment replication = 9. The left graph panel represents data after one growing season. Labels on the x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Statistical output (b) shows significant main effect terms and interactions. Main effects included in the model were: Amendment, AM inoculation, Plot Height, and Year. % explained deviance is abbreviated as % Expl. Dev. in the output. Note that model terms with negative regression slopes are indicated in parentheses around the significance levels.

(a)



(b)

Model Terms	p–value	sig. level	% Expl. Dev.
Year	<0.001	(***)	70.96%
Plot Height (Dry → Wet)	<0.001	***	1.09%
Interactions			
Plot Height × Year	0.045	*	0.38%
Amendment × Inoculation	0.081	(.)	0.93%
Year × Inoculation	0.039	(*)	0.40%

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.11: Predicted *Liatris cylindracea* biomass in the plant plug trial. Panel (a) represents the graph of raw data with error bars (± 1 SD) based on the most parsimonious linear mixed effects model. Experimental treatment replication = 9. The left graph panel represents data after one growing season. Labels on the x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Statistical output (b) shows significant main effect terms and interactions. Main effects included in the model were: Amendment, AM inoculation, Plot Height, and Year. % explained deviance is abbreviated as % Expl. Dev. in the output. Note that model terms with negative regression slopes are indicated in parentheses around the significance levels.

the study (mean % cover: 1.1%, range: from 0.0% - 16.0%). When present, non–seeded vegetation was dominated by patches of the perennial herb *Artemisia campestris* (common name: field wormwood). Pooled C_3 and C_4 grasses largely dominated vegetative cover after three growing seasons (2013 mean % cover: 17.4%, range: 2.0% – 42.0%). The cover of N–fixing forbs was second most abundant by the third growing season (2013 mean % cover: 3.0%, range: 0.0% – 19.0%). The establishment and survival of composite forbs was sparse after three growing seasons (2013 mean % cover: 0.4%, range: 0.0% – 7.0%), leading to a negligible contribution to total plant cover.

Plant cover in the seed application trial

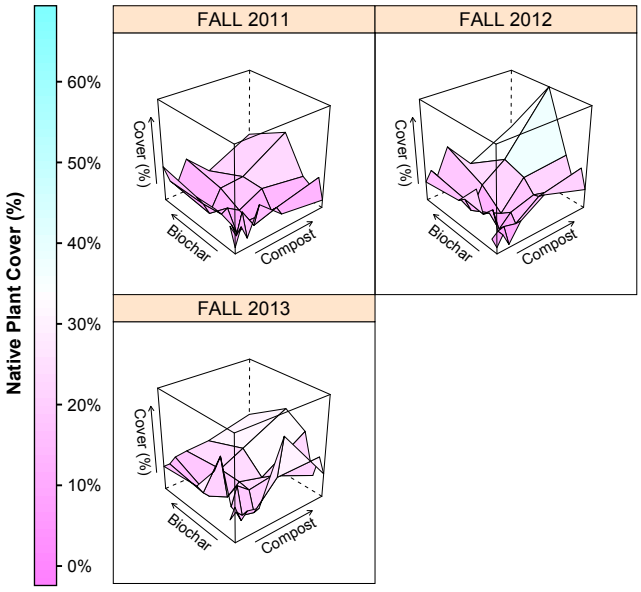
As main effects, compost rate ($p=0.025$) and growing season ($p<0.001$) were the most influential drivers of total plant cover. Model term interactions show variable positive and negative plant cover responses depending upon factor combination. Significant increases in total plant cover were largely driven by plots with three-way and four-way interactions among biochar, compost, AM fungal inoculation, and growing season (Figure 3.12a). The plot height covariate significantly influenced total cover where plots higher on the landscape had more plant cover regardless of treatment when accounting for growing season ($p=0.002$).

Biochar amendments When all other factors were held constant, increasing rates of biochar did not significantly influence total plant cover in the seed application trial ($p>0.05$)(Figure 3.13). In addition, the plot height covariate and growing season did not alter the influence of biochar in the field ($p>0.05$). The addition of biochar, regardless of application rate, resulted in no direct influence on total plant cover in this experiment.

Compost amendments With all other factors held constant, increasing rates of compost were a main driver of increasing total plant cover in the direct seeding study ($p=0.025$). The compost x year interaction ($p=0.001$) revealed a significant negative total plant cover response mainly driven by large variation in plots adding $40\,\mathrm{T\,ha^{-1}}$ compost (Figure 3.13).

AM fungal inoculation No direct influence of AM fungal inoculation was detected in the seed application trial ($p>0.05$)(Figure 3.13). A trend of decreasing plant cover was detected in plots adding AM fungal inoculum when accounting for growing season ($p=0.065$).

(a)



(b)

Model Terms	p–value	sig. level	% Expl. Dev.
Compost	0.025	*	0.92%
Year	<0.001	***	8.23%
Interactions			
Compost × Biochar	0.062	(.)	0.64%
Compost × Year	0.001	(**)	1.91%
Biochar × Inoculation	0.093	(.)	0.52%
Inoculation × Year	0.067	(.)	0.62%
Inoculation × Plot Height (Dry → Wet)	0.075	(.)	0.58%
Year × Plot Height (Dry → Wet)	0.002	(**)	1.79%
Compost × Biochar × Year	0.040	*	0.77%
Compost × Biochar × Inoculation	0.093	.	0.52%
Compost × Year × Inoculation	0.018	*	1.03%
Compost × Biochar × Inoculation × Year × Plot Height	0.012	*	1.17%

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Relationships with negative regression slopes are indicated by parentheses.

Figure 3.12: Raw data wireframe graph (a) of total native plant cover in the seed application trial based on the most parsimonious linear mixed effects model. Panels represent the three analyzed growing seasons (Fall 2011–Fall 2013). The gradient bar on the left indicates increasing % cover from magenta → cyan. Results are based on the most parsimonious statistical model. Significance levels and interactions for the model terms are given in the statistical output table(b). y–axis = % total plant cover; x–axis = biochar rate, z–axis = compost rate. AM fungal inoculation and relative plot height are not included in the graph due to visual complexity. % explained deviance accounts for the proportion of variation explained by each model term. Replication = 1.

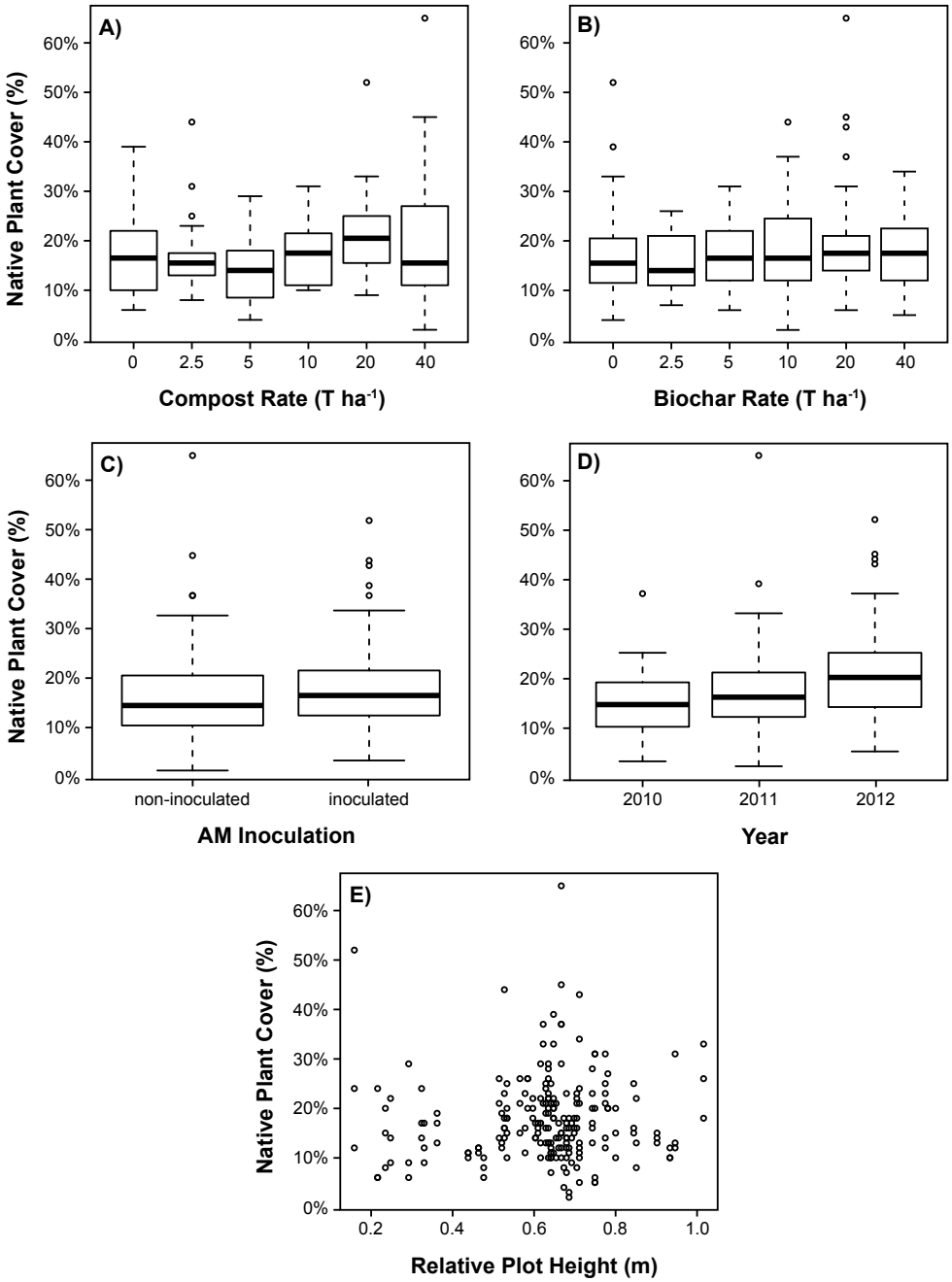


Figure 3.13: Diagnostic boxplots and a scatterplot for each main model term analyzing total native plant cover when all other factors were held constant in the seed application trial. Panels A–D are boxplots representing the raw data distribution for each categorical model term included in the linear mixed effect model. Panel E is a scatterplot of the relative plot height in meters compared to total native plant cover. The surveyed plots with relative plot height values closer to zero are higher on the landscape. At the field site, surface soils of plots higher on the landscape were observed to dry more rapidly than plots lower on the landscape.

Synergistic effects of biochar, compost, and AM fungal inoculation Plots with increasing rates of biochar and compost resulted in significant increases in total plant cover when accounting for growing season (p=0.040). When adding *R. irregularis* inoculum, total plant cover was significantly increased only when combined with increasing rates of compost + growing season (p=0.018) or compost + biochar (p=0.093). Overall, plot inoculation with AM fungal inoculum was most effective for increasing total plant cover when combined with high rates of biochar and compost while accounting for plot height and growing season influences (p=0.012).

Growing season Growing season explained the largest amount of variation in the model (Expl. Dev: 8.23%) (Figure 3.12a). Total % cover across all treatments was sparse after one growing season, gradually increasing by the second and third growing seasons (mean % cover: 14.9% (2011), 17.7% (2012), and 20.9% (2013)).

Plot height covariate The plot height covariate did not directly influence total plant cover in the model when all other factors were held constant (Figure 3.13). Plot height was a significant covariate when determining increasing plant cover in the interaction terms combining biochar, compost, AM fungal inoculation, and growing season (p=0.012). Plots with a higher position at the field site exhibited increased plant biomass when accounting for growing season (p=0.002).

3.4 Discussion

These field trials show that compost is the most influential driver to improve plant response in post-mine sandpit areas regardless of planting method. Supplementing compost with biochar and a commercial AM fungal inoculant largely accentuates its effectiveness by further increasing plant response in the seed application trial. The combination of compost, biochar, and AM inoculum are effective land management tools to restore grassland plants in severely disturbed post-mine sandpits.

Plant response to AM fungal inoculation

The effectiveness of AM fungal inoculation depended upon the restoration planting method in the field. Total plant biomass in the plant plug experiment was not significantly influenced by *R. irregularis* inoculum. Although AM colonization was significantly greater in inoculated plots in the plant plug trial, this did not translate to a consistent plant response. Comparatively, total cover in the seed application trial responded positively to the AM fungal inoculum addition only in the presence of high rates of compost and

biochar. In both trials, my hypothesis that AM fungal inoculation would directly account for increased total plant response in post–mine sandpits was not supported.

Studies indicate that the application of AM fungi in a mine restoration setting generates positive plant responses from seed compared to non–inoculated controls (Noyd et al. 1996, Johnson 1998, de Souza et al. 2010). But, these studies also suggest that the application of AM fungi does not consistently increase plant response for all plant species or restoration scenarios. In the seed application trial, the model suggests that AM fungal inoculum increased total plant cover in the presence of increasing compost and compost + biochar rates. Under similar mine area conditions, Püschel et al. (2011) shows that plant response was maximized in mine spoil banks co–amended with AM fungal isolates and organic matter treatments due to increased resource availability. Contrary to this study, plant response to AM fungal inoculum in mine areas can vary in the presence of soil conditioners such as organic matter or fertilizer (Gryndler et al. 2008). Gryndler et al. (2008) revealed that the addition of compost increased plant performance in reclaimed clay substrate to the detriment of a plant biomass effect from AM fungal inoculation. Therefore, the addition of AM fungal inoculum does not universially benefit plant growth in severely degraded mine areas.

In the plant plug trial, inoculated and non–inoculated plants were colonized by *R. irregularis* and the inoculation effect persisted in the field for two growing seasons. Yet, plant host–fungal pairings in this study yielded mixed results in terms of plant response to the commercial inoculum. *P. virgatum* and *L. capitata* responded positively to the commercial inoculum while *A. gerardii* and *L. cylindracea* biomass decreased. Plant–mycorrhizal associations range from mutualistic to parasitic depending upon the environmental context and host species (Johnson et al. 1997). Klironomos (2003) shows that AM fungal–plant pairings are known to be unpredictable in biomass responses along a mutualistic–parasitic continuum. In my mixed community of grassland plants, the effectiveness of commercial inoculum was shown to be variable depending upon plant species. Thus, the universal application of a single AM fungal isolate may not benefit all target plants when restoring grassland habitat.

In the plant plug trial, AM fungal colonization was initially present in the non–inoculated control plugs due to non–sterile greenhouse conditions. A mycorrhizal effect induced by *R. irregularis* on plant response may have been significant if inoculated plant plugs were compared to sterile, non–colonized plant plug roots. In a transplant study of *Sporobolus wrightii* into agricultural fields, greenhouse grown plants inoculated with AM fungi had greater survival, larger basal diameters, and increased tiller production after two growing seasons (Richter and Stutz 2002). That study compared plants growing in sterilized soil versus plants inoculated with AM fungi. Therefore, it is possible that the background mycorrhizal community present in non–inoculated plant plugs grown in the greenhouse may

have also benefited plant growth in the field. Thus, a mycorrhizal effect from *R. irregularis* may be minimized due to the presence of a background AM fungal community in my study.

Compared to the plant plug trial, a background community of AM fungi was not introduced into the seed application experiment in the non–inoculated controls. The sandpit did not have a legacy of plant growth prior to planting. Therefore, inoculum potential in the mine substrate was anticipated to be very low (Allen and Allen 1980, Stahl et al. 1988). Furthermore, any established hyphal networks present before the disturbance would have reduced infectivity in the severely disturbed mine system (Jasper et al. 1989). In my study, the addition of AM fungi was potentially more effective in promoting a plant response in the seed application trial due to a lack of an infective background AM fungal community at the site.

Over time, the immigration of AM fungal propagules adapted to post–mine conditions is expected to develop and benefit new plant recruits in degraded areas (Ganesan et al. 1991). Thorne et al. (2013) found no difference in dominant prairie plant growth when comparing AM fungi collected from a 30-year reclaimed mine spoil and a tallgrass prairie soil. The commercial fungal isolate used in that experiment, *R. irregularis*, may not be well–adapted to mine land conditions, thus exhibiting an inconsistent plant response in the plant plug trial. Taheri and Bever (2010) found that plants growing in mines are particularly dependent on strong AM fungal partners under harsh edaphic condition. Larger plant biomass responses were induced when locally adapted AM fungi from recovering mine areas were applied as inoculum. Inoculating plants with AM fungi isolated from sandy, post–mine habitats may increase target plant biomass due to local adaptation of the inoculum.

Plant response to biochar

Biochar as a solitary amendment was not an effective land management tool to promote plant growth in either trial. Applying biochar alone reduced plant growth compared to most compost and compost + biochar amended plots in the plant plug experiment while exhibiting no significant difference from non–amended controls. In the seed application experiment, biochar rate did not directly influence total plant cover but was an effective tool to promote plant response when used in combination with compost and AM fungal inoculum. In both trials, my hypothesis that biochar would directly account for increased total plant response in post–mine sandpits was not supported.

The high cation exchange capacity of biochar may have bound nutrients in the soil solution, thus introducing more abiotic stress into the plant plug trial (Liang et al. 2006). Positive vegetative response to biochar application in disturbed mine substrates has been attributed to increased water holding capacity, nutrient retention, and reduced soil bulk density (Fellet et al. 2011, Kelly et al. 2014). These nutrient retention properties may have

been to be detrimental to rapidly growing plants in my study. Xu et al. (2013) found that biochar application to sandy soils resulted in pH changes and mineral sorption to alter nutrient bioavailability and reduce predicted total phosphorous. In metal contaminated mine sites, Beesley et al. (2014) found high metal adsorption rates when co–amending soils with biochar and compost. The potential nutrient retention (Ding et al. 2010) and release (Mukherjee and Zimmerman 2013) of biochar alone did not result in increased total plant biomass in the plant plug trial. Further research should be conducted to determine the efficacy of applying biochar as a solitary amendment in unconsolidated mine substrates.

The properties of biochar as a solitary amendment may only be beneficial depending upon the environmental context. Experiments on biochar amendments have been largely restricted to agricultural soil systems and test plants have exhibited mixed growth results (Major et al. 2010, Jones et al. 2012, Filiberto and Gaunt 2013). A solitary study investigating biochar's effect on a native and invasive grassland plant in a greenhouse experiment indicates increased biomass of the native plant, *Andropogon gerardii*, while reducing the growth of the invasive plant, *Lespedeza cuneata* (Adams et al. 2013). The results of my study contradict Adams et al. (2013) as biochar had no significant influence on the biomass of *Andropogon gerardii*. A key difference between the studies was that the greenhouse experiment used natural soils collected from a 2–yr–old prairie restoration site on former agricultural land. Biochar's effectiveness as a soil amendment may be more promising when restoring former agricultural landscapes as compared to low quality post–mine soils.

The negative effects of biochar were not evident in the seed application trial. With all other factors held constant, biochar rate had no direct effect on total native plant cover. One reason may be that plant plug individuals grew much larger and faster than seeded plants over the same time period. Plant competition for soil nutrient resources may have outpaced a fertilization benefit gained by low rates of biochar at the site. The increased root stock associated with gains in aboveground plant tissues would have a high nutrient and water demand in the mine substrate (Craine and Dybzinski 2013). Comparatively, the growth of the grassland plants in the seed application trial was stunted over the same growing period. The nutrient requirements of plants in the seed application trial may have been reduced, thus minimizing the negative effects attributed to biochar. Therefore, restoration practitioners should approach the application of biochar to abioticially stressed mine areas with caution. Unintended reductions in plant community response may cascade through the system by creating more stressful plant growth conditions after the application of biochar in the field.

Plant response to compost

Compost as a solitary amendment was the most effective soil amendment to promote plant growth in both trials. Applying compost alone promoted positive plant growth in the plant plug trial with compared non–amended and biochar amended plots. In the seed application experiment, increasing the rate of compost directly contributed to total plant cover and its effectiveness increased when used in combination with biochar and AM fungal inoculum. In both trials, my hypothesis that compost would directly account for increased total plant response in post–mine sandpits was supported.

It is not surprising that compost resulted in improved plant growth, likely due to its well–known fertilizer effect, its ability to increase water retention, and create higher cation exchange capacity in soils (Shiralipour et al. 1992). Similarly, compost has been shown to promote plant production from seed in other severely disturbed mine restoration scenarios (Hortenstine and Rothwell 1972, Norland and Veith 1995, Noyd et al. 1996). A one–time application of compost is shown to be an effective, readily available technical reclamation tool that accelerates prairie plant growth from seed in my study. The long–term residual effects of compost are anticipated to continue to benefit plant growth in upcoming growing seasons (Diacono and Montemurro 2010).

In the seed experiment, high rates of compost were the most influential technical reclamation tool to alleviate abiotic plant stress as evident by increased total plant cover. A reduction in plant cover was detected in plots adding compost when accounting for growing season. Investigating the patterns in the raw data, high variability was detected in plots adding compost at $40\,\mathrm{T\,ha^{-1}}$ with increasing rates of biochar. This created large variance in the final predictive model for this treatment level with all other factors held constant. As the model interaction terms increased in complexity, total plant cover values showed a positive response as biochar rates and AM fungal inoculum were incorporated into the linear mixed effects model. Increased plot replication in the seed application trial would have accounted for the inevitable natural environmental stochasticity at the field site.

Synergisms among biochar, compost, and arbuscular mycorrhizas

Combining biochar, compost, and AM fungal inoculum had minimal effect in the plant plug trial in terms of total plant biomass compared to non–amended controls. The highest rates of compost + biochar, $20\,\mathrm{T\,ha^{-1}}$ of compost + $10\,\mathrm{T\,ha^{-1}}$ of biochar, had the largest positive influence on total plant biomass compared to control although not significant. In the seed application trial, the synergistic effect of increasing rates of compost and biochar combined with AM fungal inoculation was effective in promoting total plant cover in the field. My hypothesis that the synergistic effect of all amendments would account for the

largest plant response in each trial was weakly supported in the plant plug trial and strongly supported in the seed application trial.

Compost rates in the plant plug experiment may have been too low to elicit a strong plant response in the plant plug trial. Noyd et al. (1996) shows a significant increase in plant grassland plant cover in taconite mine spoils when increasing compost rates from $22.4\,\mathrm{T\,ha^{-1}}$ to $44.8\,\mathrm{T\,ha^{-1}}$ after three growing seasons. In clay spoils, Püschel et al. (2008a) indicates that three high compost amendment rates ($100\,\mathrm{T\,ha^{-1}}$, $200\,\mathrm{T\,ha^{-1}}$, $500\,\mathrm{T\,ha^{-1}}$) significantly increased the flax biomass compared to controls. However, only negligible increases in plant biomass were detected among these compost treatments. As such, Püschel et al. (2008a) suggested that lowering compost rates would be more cost effective for industrial applications when restoring mine spoil areas. When growing plant plugs in Ontario's sandpits, the optimal compost application rate should be determined to maximize plant response at an industrial scale.

In the seed application trial, the most effective amendment rates occurred when both biochar and compost were applied at $20\,\mathrm{T\,ha^{-1}}$ or greater. Studies investigating co–amended soils indicate mixed plant growth results when adding compost and biochar to soils (Ghosh et al. 2014, Schmidt et al. 2014). In my study, higher rates of compost and biochar in extremely degraded systems may be required to produce a larger plant response. More research needs to be conducted to determine optimal rates of biochar and compost when restoring plant communities in Ontario's post-mine sandpits.

The effect of biochar may have been enhanced by the concurrent addition of compost by charging biochar surfaces and promoting a plant response in both trials (Fischer and Glaser 2012). It has been shown that biochar and compost amended soils stimulates microbial growth and respiration rates thus enhancing decomposition rates and nutrient availability during the composting process (Steinbeiss et al. 2009). Fischer and Glaser (2012) show that plant growth generally increased with increasing amendment of biochar and compost amendments, especially in nutrient poor, sandy soil. I similarly suggest that mixing high biochar rates in conjunction with high compost rates compliments the nutrient retention properties of biochar in soils when restoring post–mine sandpits.

The addition of AM fungal inoculum in the seed application trial was most effective as compost and biochar rates increased. Therefore, the alleviation of stressful edaphic conditions by the soil amendments may have facilitated the effectiveness of the plant host–fungal pairings leading to a positive plant response in the field. In the plant plug trial, the addition of $20\,\mathrm{T\,ha^{-1}}$ of compost and/or $5\,\mathrm{T\,ha^{-1}}$ and $10\,\mathrm{T\,ha^{-1}}$ of biochar may have been too low favor a positive plant response due to AM fungal inoculation in all plant species. Using locally adapted inoculum may be more effective in promoting a plant response in restoration scenarios when applying lower rates of soil amendments (Gryndler et al.

2008). Further research needs to be conducted regarding the most appropriate mycorrhizal inoculum to include in the restoration of tallgrass prairie species in abandoned sandpits.

In my study, plug plants were germinated under stress–free conditions in the greenhouse. In contrast, germinating plant seeds under field conditions had to overcome the stressful environment of post-mine substrate. When restoring from seeds in field, plants benefited more from the AM fungal symbiotic association due to the alleviation of stress, especially in the presence of high compost and biochar rates. This was most likely due to increased access to soil nutrients and more favorable water balance in the plant provided by the AM fungal symbiosis. The difference in AM fungal inoculum efficacy between plugs and seedling trials may be due to early soil conditions experienced by germinating seedlings. Spontaneous plant succession of post–mine soils is often restricted by poor soils, a lack of biotic symbionts, and restricted local seed immigration (Prach and Hobbs 2008). When spontaneous succession of plants in quarries is observed, weedy annual plant species often persist (Khater et al. 2003). Thus, even with the introduction of grassland plant seeds to mine areas, plant germination would be restricted if soil amendments and AM fungal inoculum are not incorporated.

In the presence of increasing amendments in the seed application trial, AM fungi had a greater access to a pool of nutrients and greater water availability that was provided by the compost and biochar. Research by Hodge and Fitter (2010) shows that the AM fungal symbiosis with plants improves fungal nitrogen acquisition from decomposing organic matter which is especially beneficial in nitrogen limited systems. The transfer of nitrogen to the plant can therefore benefit plant establishment and production when restoring of degraded mine areas (Govindarajulu et al. 2005). These results indicate that AM fungal inoculum will be most effective after the stressful abiotic conditions are improved with soil amendments.

Plant growth dynamics

In the plant plug experiment, four of the six measured plants, (*Desmodium canadense*, *Lespedeza capitata*, *Symphyotrichum laeve*, *Liatris cylindracea*, had significantly reduced plant biomass between Fall 2011 and Fall 2012 despite the addition of soil amendments. The selection of these plant species was not optimal for long–term growth in post–mine sandpits. Only the C_4 grasses, *Andropogon gerardii* and *Panicum virgatum*, had significant biomass increases during the same time period. The large increase in C_4 grass biomass accounted for the majority of the detected increase in total plant biomass between Fall 2011 and Fall 2012. Long–term monitoring of these plots needs to be conducted in order to track the potential biodiversity loss in this trial.

In the seed application experiment, relative plot height had a significant influence on total native plant cover. Conflicting results were identified in the model depending upon

interactions among model terms. The interaction between growing season and relative plot height led to a significant decline total native plant cover. High plant cover variation in plots adding no compost exists within this two–way interaction, thus driving this significant decline in the model. Conversely, plot height was a significant term in the five–way interaction among the experimental variables and growing season. The resolution of the five–way interaction model most accurately represents the true experimental design in this trial. I speculate that a decrease in the relative plot height would increase water availability at the site. In this trial, higher water retention would be expected as compost rates and biochar rates increase (Aggelides and Londra 2000, Movahedi-Naeini and Cook 2000, Abel et al. 2013). As abiotic soil measurements were not collected during the study, a detailed soil moisture analysis needs to be determine to understand the driving abiotic factors determining variation in the total plant cover measurements.

3.5 Summary

The restoration of tallgrass prairie plants in post–mine aggregate sites is a viable management option in southern Ontario. Increasing grassland community diversity on marginal, anthropogenically influenced lands through prairie restoration will ensure the survival of sensitive habitat in addition to supporting species at risk. But, the harsh edaphic characteristics present in the post–mine sandpits restricts plant community development in the field. Thus, restoring self–sustaining and diverse grassland communities is not possible without acknowledging the impoverished conditions of post–mine substrates.

Incorporating land management tools to mitigate the harsh abiotic conditions of post–extraction substrate is therefore necessary to increase plant production in target grassland communities. The technical reclamation tools investigated in this study led to higher seed establishment rates and total plant biomass when used in combination. A single application of high rates ($20\,\mathrm{T\,ha^{-1}}$) of biochar and compost at the onset of an industrial–scale restoration project can lessen site maintenance costs, increase plant community recovery time, and promote vegetative biodiversity.

Chapter 4

Soil Food Webs

4.1 Background

Ecosystem productivity and fertility are characterized by organic matter inputs, primary production, and microbial energy pathways (Wardle et al. 2004). The high density and diversity of biota in soils influence these ecosystem services by altering soil water storage, litter decomposition rates, and nutrient cycles (Doran and Zeiss 2000). Ultimately, ecosystem services provided by microbial communities and soil animals in belowground food webs can influence plant community growth in mine areas restoration (Wardle 1999). Thus, this study explored the bottom–up effects of multi–trophic group interactions in belowground food webs during a grassland restoration in a post–extraction sand pit.

Grassland restoration projects focused on recovering soil food webs in post–mine sandpits must overcome harsh edaphic conditions and the lack of soil organic matter. Over time, soil invertebrates and microorganisms recolonize restored areas without assistance, but management intervention can accelerate the establishment of desirable species (Curry and Good 1992). Incorporating soil amendments into post–mine areas can increase soil organic matter, fertility, and water-holding capacity, ultimately influencing soil food web development. Increasing plant production with amendments and arbuscular mycorrhizas can further contribute to the development of an active plant rhizosphere, thus further stimulating the soil food web (Kuzyakov 2002).

4.1.1 Arbuscular mycorrhizal fungal inoculum

The plant–AM fungal association extends the biologically active zone around plant roots (i.e. rhizosphere) to incorporate the influence of the plant symbiont (i.e mycorrhizosphere). Soil areas under the influence of the mycorrhizosphere have rapid water and nutrient uptake, high concentrations of exudates, high root turnover, and increased respiration (Garbaye 1991). The plant-AM fungal symbiosis has been shown to enhance rhizobial N–fixation by legumes (Amora-Lazcano et al. 1998) and increase bacterial populations (Johansson et al. 2004) compared to non–mycorrhizal plants. Therefore, soil animals may benefit from increased microbial activity associated with the mycorrhizosphere due to higher food availability. Thus, increased plant root biomass and mycorrhizosphere activity due to AM

fungal inoculation can translate into greater soil food web production in these zones of high root exudates and plant root turnover.

4.1.2 Biochar

Biochar as a soil amendment is anticipated to prime fungal and bacterial biomass by providing carbon substrates, retaining soil macro–nutrients, and/or providing suitable microbial refugia (Lehmann et al. 2011, Lou et al. 2014). Lower temperature biochar (250 °C to 400 °C) is anticipated to most improve soil fertility and stimulate microbial communities (Novak et al. 2009).

The mechanisms of biochar's influence on microbial biomass and soil animal abundance are understudied (McCormack et al. 2013). Thus, it is difficult to predict the effects of biochar in post–mine restoration on soil biotic communities due to the dearth of research and the complexity of biological and physical interactions in soil. But, I anticipate that biochar's projected physiochemical benefits in soils will increase plant growth and soil microbial communities, thus increasing food resources for higher trophic levels in the soil food web.

4.1.3 Compost

Compost increases microbial community biomass, soil respiration rates, and soil enzyme activity by providing bacteria and fungi with decomposable substrate (Allievi et al. 1993). Compost has been shown to favor the development of fungal–dominated systems by adding complex organic matter to degraded systems (Biederman and Whisenant 2009, Biederman 2013). Microbial growth and soil fertility are closely related as compost is decomposed by the microbial community. Thus, the decomposition of compost releases important elements (N, P) into the soil solution to be taken up by organisms (Frankenberger and Dick 1983).

Compost can be considered a bio–inoculant with an associated community of fungi and bacteria, nematodes, and microarthropods (Cernova 1970, Streit et al. 1985, Steel et al. 2013a;b). As soil animal communities are severely diminished in post–mine habitats, the addition of compost is anticipated to give compost amended areas a soil food web head start. Furthermore, increased microbial biomass associated with compost should provide soil animals with an increased food supply. Thus, compost is expected to be an essential soil amendment crucial to the development of soil food webs.

4.1.4 Synergisms among biochar, compost, and arbuscular mycorrhizas

This study is the first to test the concurrent application of biochar, compost, and commercial inoculum on soil food webs. Although conceptually recommended to incorporate biochar and compost simultaneously (Fischer and Glaser 2012), no studies have researched

the soil food web response to the application of compost, biochar, and AM fungal inoculum in a restoration setting. As compost is expected to positively ameliorate soils, co–amended soils with biochar can further increase water and nutrient retention. Increased mycorrhizosphere activity and plant root biomass associated with AM fungal inoculation is anticipated to further accelerate the development of soil food webs in post–mine sandpits. Thus, co–amended soils with biochar, compost, and AM fungal inoculum are expected to exhibit the greatest increase in soil food web response.

4.1.5 Hypotheses

To my knowledge, the relationship among compost, biochar, arbuscular mycorrhizal (AM) fungal inoculation, and the soil food web has never been tested when restoring grassland vegetation in post–mine sandpits. I hypothesized that compost, biochar, and AM fungal inoculation would individually increase soil microbial biomass compared to non–amended controls. In addition, I hypothesized that the concurrent addition of compost, biochar and AM fungal inoculum would yield the largest response in plant functional group biomass and soil microbial community biomass. The rationale for these hypotheses is that the amendments and AM fungal inoculation would alleviate of water and nutrient stress and increase mycorhizosphere activity in post–mine sandpit substrates. A structural equation model was constructed to describe the direct, indirect, and total effects driving soil food web response among plant functional biomass, soil organisms abundance, and soil amendments (i.e. compost, biochar, and AM fungal inoculum). As a guide to create my structural equation model, trophic interactions from Hunt et al. (1987) form the basis in determining relationships within soil food webs. The goal of this study was to prescribe industrially feasible abiotic and biotic soil amendments to facilitate soil development and determine the trajectory of soils in recovering post–mine sandpits when restoring grassland communities.

4.2 Methods

I tested the fully factorial effects of soil amendments (biochar, compost) and AM fungal inoculation on soil microbial biomass and soil animal abundance in the post-mine sandpit. Soils were collected from plots in the plant plug trial after two growing seasons. Full experimental design details of the plant plug trial are described in *Chapter 3: Methods*.

4.2.1 Soil collection and organismal analyses

In September 2012, sixteen soil cores (2.54 cm diameter) were collected to a depth of 12 cm from each plot in the plant plug experiment. Soil cores were sampled directly adjacent

to aboveground plant tissue for each of the six C_4 grasses, five N–fixing forbs, and five composite forbs within the core sampling area. Soil cores were sampled directly adjacent to each plant at each plug location to minimize plant destruction. The soil corer was cleaned of substrate with a clean cloth and water between plot sampling to minimize contamination. Collected soils were pooled at the plot level and homogenized. Soil samples were stored in a cooler on ice in the field until final storage at 4°C. Soil microbial biomass and soil animal abundance were analyzed at the Soil Analysis Laboratory, University of California, Riverside.

Bacterial and fungal biomass

Bacterial and fungal biomass was estimated by differential fluorescent staining (DFS) following an adapted protocol by Klironomos et al. (1996). The DFS was composed of a mixture of europium(III)thenoyl–trifluoroacetonate and a fluorescent intensifier (Anderson and Westmoreland 1971). For fungal biomass estimation, 200 ml of soil was suspended with 1 ml of DFS stain for 1 hour. Once stained, the suspension was filtered through nitrocellulose filter paper using a 50% ethanol wash. Filters were then mounted on microscope slides for visual inspection under UV light at 620 nm. Active cellular material was visually highlighted with red fluorescence under UV light. Fungal biomass was calculated from images taken by computer imaging software. Hyphal length was measured to estimate milligrams (mg) of fungal biomass kilogram $(kg)^{-1}$ soil using conversion factors [hyphal diameter: 1.65 micrometers (μm)(Kjøller and Struwe 1982), density: $0.33\,g\,cm^{-3}$ (van Veen and Paul 1979), C content: 45% (Swift et al. 1979)].

To estimate bacterial biomass, soil dilution aliquots were stained with DFS for 1 hour. After staining, filters were rinsed with a 50% ethanol wash and slides mounted for visual inspection using UV microscopy at 620 nm. Active cellular material was visualized by red fluorescence and images taken with computer imaging software. Bacterial biomass was estimated using a conversion factor of 6.4×10^{-14} g carbon $cell^{-1}$ calculated by Hunt and Fogel (1983). Results are given in mg of bacterial biomass kg^{-1} soil.

Nematode enumeration

Nematodes were extracted by the same wet sieve sucrose centrifuge approach for extracting arbuscular mycorrhizal spores as described in Klironomos et al. (1993). Soil samples were suspended in water and passed through a series of mesh sieves decreasing in pore size (1.0 mm – 45μm). After rinsing with water, the material retained in the 45μm sieve was suspended on top of a 60% sucrose solution and centrifuged for 20 minutes. Nematodes were collected via a pipette at the sucrose–water interface. Nematode individuals were sorted

and counted under a microscope. Individuals were classified into one of three functional feeding groups: bacterial feeding, fungal feeding, and predatory. Categorization was based upon nematode morphological characteristics. Abundance is reported as # of individuals g^{-1} soil.

Soil arthropod enumeration

A high efficiency canister–type soil arthropod extractor (Lussenhop 1971) was used to extract mites and Collembola onto dishes containing picric acid as described in Klironomos et al. (1996). Soil arthropods were counted and classified into Collembola, microbial feeding (Oribatid) mites, and predatory mites based on morphological characteristics. Abundance is reported as # of individuals g^{-1} soil.

4.2.2 Statistical analyses for soil biota

Linear models were used to test treatment–level effects on bacterial and fungal biomass. Fungal and bacterial biomass data approximated a normal data distribution after a log transformation. Generalized linear models with a negative binomial distribution link function were used to test the treatment effects on soil animal abundance (i.e. nematodes, Collembola, mites). As with most count data, all soil animal abundance data displayed characteristics of over–dispersion (the variance of the response variable exceeded the mean) and is best analyzed using a generalized linear models (Bolker et al. 2009).

Linear and generalized linear models for each soil functional group were reduced using model selection procedures. Full models including treatment factors, covariates, and interactions were iteratively reduced to remove non–significant variables using χ^2 tests. This resulted in the most parsimonious model to test differences in the response variable. Linear models were analyzed using the *base* package in R (R-Core-Team 2013). Generalized linear models were analyzed with the *glm.nb* function from the MASS package to calculate estimates from a negative binomial distribution.

4.2.3 Soil food web analysis with structural equation modeling

Structural equation modeling was used to test multivariate hypotheses and their interdependencies among soil functional groups, experimental treatments, and plant functional group biomass. *A priori* soil food web hypotheses, developed from prior literature knowledge, were determined among exogenous variables (i.e. soil amendments, AM fungal inoculation) and endogenous variables (i.e. biota in the soil food web, plant functional group biomass)(Figure 4.1). Before running the analysis, a covariance matrix of relation-

ships among all variables suggested that log transformations were appropriate to better approximate normality of residuals.

Plant biomass was estimated by partial least squares regression (PLS) for C_4 grasses, C_3 grasses, and N–fixing forbs (See *Chapter 3* for full details). Plant functional group biomass was pooled at the plot–level in September 2012 and used as a variable in the structural equation model.

Structural equation models were conducted in IBM's SPSS program extension AMOS. Overall fit of the *a priori* hypotheses was tested by evaluating χ^2 tests and the comparative fit index (CFI). An acceptable fit of a model is indicated by non–significant p–values in χ^2 tests and CFI values over 0.93 (Byrne 2013). Generalized least squares (GLS) with a bootstrap correction was used to calculate squared multiple correlations and standardized path coefficients. Squared multiple correlations were calculated for each endogenous variable to determine explained variance by incoming predictor variables. Standardized path coefficients were calculated from maximum likelihood distributions based on each variable's standard deviations (Grace 2006). Standardized path coefficients are interpreted as the expected change in response variable for each unit increase of the explanatory variable.

Three structural equation models were compared by selecting the model with the lowest AIC value to remove non–significant explanatory variables in the soil food web. Removing non–significant variables increases statistical power and improves structural equation model parsimony. Note that predatory and Oribatid mites were pooled in all models due to low correlations and large variances in this data.

Model 1: *Exogenous variables:* compost rate, biochar rate, AM fungal inoculation; *Endogenous plant biomass variables:* composite forbs, N-fixing forbs, C_4 grasses

Model 2: *Exogenous variables:* compost rate, biochar rate; *Endogenous plant biomass variables:* composite forbs, N-fixing forbs, C_4 grasses

Model 3: *Exogenous variables:* compost rate, biochar rate; *Endogenous plant biomass variables:* N-fixing forbs

Only biologically relevant standardized direct pathways with estimates greater than 0.1 or less than -0.1 were included in Figure 4.5 – Figure 4.12. Pathway significance ($p<0.05$) and trends ($0.05<p<0.1$) are indicated by colored arrows. Non–significant pathways ($p>0.1$) are indicated by dashed lines. All direct, indirect, and total standardized regression coefficients are given in Tables 4.1 – Table 4.3.

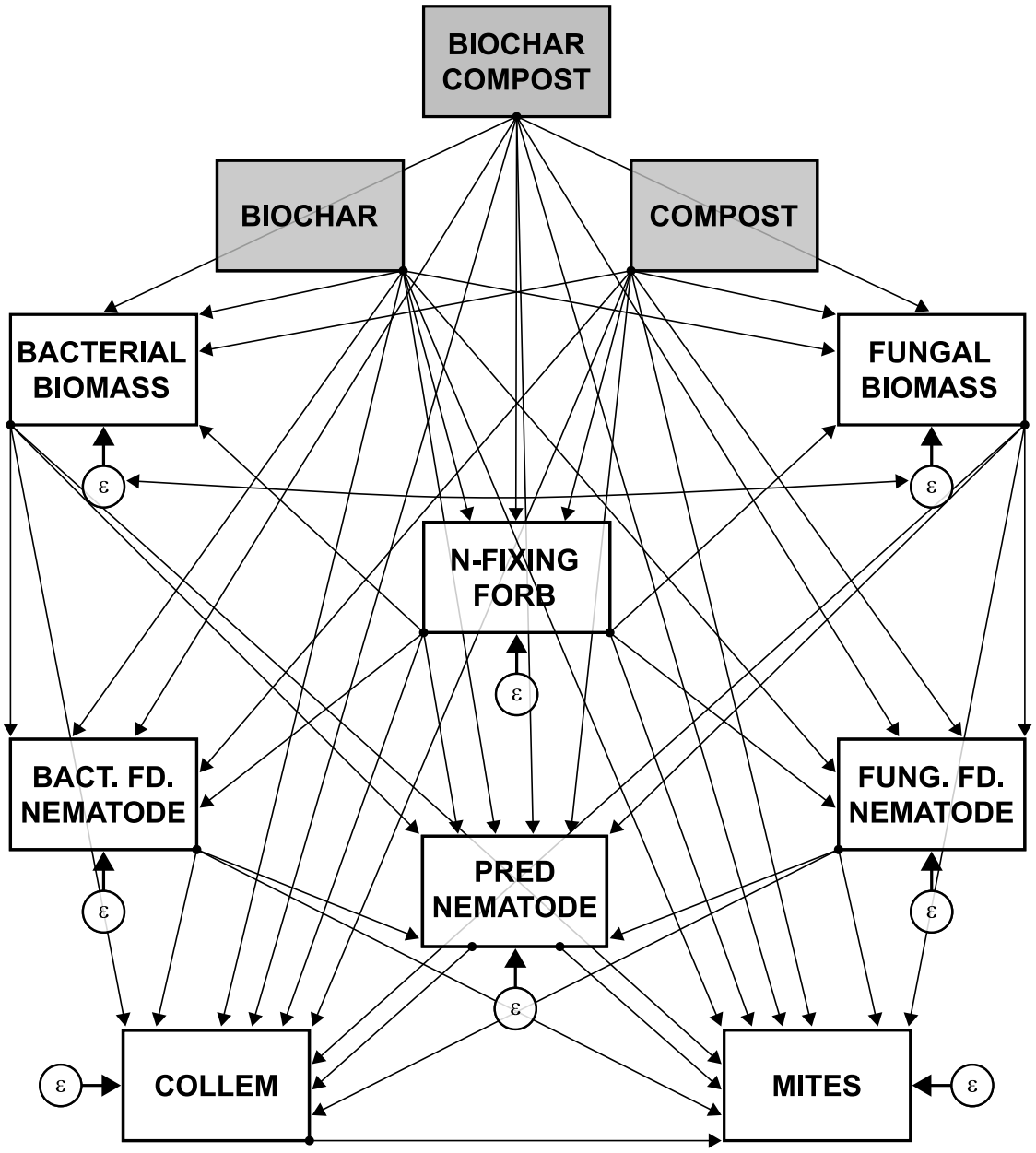


Figure 4.1: *A priori* hypotheses used to construct the most parsimonious structural equation soil food web model (Model 3). Exogenous variables are displayed in shaded gray boxes. Endogenous variables are displayed in white boxes. The residual error associated with each endogenous variable is displayed as (ε). Single headed arrows indicate direct pathways.

4.3 Results

4.3.1 Soil food web structural equation model selection

Model #1 (χ^2=17.381, df=13, p=0.182; CFI=0.955; AIC=201.4), Model #2 (χ^2=17.276, df=11, p=0.100; CFI=0.940; AIC=177.3), and Model #3 (χ^2=7.9, df=5, p=0.162; CFI=0.970; AIC=129.9) had an acceptable fit for testing *a priori* hypotheses as shown by non–significant χ^2 comparisons and CFI values > 0.93. Nevertheless, Model #3 (*Exogenous variables:* soil amendment rate; *Endogenous plant biomass variables:* N-fixing forbs) was chosen as the final model because it had the highest statistical power and parsimony as shown by lowest AIC values. Therefore, AM fungal inoculation, composite forb biomass, and C_4 grass variables were removed from the final model as they did not have sufficient explanatory power. Inoculated and non–inoculated plots were pooled by soil amendment rate (replication = 18).

4.3.2 Microbial community biomass and soil animal abundance

The influence of AM fungal inoculum on soil biota

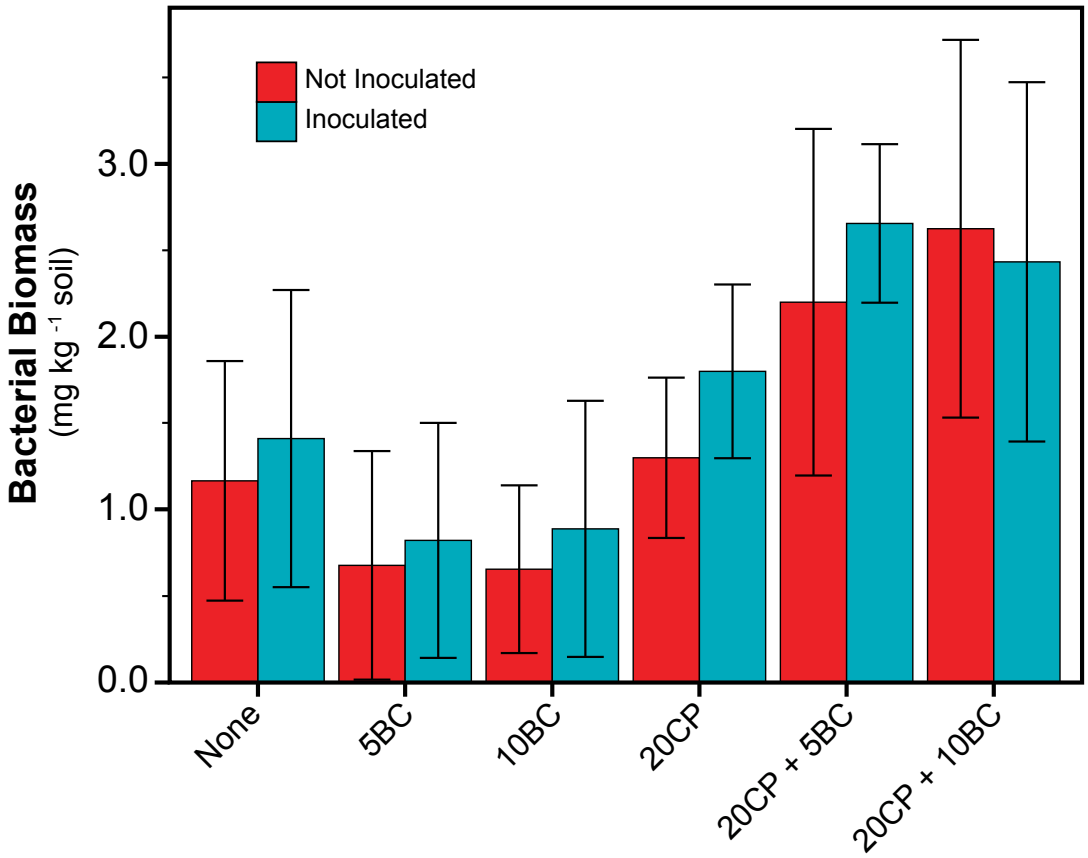
The response of the soil microbial community to the AM fungal inoculation of plant plugs was not significant in this study. No significant effect of AM fungal inoculation was detected for fungal biomass or the fungal:bacterial ratio as the AM fungal inoculation term was dropped from both linear models. Plots inoculated with AM fungi exhibited a positive trend in bacterial biomass at the study site (p=0.085)(Figure 4.2a).

Fungal feeding nematode (p=0.035) and Collembola (p=0.043) abundance was significantly reduced in plots inoculated with AM fungi (Figure 4.3a & 4.4a respectively). All other grazing and predatory soil animals were not significantly influenced by *R. irregularis* inoculation.

The influence of biochar on soil biota

Biochar's influence on soil microbial community biomass and soil animal abundance was consistently negative compared to non–amended control plots. Although not always significant, both biochar rates reduced microbial biomass and soil animal abundance in all cases as evident by negative coefficient estimates in the statistical output of the models. Biochar as a solitary amendment had no significant direct impact on the soil microbial biomass, soil animal abundance, or N–fixing forb biomass in the structural equation model (Figure 4.5 & Table 4.1).

(a)



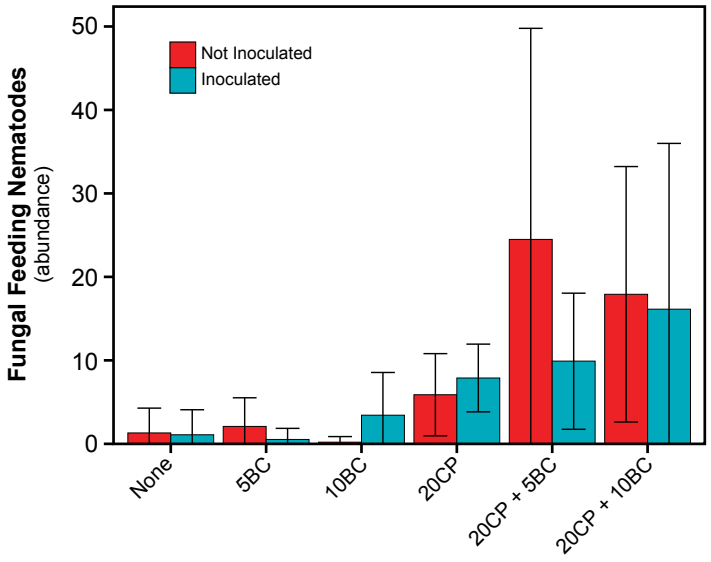
(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	0.312	0.033	<0.001	***
5BC	-0.116	0.044	0.010	(**)
10BC	-0.109	0.044	0.015	(*)
20CP	0.064	0.044	0.149	n.s.
20CP + 5BC	0.190	0.044	<0.001	***
20CP + 10BC	0.194	0.047	<0.001	***
AM Fungi	0.044	0.025	0.085	.

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Significantly different intercepts with negative values in parentheses

Figure 4.2: Bacterial biomass collected during the second growing season of the prairie restoration (September 2012). Data were analyzed with linear models to test treatment–level effects. Panel (a) represents raw data ± 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model term estimates represent the expected change from the model intercept (i.e. control plots).

(a)



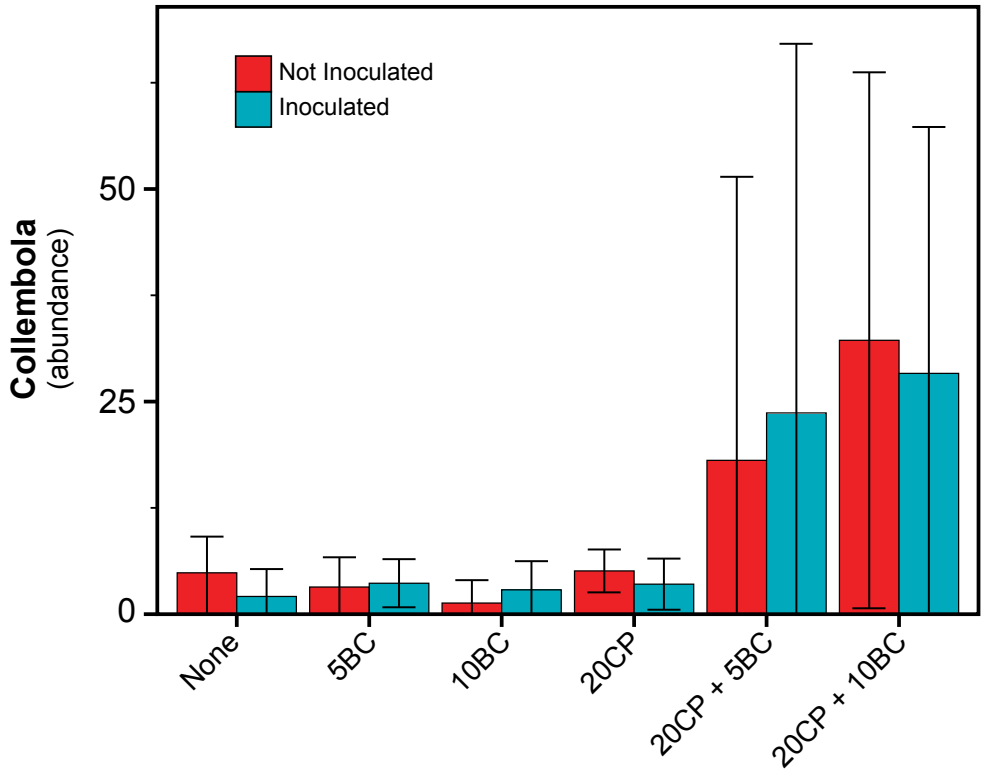
(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	0.400	0.837	0.633	n.s.
5BC	-0.844	1.205	0.484	n.s.
10BC	-10.130	10.204	0.321	n.s.
20CP	1.301	1.114	0.243	n.s.
20CP + 5BC	2.390	1.034	0.021	*
20CP + 10BC	1.620	1.067	0.129	n.s.
AM Inoculation	-5.700	2.703	0.035	(*)
Interactions				
5BC × AM Inoculation	5.741	3.069	0.061	.
10BC × AM Inoculation	17.930	10.559	0.090	.
20CP × AM Inoculation	5.938	2.917	0.042	*
20CP + 5BC × AM Inoculation	5.960	2.931	0.042	*
20CP + 10BC × AM Inoculation	7.006	2.860	0.014	*
AM Inoculation × Plot Height	0.134	0.062	0.030	*
5BC × AM Inoculation × Plot Height	-0.163	0.070	0.021	(*)
10BC × AM Inoculation × Plot Height	-0.331	0.192	0.086	(.)
20CP × AM Inoculation × Plot Height	-0.134	0.071	0.059	(.)
20CP + 5BC × AM Inoculation × Plot Height	-0.165	0.067	0.014	(*)
20CP + 10BC × AM Inoculation × Plot Height	-0.184	0.068	0.007	(**)

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Significantly different intercepts with negative values in parentheses

Figure 4.3: Fungivorous nematode abundance collected during the second season (September 2012). Generalized linear models with a negative binomial distribution link function were used to test the treatment effects. Panel (a) represents raw data ± 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model estimates represent the expected change from the intercept.

(a)



(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	1.591	0.456	<0.001	***
5BC	0.071	0.478	0.882	n.s.
10BC	-0.560	0.492	0.256	n.s.
20CP	-0.105	0.477	0.825	n.s.
20CP + 5BC	1.562	0.469	<0.001	***
20CP + 10BC	2.151	0.461	<0.001	***
AM Inoculation	-1.071	0.530	0.043	(*)
Interactions				
AM Inoculation × Plot Height	0.029	0.014	0.032	*

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Significantly different intercepts with negative values in parentheses

Figure 4.4: Collembola abundance collected during the second growing season of the prairie restoration (September 2012). Generalized linear models with a negative binomial distribution link function were used to test the treatment effects. Panel (a) represents raw data ± 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model term estimates represent the expected change from the model intercept (i.e. control plots).

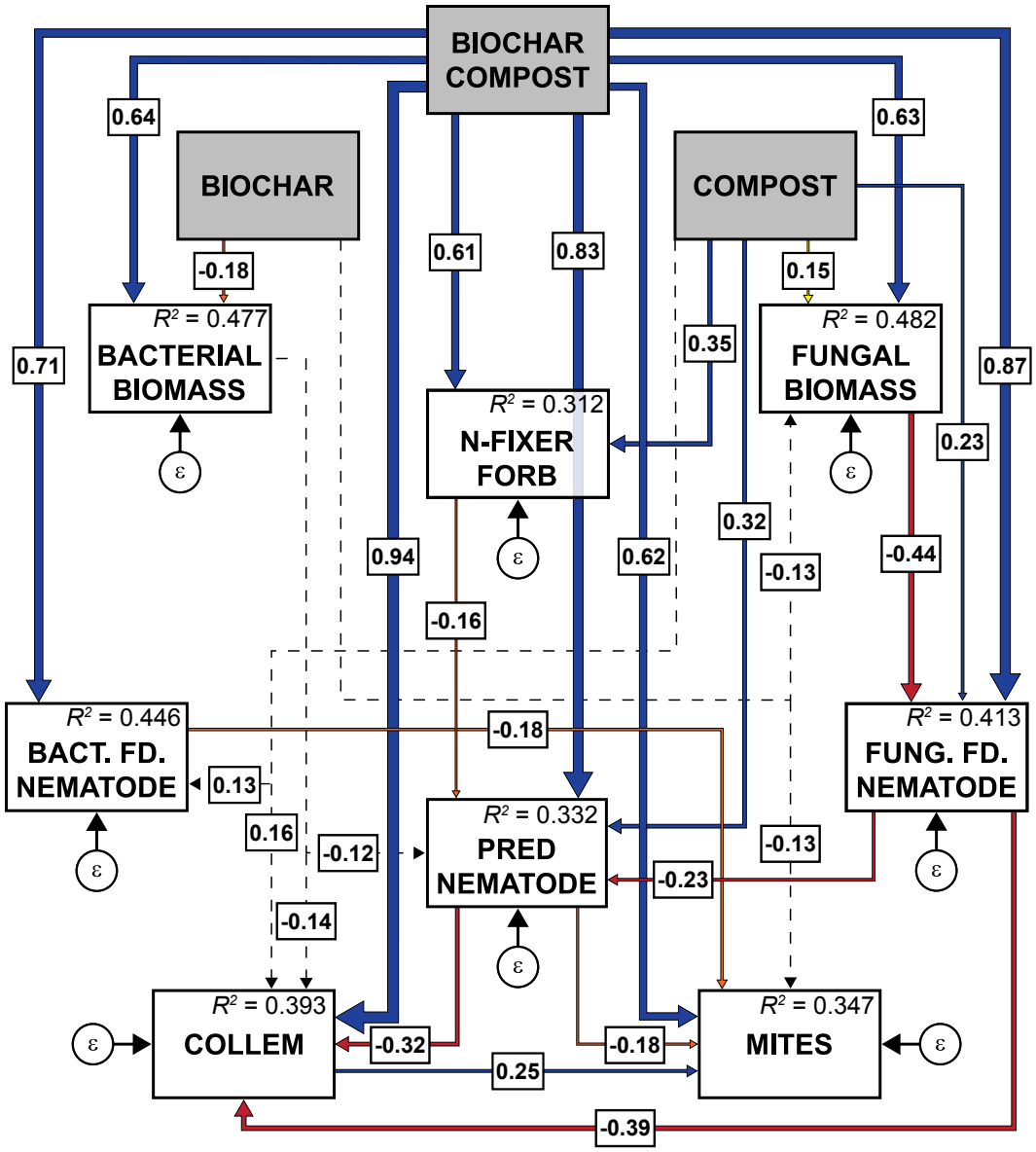


Figure 4.5: Structural equation soil food web model for the plant plug experiment. Exogenous variables are displayed in shaded gray boxes. Endogenous variables are displayed in white boxes. The residual error associated with each endogenous variable is displayed as (ε). Structural equation model line weights are scaled to the direct pathway standardized regression estimates given in each boxes. Blue (positive) and red (negative) arrows indicate significant standardized regression estimates ($p < 0.05$). Yellow (positive) and orange (negative) arrows indicate trends in standardized regression estimates ($0.05 < p < 0.1$). Dashed lines are non–significant paths with standardized regression estimates > 0.1. Regression estimates < 0.1 are not included to simplify the data presentation. A full description of direct, indirect, and total model estimates are given in Table 4.1 – 4.3. Squared multiple correlations are reported within endogenous variable boxes. Squared multiple correlations were calculated for each endogenous variable to determine explained variance.

Table 4.1: Direct, indirect, and total standardized regression estimates of soil amendments on the soil community and N–fixing forbs generated by the structural equation model. Significant direct pathway estimates are given in bold text ($p < 0.05$).

	Observed Effects(λ)		
Predictor → Response	**Direct**	**Indirect**	**Total**
Amendments → Microbes			
BC → bacterial biomass	-0.18	-0.01	-0.18
BC → fungal biomass	-0.13	0.00	-0.12
CP → bacterial biomass	0.15	-0.03	0.12
CP → fungal biomass	0.15	0.01	0.17
BC + CP → bacterial biomass	**0.64**	-0.04	0.60
BC + CP → fungal biomass	**0.63**	0.02	0.65
Amendments → N–fixing forbs			
BC → N–fixing forbs	0.07	0.00	0.07
CP → N–fixing forbs	**0.35**	0.00	0.35
BC + CP → N–fixing forbs	**0.61**	0.00	0.61
Amendments → Nematodes			
BC → bact. feeding nematodes	-0.07	0.02	-0.05
BC → fungal feeding nematodes	-0.05	0.06	0.01
BC → predatory nematodes	-0.02	0.01	-0.02
CP → bact. feeding nematodes	0.13	-0.01	0.12
CP → fungal feeding nematodes	**0.23**	-0.07	0.16
CP → predatory nematodes	**0.32**	-0.11	0.21
BC + CP → bact. feeding nematodes	**0.71**	-0.05	0.66
BC + CP → fungal feeding nematodes	**0.87**	-0.28	0.59
BC + CP → predatory nematodes	**0.83**	-0.30	0.54
Amendments → Microarthropods			
BC → Collembola	-0.04	0.02	-0.02
BC → mites	-0.13	0.04	-0.10
CP → Collembola	0.16	-0.14	0.02
CP → mites	0.07	-0.08	-0.01
BC + CP → Collembola	**0.94**	-0.45	0.49
BC + CP → mites	**0.62**	-0.19	0.43

When applied as a solitary amendment, both biochar rates negatively influenced bacterial biomass compared to non–amended control plots (5 T ha^{-1} biochar, p=0.010; 10 T ha^{-1} biochar, p=0.015)(Figure 4.2a). Fungal biomass was not significantly influenced by either biochar rate in this study (5 T ha^{-1} biochar, p=0.441; 10 T ha^{-1} biochar, p=0.119)(Figure 4.6a). Across all soil amendment treatments, the only significant increase in fungal:bacterial ratios compared to non–amended controls occurred at the 5 T ha^{-1} biochar amendment rate (p=0.002)(Figure 4.7a).

Biochar's influence on soil nematode abundance was variable, ranging from a neutral to significantly negative responses. Bacterial feeding nematode abundance was significantly reduced in the 10 T ha^{-1} biochar application rate (p=0.005) but unaffected in 5 T ha^{-1} biochar treatments when compared to non–amended controls (Figure 4.8a). Fungal feeding nematode abundance was not directly influenced by biochar (5 T ha^{-1} biochar, p=0.484; 10 T ha^{-1} biochar, p=0.321)(Figure 4.3a). No predatory nematodes were detected in either biochar amendment rates (Figure 4.9a).

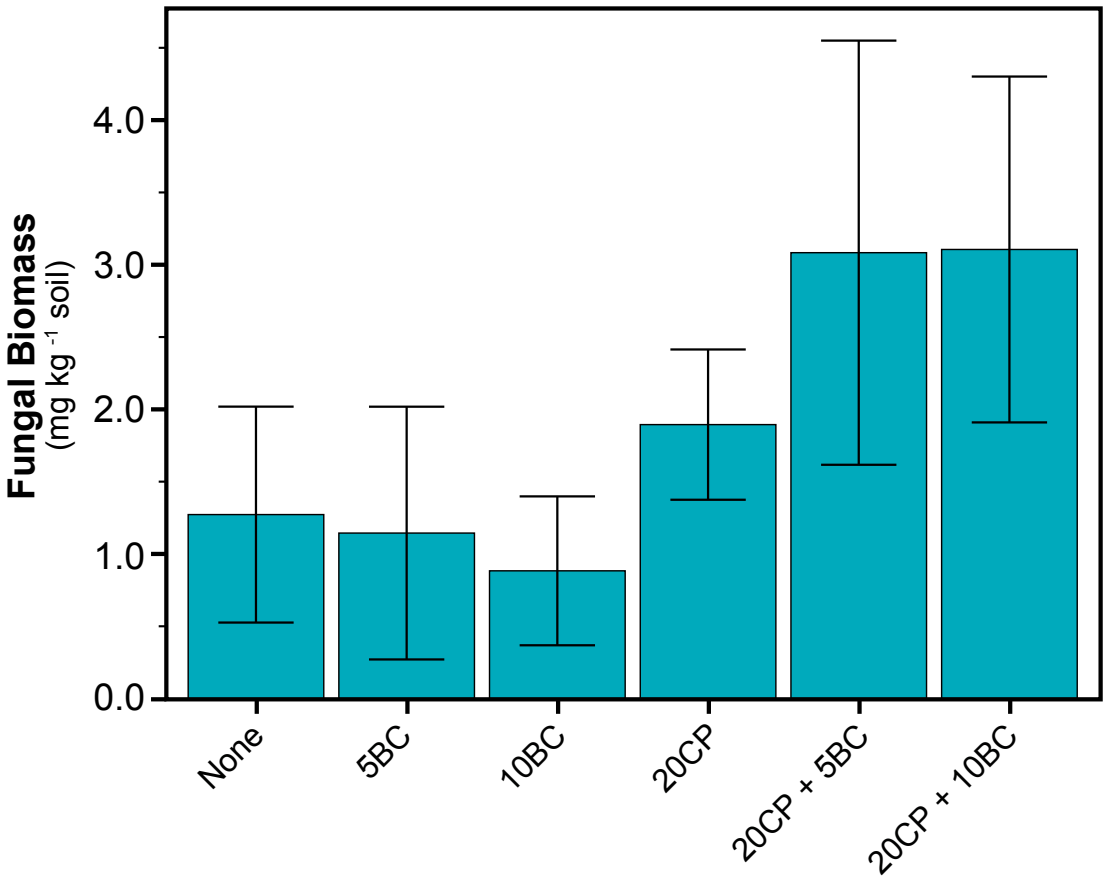
Direct effects of biochar on the soil arthropods was not predictable after generalized linear model analysis. Oribatid mite abundance was significantly reduced in plots with both biochar rates (5 T ha^{-1} biochar, p=0.014; 10 T ha^{-1} biochar, p=0.047) while Collembola abundance (5 T ha^{-1} biochar, p=0.882; 10 T ha^{-1} biochar, p=0.256) and predatory mite abundance (5 T ha^{-1} biochar, p=0.146; 10 T ha^{-1} biochar, p=0.363) was unaffected by biochar application rate(Figures 4.4a, 4.10a, & 4.11a).

Bacterial biomass tended to be reduced in the presence of biochar in the structural equation model (direct pathway coefficient = -0.18, p=0.059) (Figure 4.12). Soil fungal biomass (p=0.171) and soil animal abundance (bacterial feeding nematodes (p=0.497); fungal feeding nematodes (p=0.648); predatory nematodes (p=0.827); Collembola (p=0.723); mites (p=0.219)) remained largely unaffected by the introduction of biochar in the post–mine sandpit. Although not significant, all direct pathway coefficients were negative for soil microbial biomass and soil animal abundance relationships suggesting a negative impact on the soil food web compared to control plots (See Table 4.1 for direct pathway values).

The influence of compost on soil biota

The effect of 20 T ha^{-1} compost on the soil food web was largely neutral in this post–mine restoration trial compared to non–amended plots. Only fungal biomass was significantly increased by compost addition in the study (p=0.010) (Figure 4.6a). Bacterial biomass (p=0.149) and fungal:bacterial ratios (p=0.822) were not influenced by the addition of compost. Among soil animals, the abundance of predatory nematodes was significantly increased by 20 T ha^{-1} of compost (p<0.001). All other soil animals had no significant response to the compost amendment compared to non–amended controls (bacterial feeding

(a)

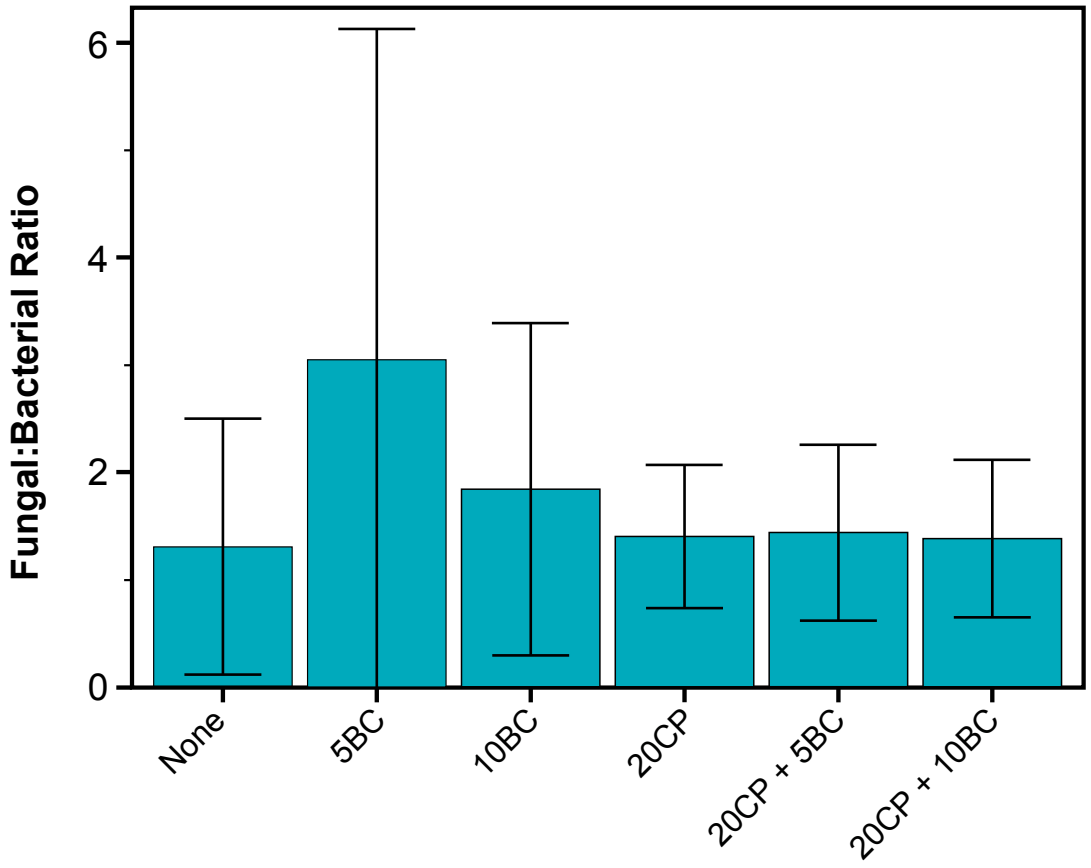


(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	0.333	0.033	<0.001	***
5BC	-0.036	0.046	0.441	n.s.
10BC	-0.073	0.046	0.119	n.s.
20CP	0.122	0.046	0.010	**
20CP + 5BC	0.252	0.046	<0.001	***
20CP + 10BC	0.261	0.046	<0.001	***
Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**				

Figure 4.6: Fungal biomass collected during the second growing season of the prairie restoration (September 2012). Data were analyzed with linear models to test treatment–level effects. Panel (a) represents raw data ± 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model term estimates represent the expected change from the model intercept (i.e. control plots).

(a)

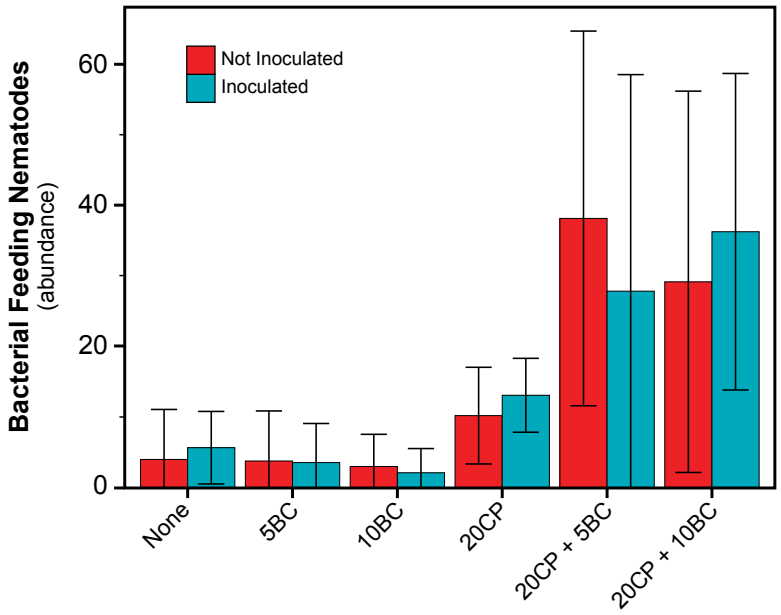


(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	0.267	0.226	0.238	n.s.
5BC	0.847	0.274	0.002	**
10BC	0.343	0.299	0.251	n.s.
20CP	0.069	0.311	0.822	n.s.
20CP + 5BC	0.094	0.309	0.760	n.s.
20CP + 10BC	0.055	0.311	0.860	n.s.
Significance: * $\leq$ 0.001 | ** $\leq$ 0.010 | * $\leq$ 0.050 | . $\leq$ 0.100**				

Figure 4.7: Fungal:bacterial biomass ratios collected during the second growing season of the prairie restoration (September 2012). Generalized linear models with a negative binomial distribution link function were used to test the treatment effects. Panel (a) represents raw data $\pm$ 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model term estimates represent the expected change from the model intercept (i.e. control plots).

(a)



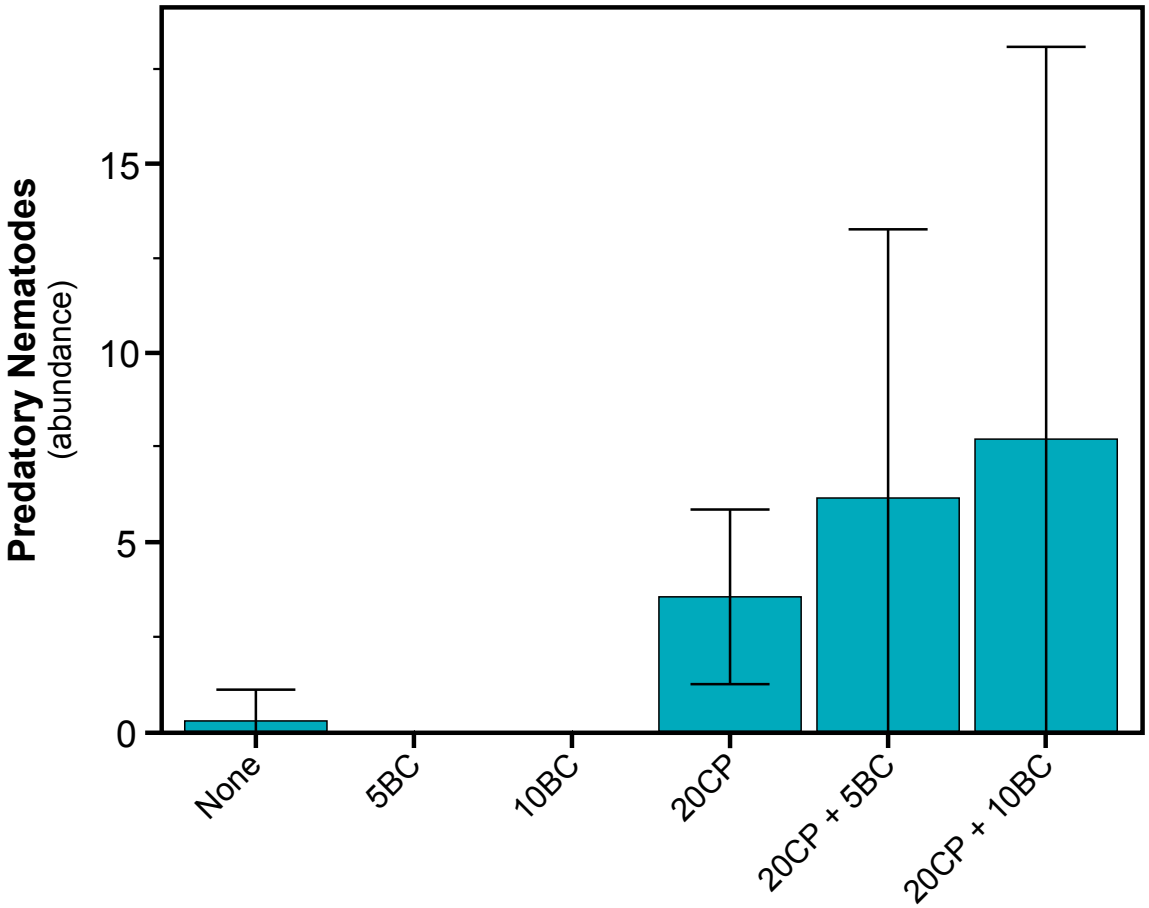
(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	2.448	0.712	0.001	***
5BC	-1.380	0.952	0.147	n.s.
10BC	-3.412	1.210	0.005	(**)
20CP	-0.316	0.960	0.742	n.s.
20CP + 5BC	1.814	0.893	0.042	*
20CP + 10BC	0.684	0.915	0.046	*
Plot Height (Dry → Wet)	-0.052	0.025	0.040	(*)
Interactions				
10BC × AM Inoculation	4.642	1.668	0.005	**
5BC × Plot Height	0.058	0.029	0.046	*
10BC × Plot Height	0.100	0.034	0.003	**
20CP × Plot Height	0.060	0.036	0.096	.
20CP + 10BC × Plot Height	0.060	0.0305	0.049	*
AM Inoculation × Plot Height	0.071	0.033	0.030	*
10BC × AM Inoculation × Plot Height	-0.169	0.047	<0.001	(***)
20CP × AM Inoculation × Plot Height	-0.076	0.045	0.090	(.)
20CP + 10BC × AM Inoculation × Plot Height	-0.093	0.041	0.025	(*)

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Significantly different intercepts with negative values in parentheses

Figure 4.8: Bacteriovorus nematode abundance collected during the second growing season of the prairie restoration (September 2012). Generalized linear models with a negative binomial distribution link function were used to test the treatment effects. Panel (a) represents raw data ± 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model term estimates represent the expected change from the model intercept (i.e. control plots).

(a)

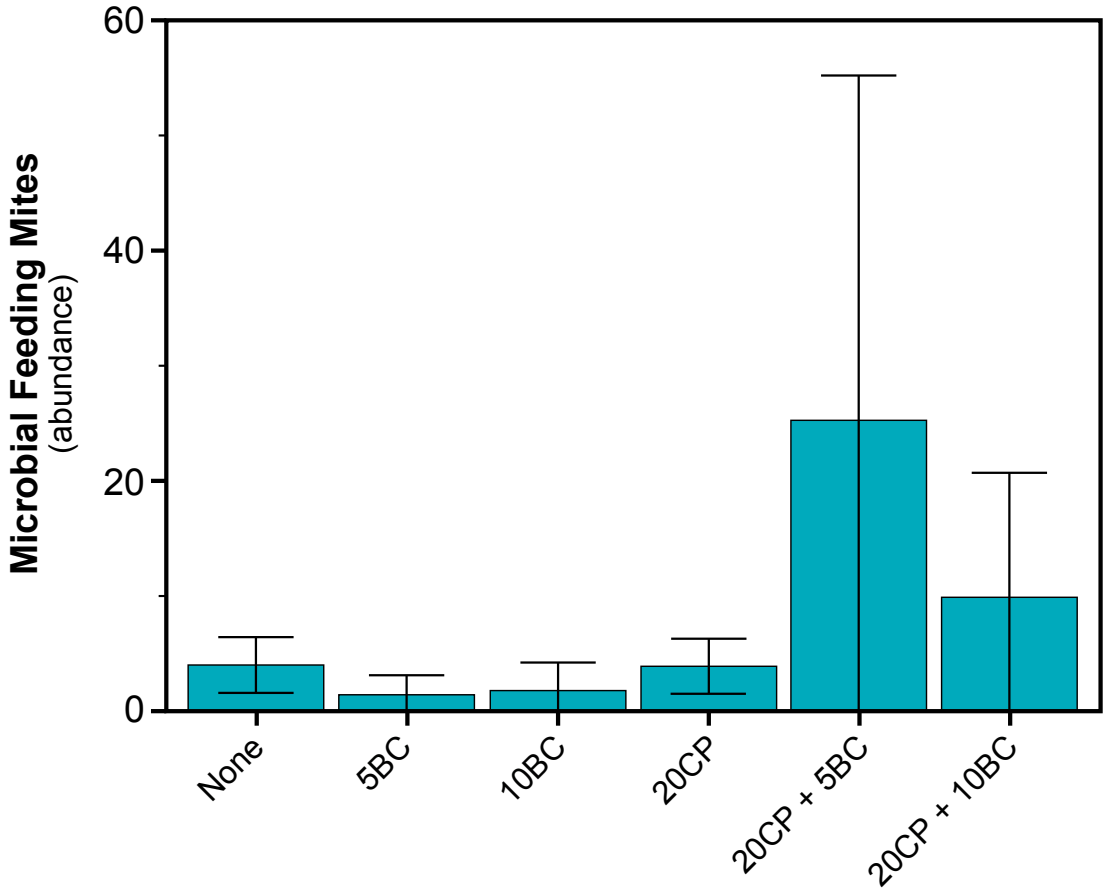


(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	-2.140	0.749	0.004	**
5BC	-	-	-	-
10BC	-	-	-	-
20CP	3.401	0.796	<0.001	***
20CP + 5BC	3.959	0.792	<0.001	***
20CP + 10BC	4.184	0.791	<0.001	***
Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**				

Figure 4.9: Predatory nematode abundance collected during the second growing season of the prairie restoration (September 2012). Generalized linear models with a negative binomial distribution link function were used to test the treatment effects. Panel (a) represents raw data ± 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model term estimates represent the expected change from the model intercept (i.e. control plots).

(a)



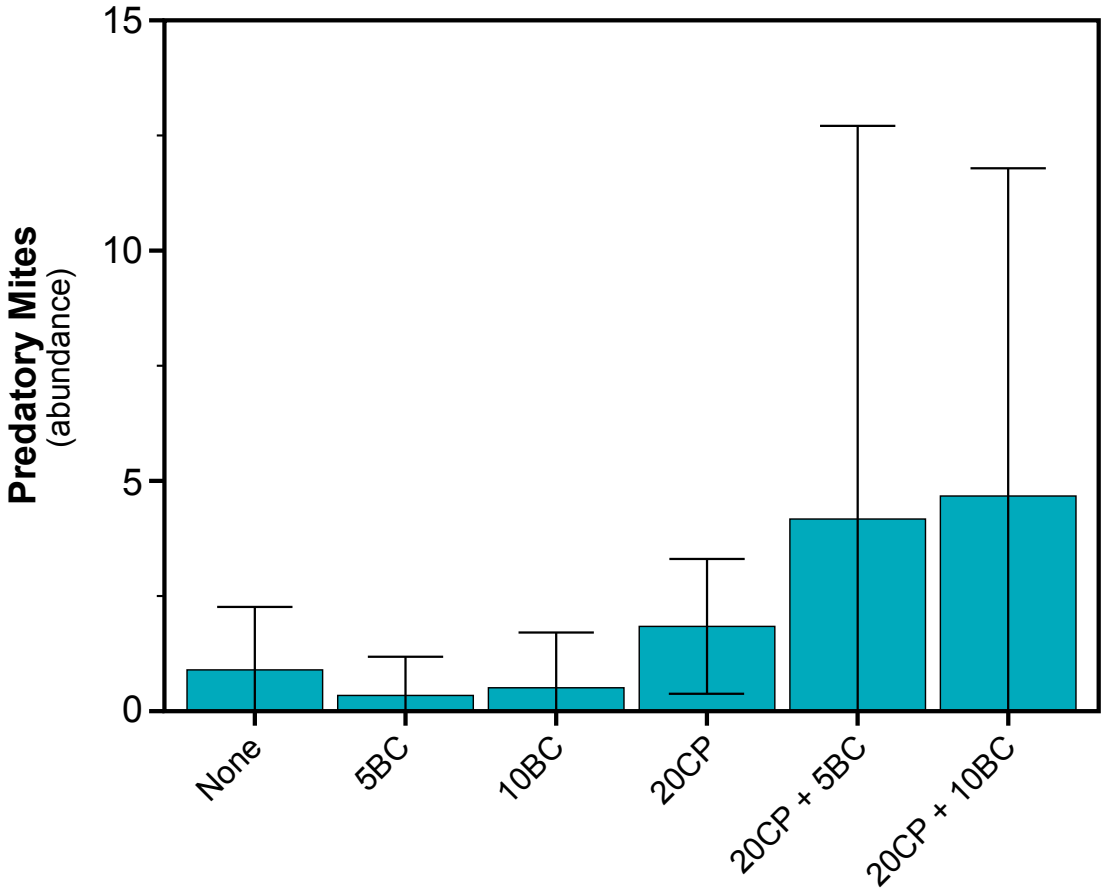
(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	1.386	0.273	<0.001	***
5BC	-1.041	0.424	0.014	(*)
10BC	-0.811	0.407	0.047	(*)
20CP	-0.028	0.386	0.942	n.s
20CP + 5BC	1.843	0.370	<0.001	***
20CP + 10BC	0.905	0.375	0.016	*

Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**
Note: Significantly different intercepts with negative values in parentheses

Figure 4.10: Oribatid mite abundance collected during the second growing season of the prairie restoration (September 2012). Generalized linear models with a negative binomial distribution link function were used to test the treatment effects. Panel (a) represents raw data ± 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model term estimates represent the expected change from the model intercept (i.e. control plots).

(a)



(b)

Model Terms	Estimate	SE	p–value	sig. level
Intercept	-0.118	0.419	0.778	n.s.
5BC	-0.981	0.674	0.146	n.s.
10BC	-0.575	0.632	0.363	n.s.
20CP	-0.724	0.564	0.893	n.s.
20CP + 5BC	1.545	0.549	0.005	**
20CP + 10BC	1.658	0.548	0.002	**
Significance: * ≤ 0.001 | ** ≤ 0.010 | * ≤ 0.050 | . ≤ 0.100**				

Figure 4.11: Predatory mite abundance collected during the second growing season of the prairie restoration (September 2012). Generalized linear models with a negative binomial distribution link function were used to test the treatment effects. Panel (a) represents raw data ± 1 SD; n = 9. x–axis: None = no soil amendment, 5BC = 5 T ha^{-1} biochar, 10BC = 10 T ha^{-1} biochar, 20CP = 20 T ha^{-1} compost, 5BC +20CP = 5 T ha^{-1} biochar + 20 T ha^{-1} compost, 10BC +20CP = 10 T ha^{-1} biochar + 20 T ha^{-1} compost. Significant main effect terms and interactions shown in (b). Model term estimates represent the expected change from the model intercept (i.e. control plots).

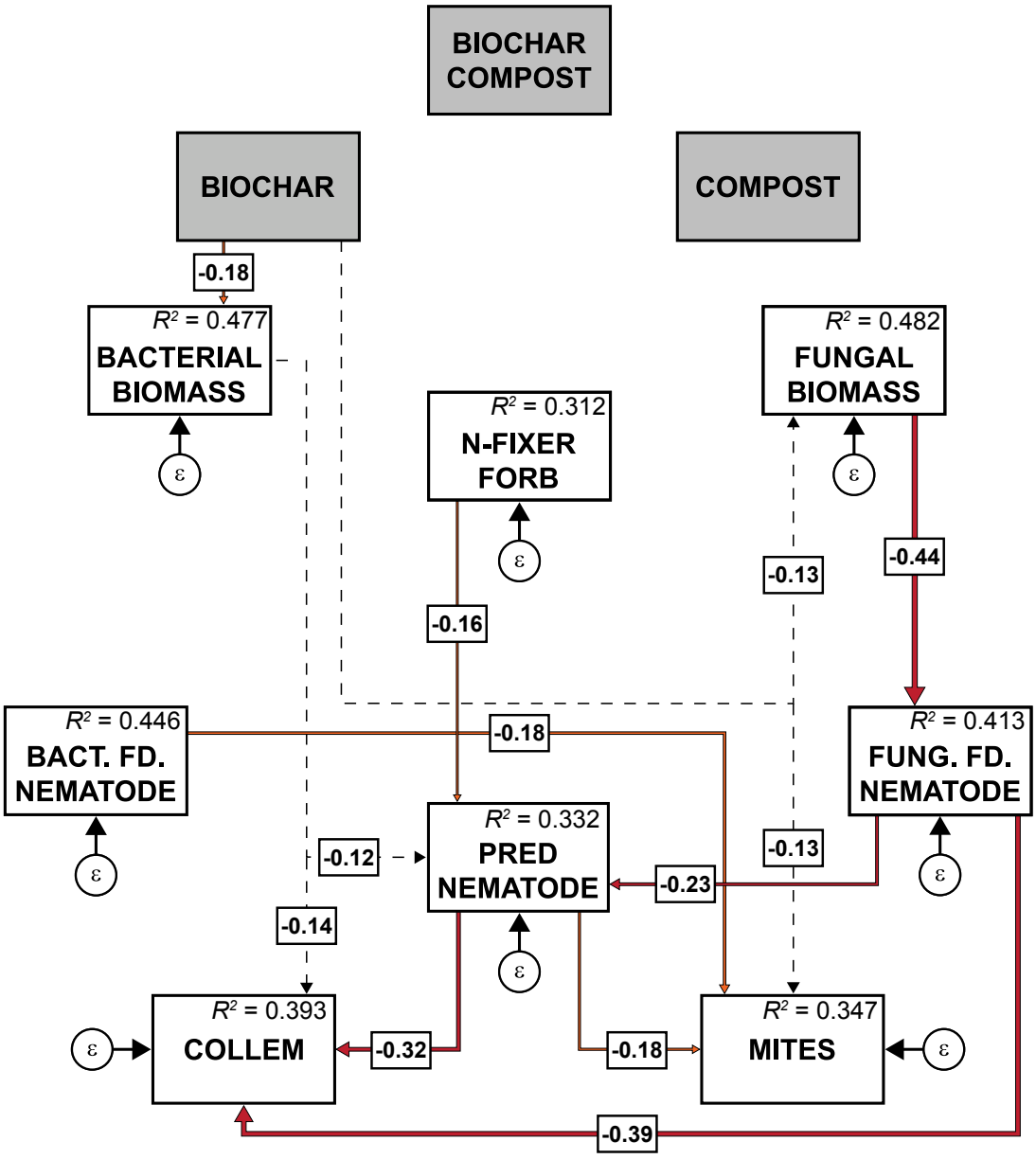


Figure 4.12: Negative standardized regression estimates in the soil food web model for the grassland restoration plant plug experiment Exogenous variables are displayed in shaded gray boxes. Endogenous variables are displayed in white boxes. Structural equation model line weights are scaled to the direct pathway standardized regression estimates given in each boxes. Red arrows indicate significant standardized regression estimates ($p < 0.05$) and orange arrows indicate trends in standardized regression estimates ($0.05 < p < 0.1$). Dashed lines are non–significant paths with standardized regression estimates > 0.1. Regression estimates < 0.1 are not included to simplify the data presentation. A full description of direct, indirect, and total model estimates are given in Table 4.1 – 4.3. Squared multiple correlations are reported within endogenous variable boxes. Squared multiple correlations were calculated for each endogenous variable to determine explained variance.

nematodes, p=0.742; fungal feeding nematodes, p=0.243; Collembola, p=0.825; Oribatid mites, p=0.942; predatory mites, p=0.893).

As a solitary amendment, compost had a significant positive influence on the soil food web for several biotic variables in the structural equation model (direct pathway coefficients: N–fixing forb biomass = 0.35 (p<0.001); predatory nematode abundance = 0.32 (p=0.004); fungal feeding nematode abundance = 0.23 (p=0.026))(Figure 4.13 & Table 4.1). A positive trend in increasing fungal biomass was detected in plots with compost only addition (direct pathway coefficients = 0.15 (p=0.089)).

Synergistic effects of biochar, compost, and AM fungal inoculation

The synergistic interaction of compost and biochar had a large positive effect on the soil microbial community, but no such interaction occurred with AM fungal inoculation and soil amendments. Regardless of biochar rate, compost + biochar significantly increased bacterial and fungal biomass (p<0.001) compared to non–amended controls (Figures 4.2a & 4.6a). Biochar + compost significantly increased all biotic variables (i.e. soil microbial community biomass, soil animal abundance, and N–fixing forb biomass) in the soil food web structural equation model (all p–values<0.001)(Figure 4.5 & Table 4.1). No significant change in fungal:bacterial ratios was detected in the compost + biochar treatments compared to non–amended controls.

Soil animal abundance responded positively in plots with added compost + biochar. Bacterial feeding nematode abundance increased significantly in both compost + biochar treatments ($20\,\mathrm{T\,ha^{-1}}$ compost + $5\,\mathrm{T\,ha^{-1}}$ biochar, p=0.042; $20\,\mathrm{T\,ha^{-1}}$ compost + $10\,\mathrm{T\,ha^{-1}}$ biochar, p=0.046) while fungal feeding nematode abundance was only significantly increased in the $20\,\mathrm{T\,ha^{-1}}$ compost + $5\,\mathrm{T\,ha^{-1}}$ biochar treatment (p=0.021) and no significant response in the $20\,\mathrm{T\,ha^{-1}}$ compost + $10\,\mathrm{T\,ha^{-1}}$ biochar treatment (p=0.129) (Figures 4.3a & 4.8a). Predatory nematode abundance was significantly increased in both compost + biochar treatments compared to controls (p<0.001)(Figure 4.9a). A consistent positive response in fungal feeding nematode abundance was also detected when adding AM fungal inoculum in conjunction with all amendment rates (Figure 4.3a).

Collembola abundance ($20\,\mathrm{T\,ha^{-1}}$ compost + $5\,\mathrm{T\,ha^{-1}}$ biochar, p<0.001; $20\,\mathrm{T\,ha^{-1}}$ compost + $10\,\mathrm{T\,ha^{-1}}$ biochar, p<0.001), Oribatid mite abundance ($20\,\mathrm{T\,ha^{-1}}$ compost + $5\,\mathrm{T\,ha^{-1}}$ biochar, p<0.001; $20\,\mathrm{T\,ha^{-1}}$ compost + $10\,\mathrm{T\,ha^{-1}}$ biochar, p=0.016), and predatory mite abundance ($20\,\mathrm{T\,ha^{-1}}$ compost + $5\,\mathrm{T\,ha^{-1}}$ biochar, p<0.005; $20\,\mathrm{T\,ha^{-1}}$ compost + $10\,\mathrm{T\,ha^{-1}}$ biochar, p=0.002) responded positively to both compost + biochar treatments (Figures 4.4a, 4.10a, & 4.11a).

The standardized direct pathway coefficients in biochar + compost treatments were demonstrably larger (range of direct pathway coefficients: N–fixing forb biomass (0.61) –

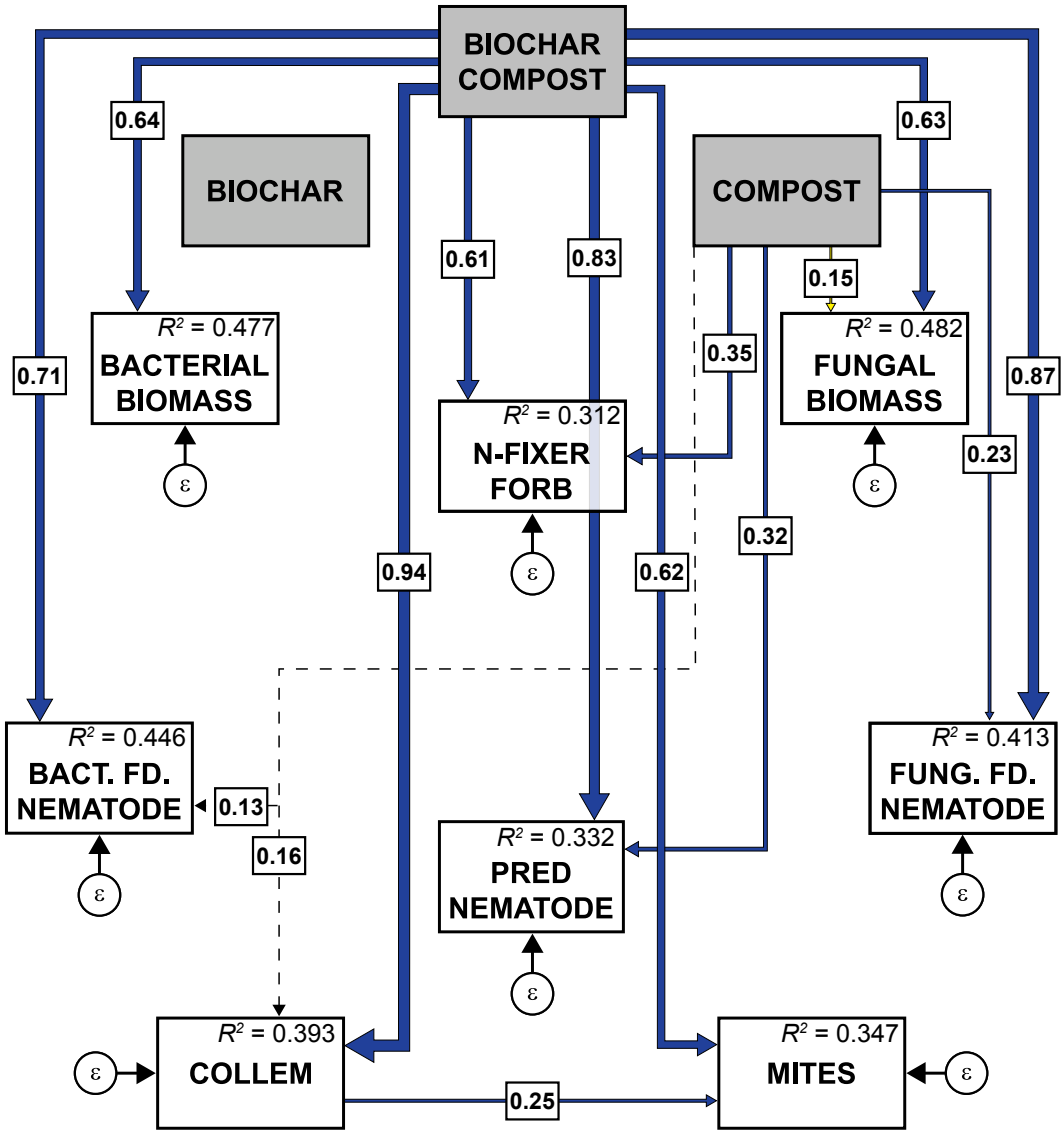


Figure 4.13: Positive standardized regression estimates in the soil food web model for the grassland restoration plant plug experiment. Exogenous variables are displayed in shaded gray boxes. Endogenous variables are displayed in white boxes. The residual error associated with each endogenous variable is displayed as (ε). Structural equation model line weights are scaled to the direct pathway standardized regression estimates given in each boxes. Blue arrows indicate significant standardized regression estimates ($p < 0.05$) and yellow arrows indicate trends in standardized regression estimates ($0.05 < p < 0.1$). Dashed lines are non–significant paths with standardized regression estimates > 0.1. Regression estimates < 0.1 are not included to simplify the data presentation. A full description of direct, indirect, and total model estimates are given in Table 4.1 – 4.3. Squared multiple correlations are reported within endogenous variable boxes. Squared multiple correlations were calculated for each endogenous variable to determine explained variance.

Collembola abundance (0.94)) compared to biochar only and compost only plots (Table 4.1). In the biochar + compost treatments, large indirect effects were detected due to negative interactions among the soil animals (Tables 4.1 & 4.3). Thus, the biochar + compost treatments resulted in reduced total effects in the soil food web model compared to direct effects (range of total pathway coefficients: mite abundance (0.43) – bacterial feeding nematode abundance (0.66)). Comparatively, the total effects of the biochar + compost treatments are consistently more influential on all soil biota than compost alone (range of total pathway coefficients: mite abundance (-0.01) – N–fixing forb biomass (0.35)) and biochar alone (range of total pathway coefficients: bacterial biomass (-0.18) – N–fixing forb biomass (0.07))(Table 4.1).

The influence of plot height on soil biota

Relative plot height had no significant direct effect on the majority of soil food web biota. Only bacterial feeding nematode abundance had a direct negative response to decreasing height of plots on the landscape. The interaction of soil amendments, AM fungal inoculation, and decreasing plot height indicated a consistent reduction in fungal feeding nematode abundance (Figure 4.3a). Collembola abundance increased significantly with plot height and the inoculated plots (p=0.043).

The influence of plant functional group biomass on the soil food web

Model #2 containing the C_4 grass and composite forb biomass variables had an acceptable fit to the proposed *a priori* hypotheses (χ^2=17.276, df=11, p=0.100; CFI=0.940; AIC=177.3). Compared to Model #3 (AIC=129.9), dropping the C_4 grass and composite forb biomass variables greatly improved model performance Therefore, these variables were dropped from the final structural equation model. N–fixing biomass was significantly increased as a result of compost addition (direct pathway coefficient: 0.35, p<0.001) and biochar + compost addition (direct pathway coefficient: 0.61, p<0.001). Although soil amendment rate influenced N–fixing biomass, no significant direct effect of N–fixing biomass was observed on soil microbial biomass or soil animal abundance (Table 4.2).

The influence of bacterial and fungal biomass on soil animal abundance

No significant covariance effect was detected between bacterial biomass and fungal biomass (direct pathway coefficient: 0.00, p=0.950)(Table 4.2). Bacterial biomass did not significantly influence the abundance of soil animals in the study (bacterial feeding nematodes (p=0.393), predatory nematodes (p=0.292); Collembola (p=0.197); mites (p=0.476))(Table 4.2). Comparatively, fungal biomass significantly reduced fungal feed-

Table 4.2: Direct, indirect, and total standardized regression estimates of soil microbes and N–fixing forbs on the soil community generated by the structural equation model. Significant direct pathway estimates are given in bold text ($p < 0.05$).

	Observed Effects(λ)		
Predictor → Response	**Direct**	**Indirect**	**Total**
N–fixing forbs → Soil Community			
N–fixing forbs → bacterial biomass	-0.07	0.00	-0.07
N–fixing forbs → fungal biomass	0.04	0.00	0.04
N–fixing forbs → bact. feeding nematodes	0.00	0.01	0.01
N–fixing forbs → fungal feeding nematodes	0.01	-0.02	0.00
N–fixing forbs → predatory nematodes	-0.16	0.01	-0.15
N–fixing forbs → Collembola	0.00	0.06	0.06
N–fixing forbs → mites	0.00	0.04	0.04
Bacteria → Soil Community			
bacterial biomass → fungal biomass	0.00	0.00	0.00
bacterial biomass → bact. feeding nematodes	-0.09	0.00	-0.09
bacterial biomass → predatory nematodes	-0.12	0.00	-0.12
bacterial biomass → Collembola	-0.14	0.04	-0.10
bacterial biomass → mites	-0.08	0.02	-0.06
Fungi → Soil Community			
fungal biomass → bacterial biomass	0.00	0.00	0.00
fungal biomass → fungal feeding nematodes	**-0.44**	0.00	-0.44
fungal biomass → predatory nematodes	-0.01	0.10	0.10
fungal biomass → Collembola	0.04	0.14	0.18
fungal biomass → mites	-0.08	0.01	-0.07

Table 4.3: Direct, indirect, and total standardized regression estimates among the soil animals generated by the structural equation model. Significant direct pathway estimates are given in bold text ($p < 0.05$).

	Observed Effects(λ)		
Predictor → Response	**Direct**	**Indirect**	**Total**
Grazing nematodes → Predatory nematodes			
bacterial feeding nematodes → predatory nematodes	0.02	0.00	0.02
fungal feeding nematodes → predatory nematodes	**-0.23**	0.00	-0.23
Grazing nematodes → Microarthropods			
bacterial feeding nematodes → Collembola	0.02	-0.01	0.01
bacterial feeding nematodes → mites	-0.21	0.00	-0.21
fungal feeding nematodes → Collembola	**-0.39**	0.07	-0.32
fungal feeding nematodes → mites	0.04	-0.04	0.00
Predatory nematodes → Microarthropods			
predatory nematodes → Collembola	**-0.32**	0.00	-0.32
predatory nematodes → mites	-0.18	-0.08	-0.26
Collembola → Mites			
Collembola → mites	**0.25**	0.00	0.25

ing nematode abundance (direct pathway coefficient: -0.44, p<0.001), but did not significantly influence any other soil animal group (predatory nematodes (p=0.962); Collembola (p=0.724); mites (p=0.507)) (Table 4.2).

Soil animal functional feeding group interactions

Significant negative direct effects were observed in the proposed feeding hierarchy based on my *a priori* hypotheses (Figure 4.12 & Table 4.3). Fungal feeding nematode abundance was negatively correlated with predatory nematode abundance (direct pathway coefficient: -0.23, p=0.033) and Collembola abundance (direct pathway coefficient: -0.39, p<0.001). In addition, predatory nematode abundance had a significant negative effect on Collembola abundance (direct pathway coefficient: -0.32, p<0.001). Negative trends were detected between bacterial feeding nematodes → mites (direct pathway coefficient: -0.18, p=0.055) and predatory nematodes → mites (direct pathway coefficient: -0.18, p=0.076). The only significant, positive direct path coefficient detected in soil animal interactions was Collembola → mites (direct pathway coefficient: 0.25, p=0.015).

4.4 Discussion

After two growing seasons, this study showed that soil amendments can significantly affect soil microbial biomass and soil animal abundance in post–mine sandpits. I found that AM fungal inoculation of plant plug roots had little influence on soil food web structure compared to non–inoculated controls. As a solitary amendment, biochar had a largely negative effect on the soil food web although this effect was not always significant. Conversely, biochar mixed with compost promoted large significant increases in the soil microbial community and soil animal abundance. Thus, I clearly show that soil food web development is highly dependent upon amendment choice during grassland restoration in the degraded sandpit.

4.4.1 Soil food web response to AM fungal inoculation

N–fixing plant biomass, soil microbial biomass, and soil animal abundance were unaffected by AM fungal inoculation. The structural equation model clearly showed that AM fungal inoculation did not have sufficient explanatory power to describe any direct effect influences on soil microbial biomass and soil animal abundance. Thus, my hypothesis that AM fungal inoculation of plant plugs would increase the response of microbial community biomass and soil animal abundance in post–mine sandpits was not supported.

Arbuscular mycorrhizal fungi are key components of the soil microbiota and interact with other microorganisms in the rhizosphere (Bowen and Rovira 1999). AM fungi can influence belowground soil food webs by increasing plant root biomass through nutrient acquisition, therefore increasing litter inputs (Langley and Hungate 2003). Jastrow et al. (1998) determined strong positive direct and indirect effects by AM fungi on fine roots, microorganisms, and soil aggregation using path analysis. Increased microbial biomass and rhizosphere activity can subsequently support soil nematodes and microarthropods through bottom–up cascading mechanisms (Scherber et al. 2010).

One reason for no effect may be due to the presence of a background AM fungal community in non–inoculated plant plug roots at the time of planting due to unsterilzed greenhouse conditions in the nursery. Rowe et al. (2007) has suggested that locally collected field inoculum is more effective than commercial inoculum for establishing late-successional species. This background community of local AM fungi may have formed a strong partnership with the plants used in this restoration, thus contributing to a positive plant biomass response in the field. In desertified semi–arid systems, the establishment of the shrub with local vs. non–native AM fungal inoculum showed increases in soil enzyme activity compared to controls but no difference between inoculum source (Alguacil et al. 2005). Therefore, the influence of the AM fungal inoculum on the development of the soil food web may have

gone undetected because of a beneficial belowground plant response by the background AM fungal community in the non–inoculated plots. Thus, the commercial AM fungal inoculum, *R. irregularis*, was not an effective land management tool to increase total plant biomass, plant functional group biomass, or soil food web biomass and abundance in this study.

4.4.2 Soil food web response to biochar

Biochar had neutral to negative influence soil microbial biomass and soil animal abundance in this study. My hypothesis that both biochar amendments would positively increase soil microbial biomass, soil animal abundance, and N–fixing plant abundance due to ameliorative effects in sandpit substrate was not supported.

Biochar application as a land management tool has been proposed to assist soil recovery in severely degraded systems (Blackwell et al. 2009). To date, most mine reclamation studies using biochar have investigated soil chemical properties under laboratory conditions (i.e. pH, cation exchange capacity, heavy metal sequestration) (Fellet et al. 2011; 2014, Kelly et al. 2014). When investigated, the response of soil microbes to biochar in the literature are mixed. Kelly et al. (2014) found that microbial biomass was not altered by biochar amendments in mine tailings while a meta–analysis of plant and microbial biomass by Biederman and Harpole (2012) found that biochar addition increased aboveground productivity, crop yield, soil microbial biomass, and favorable tissue macro–nutrients across all soil types and climates. Graber et al. (2010) suggested that shifts in soil microbial activity were indirect and arose from biochar stimulating plant growth, thus inducing a plant exudate effect in the rhizosphere. My study suggests that this mechanism is unlikely as plant biomass decreases were not detected by the structural equation model yet decreases in the soil microbial community were detected.

In my study, biochar may have introduced nutrient stress associated with post–mine sandpits, reducing microbial biomass and soil animal abundance. One potential mechanism for reduced biotic response in reduced soil microbial abundance is biochar's high cation exchange capacity strongly adhering limited nutrients in post–mine soils (Steiner et al. 2007, Xu et al. 2013). Ultimately, feedstock source and pyrolysis time determines nutrient leaching rates, chemical properties, and hydrophobicity of biochar in soils, thus dictating a soil biotic response (Singh et al. 2010, Kinney et al. 2012). Furthermore, biochar's hydrophobic nature may have repelled soil moisture, causing negative trends in soil microbial biomass (Kinney et al. 2012). Complex biogeochemical interactions will ultimately determine resource availability for biotic communities and may need to be optimized for soil conditions. The long–term ecological effect of biochar application needs to be investigated in terms of soil community development and plant response under field conditions.

Reductions in soil animal abundance were also detected when applying biochar although not significant. This is most likely due to reduced fungal and bacterial biomass creating a limiting microbial food resource for grazing soil animals. Mikola and Setälä (1998) established a trophic dynamic microcosm experiment testing the soil interactions among microbes, microbivorous nematodes, and predatory nematodes. This study suggested that increased microbial productivity leads to the increased biomass of microbes followed by a lagging response time in the microbivorous nematode trophic level. As nematode recovery after severe disturbance is slow (Bongers and Ferris 1999), the reduced abundance of grazing soil animals in biochar only plots most likely contributed to the minimized abundance of predatory soil animals. To date, biochar's influence on soil food web structure is relatively unknown (Lehmann et al. 2011). My study results are contrary to the hypotheses proposed by McCormack et al. (2013) where the addition of biochar was anticipated to increase microbial and soil animal resource availability. The direct influence of biochar on the multi–trophic interactions warrants further study in restoration and agriculture. Biochar's use as a land management tool to assist soil food web development is questionable as shown by my results.

4.4.3 Soil food web response to compost

Compost application did not affect most of the organisms in my study. While it increased the abundance of N–fixing plant biomass, fungal biomass, and predatory nematode abundance, I could not detect an influence of compost in other groups. This is contrary to my hypothesis which predicted that compost would increase soil microbial biomass and soil animal abundance due to ameliorated soil conditions.

Compost was expected to promote increased soil microbial community growth due to improved nutrient and water retention profiles in compost amended soils (Bastida et al. 2008, Larney and Angers 2012). Long–term and short–term studies indicate that urban compost primes microbial community decomposition and increases plant–available macro–nutrients in agricultural soils (Weber et al. 2007, Hadas and Portnoy 1997). Jones et al. (2010) concluded that compost additions in bauxite–processing residue sand positively influenced water retention and nutrient profiles, thus increasing soil microbial activity. The influence of organic amendments has been shown to favor the growth of fungal community compared to bacteria communities (Jastrow et al. 2007). Fungi are more efficient decomposers of compost amendments compared to bacteria due to large hyphal networks and efficient nutrient acquisition and translocation mechanisms (Lucas et al. 2014). My study confirms that fungal biomass was significantly increased by the addition of compost compared to a positive, but non–significant influence on bacterial biomass.

An important indicator in re–establishing a soil microbial community is the relative proportions of bacterial and fungal biomass (Bardgett and McAlister 1999) with natural grassland systems being dominated by fungal communities (Harris 2009). Mummey et al. (2002) suggested that fungal:bacterial ratios in mine spoils can approach ratios in natural soils after 20 years following restoration although total biomass is comparatively reduced. Compost addition in my study did not significantly influence fungal:bacterial biomass ratios compared to non–amended controls. This is surprising as the compost amendment resulted in significant increases in fungal biomass but did not significantly influence bacterial biomass. This result suggests that soil conditions improved due to the compost amendment but gains in fungal biomass were not pronounced compared to gains in bacterial biomass. As fungal:bacterial biomass ratios were only measured after two growing seasons, these ratios are expected to increase as above– and belowground litter inputs accumulate, ultimately favoring fungal dominance in the soil food web (Holtkamp et al. 2008).

I had anticipated that compost would have a larger influence on soil food web structure due to the alleviation abiotic stress in impoverished soils, increased microbial food resources, and more feeding substrate. The addition of farm composts containing crop residue and manure showed increased fungal feeding nematodes in a soil incubation study due to increased food resources (Steel et al. 2012). Jørgensen and Hedlund (2013) showed that Collembola and predatory mites had increased fecundity when adding a fungal inoculated clover amendment to soils, highlighting the importance of fungal biomass for grazing animal fecundity and prey attraction.

After two growing seasons, my study showed that no direct influence of compost was detected on soil animal abundance. This is surprising as compost significantly increased fungal biomass, a food source for fungal feeding nematodes, Collembola, and Oribatid mites. As Collembola and Oribatid mites also consume litter, the addition of composted plant material had little influence on population densities. Therefore, predatory nematodes and mites subsequently had low population densities most likely attributed to low prey abundances. Several studies have linked rates of organic matter mineralization to microbial production and the biomass of soil microbivores and predators (Seastedt 1984, Bardgett et al. 1998, Laakso et al. 2000).

As increases in resource availability drives microbial production (Baer et al. 2003), the chosen compost rate may have been too low to overcome the harsh abiotic sandpit conditions, resulting in a subdued response in the soil food web. Thus, I speculate that largest obstacle for soil food web development is the harsh conditions in post–mine areas. Increasing the amount of organic matter may have a stronger influence on the development of the soil food web by further ameliorating soil conditions and increasing food resources for grazing soil animals.

4.4.4 Soil food web response to compost and biochar

Soil food web responses were much more pronounced when compost and biochar were applied together. My hypothesis that co–amending soils with compost + biochar and AM fungal inoculation would positively influence soil microbial biomass and soil animal abundance was partially supported in my study. The compost + biochar treatments had a large positive influence on soil food web development but an effect of AM fungal inoculation was not detected.

This is the first study to investigate the influence of compost and biochar on soil food web structure. As suggested by Fischer and Glaser (2012), a synergistic interaction of compost + biochar can positively influence soil conditions, leading to a large positive impact on soil food web structure. Potential mechanisms may be increased water and nutrient retention, buffered pH, or creation of microbial refugia in biochar's highly porous structure with a large nutrient pulse supplied by compost. Fischer and Glaser (2012) also indicate that compost may charge biochar's surfaces to slowly release nutrients to soils, increase soil aeration, and reduce leaching losses. These soil amelioration mechanisms may have overcome the abiotic conditions associated with post–mine sandpits, significantly increasing soil microbial biomass followed by soil animal abundance compared to solitary amendments or controls.

Food resources are key to the development of multi–trophic belowground food webs (Hunt et al. 1987). Increased knowledge regarding the linkages within decomposer food webs requires an understanding of the importance of resource availability upon the growth and abundance microbial communities and the associated consumer trophic levels (Wardle 2006). Adding compost with biochar clearly influenced bacterial and fungal biomass by creating a bottom–up trophic cascade effect within the soil food web. By co–mixing amendments, large increases in soil microbial communities translated to increased abundance of grazing nematodes, Collembola, Oribatid mites. Increased abundance of grazing soil animals created a prey resource for predatory nematodes and mites leading to a more complete trophic hierarchy in the soil food web. Thus, the soil environment with compost + biochar was more tolerable for the growth and development of a belowground soil microbial communities, inducing a positive response in soil nematode and microarthropod abundance.

Compost + biochar treatments did not influence fungal:bacterial biomass ratios when compared to control plots even though substantial increases in fungal and bacterial biomass were detected. The addition of organic amendments to severely degraded areas drives positive changes in microbial activity as estimated by soil microbial biomass carbon (Ros et al. 2003). This suggests that an influence of increased resource availability may have equally benefited the growth of both fungal and bacterial functional groups. As biochar persists in soils for over 100+ years (Lehmann et al. 2009), the growth benefits gained by

soil microbial communities and belowground plant response from biochar's physiochemical properties are anticipated to be long–term (Glaser et al. 2002).

My study results clearly confirm that the development of soil microbial communities and soil animals is significantly enhanced when biochar and compost are used in tandem. When used as a land management tool, co–amending soils with compost + biochar can accelerate the development of soil food webs, indicating that post–mine substrate recovery is substantial when compared to non–amended controls. As the soil food web is developed, functioning decomposition and nutrient cycles can translate into greater plant response in the field due to the ecosystem services provided by multi–trophic interactions of soil biota (de Vries et al. 2013). Thus, when a researcher approaches restoration with a holistic ecosystem perspective, the incorporation of compost + biochar is an essential to soil regeneration in degraded mine areas.

4.4.5 Interactions among soil microbial biomass and soil animal abundance

This study showed that increasing fungal biomass significantly reduced fungal feeding nematode abundance while increasing Collembola abundance. Bacterial biomass had no significant effect on bacterial feeding nematode and Collembola grazers. These relationships in the fungal and bacterial energy pathways channels did not coincide with my hypothesis that increased food resources in the soil microbial community would correlate to higher grazing soil animal densities.

As described previously, amendment choice had a direct influence on microbial biomass and soil animal density with compost + biochar amendments significantly increasing all soil animals in the belowground food web. These results indicate that the soil edaphic conditions improved the growth of food resources for grazing and predatory animals. Conversely, when investigating direct relationships in the fungal and bacterial energy channels, unexpected patterns emerged. Increasing fungal biomass resulted in a significant negative correlation with fungal feeding nematodes while no significant correlation was detected between bacterial biomass and bacterial feeding nematodes.

Bacterial and fungal nematode response to soil amendments is not equivocal due to different life history strategies in these groups (Ferris and Bongers 2006). Generally, bacterial feeding nematodes have short life cycles and high reproductive potential to quickly respond to bacterial blooms in soils (Bongers and Ferris 1999). Fungal feeding nematodes, on the other hand, are longer–lived and reproduce more slowly compared to bacterial feeding nematodes, thus are less likely to respond to changing conditions (Ferris et al. 2001). Disturbance severity will ultimately dictate the composition of bacterial vs. fungal feeding nematodes with severe disturbance shifting soil systems towards bacterial feeding energy

channels (Bongers 1990). Thus, in the fungal pathway, the negative correlation between fungal biomass and fungal feeding nematodes may be due to a lag in rapid nematode response to the available fungal food resource. Bacterial colonizers present in the sandpit may have a more ephemeral response to bacterial biomass leading to no direct correlation between these trophic levels. Long–term monitoring of the grazing nematodes populations is needed to determine the recovery trajectory of this system.

Furthermore, Yeates et al. (1993) points out that nematode feeding group identification may not be well delineated in practice. Bacterial feeding nematodes are generally classified by having a wide mouth, but bacterial feeders have been known to feed upon fungal food resources (Gupta et al. 1979). Fungal feeding nematodes are classified as possessing a stylet but some genera with this feature are known to feed on plant roots or be vertebrate predators (Bongers and Bongers 1998). The large nematode family, Tylenchidae, is commonly considered root feeding nematodes but have been shown to feed on fungal food resources (Okada et al. 2005). Distinguishing between plant feeding nematodes and fungal feeding nematodes is important to understand interactions between food resources. In my study, nematodes classified as fungal feeding nematodes may feed upon plant roots in the rhizosphere, leading to unexpected correlations between fungal biomass and fungal feeding nematodes by inaccurately attributing food resource being consumed.

Fungal biomass had a positive total effect on predatory nematode abundance. Conversely, bacterial biomass had a negative total effect on predatory nematode abundance. Wardle et al. (1995) showed that top predatory nematodes were regulated by microbial biomass while fungal and bacterial feeding nematode responses were more variable in the belowground interactions. Li et al. (2014) found that organic enrichment in an agricultural setting shifted grazing fungal dominance to the fungal energy channel while increasing predatory nematodes. In my study, the ephemeral life cycle response of bacterial feeding nematodes may not be a stable food source for predatory nematodes. Thus, the longer–lived fungal feeding nematodes in the fungal energy channel may be a more nutritious food source with a more stable population to support nematode predators.

Compost + biochar soil amendments showed large increases in soil microarthropods abundance compared to controls. This indicates that a food resource is available to support these soil animals in the food web. When investigating the relationship in the structural equation model, Collembola responded positively to increases in fungal biomass while a negative response was detected for bacterial biomass although no strong relationship exists. Soil mites had no strong response to the microbial community. Results from field studies in microarthropod populations are often ambiguous as invertebrate population size can increase or decrease with amendment type (Bardgett and Cook 1998, Jørgensen and Hedlund 2013). Complex trophic interactions present in soil food webs occur as a number

of direct and indirect interactions occur between species is difficult to predict (Bengtsson et al. 1996). More research needs to be conducted on the reliability of fungal and bacterial biomass measurements when determining resource availability for grazing soil animals and predators.

Sandpit resource extraction strongly diminishes populations of soil animals in the system. Slow recovery may be expected as the soil environment develops over time and response of microarthropods can be highly variable due to environmental heterogeneity (Curry and Good 1992, Menta 2012). As evident in control plots, soil animal abundances were extremely low after two growing seasons. In former agricultural lands, colonization of new areas is unpredictable and responds differently to successional changes in plant communities (Scheu and Schulz 1996, Korthals et al. 2001). Most soil organisms are considered to have limited abilities of overcoming soil heterogeneity and have restricted movement (Ojala and Huhta 2001). In my study system, compost + biochar amended plots ultimately improved soil conditions to support a higher abundance of soil microbes and soil animals compared to control plots. Thus, as soil organisms disperse to the field site, these amended plots are anticipated to better support the survival of newly arriving immigrants compared to unamended plots.

To accelerate recovery of soil food webs, post–mine sandpits may benefit from an inoculation of soil food web biota. Assuming soil animals survive the composting process, it is reasonable to conclude that compost would act as soil food web inoculant that contained a high abundance of bacteria, fungi, nematodes, and soil microarthropods (Cernova 1970, Streit et al. 1985, Steel et al. 2013b). After two growing seasons, compost additions alone unexpectedly did not alter soil animal abundance significantly as shown by marginally improved soil animal densities compared to non–amended plots. Conversely, improved biogeochemical conditions of compost + biochar amended soils may have allowed for soil animal survival associated with compost. Further research on the potential of using compost as a soil food web inoculum should be conducted under various restoration scenarios.

4.4.6 Summary

Based on my results, the recovery of soil food webs when restoring grasslands in post–mine aggregate sites is a viable management option in southern Ontario. As ecosystem productivity and soil fertility are closely tied to soil biota, land managers should target the development of soil food webs in tandem with phyto–centric goals to maximize plant production in a restoration project. As shown in this study, mining sand strongly reduces soil microbial communities and soil animal abundance even after two years of habitat recovery with grassland plant plugs. The harsh substrate conditions in non–amended control plots suggests that the recovery time of a soil food web would be slow if no management

action was implemented. Therefore, land management tools are necessary to accelerate the development of functioning soil food webs in severely disturbed habitats.

In my study, the technical reclamation tools (i.e. compost, biochar, and AM fungal inoculation) induced a variable response within the soil food web. The application of biochar alone added stress to the post–mine substrate and further restricted soil food web development in the field while compost had a negligible effect compared to control plots. In contrast, co–amending soils with compost + biochar led to large increases in soil food web development in the field. The application of compost + biochar in an industrial–scale restoration project should promote increases in soil microbial biomass and soil animal production leading by improving soil conditions at the site. Increasing the function of soil food webs can ultimately drive aboveground plant community production due to the ecosystem services provided. These ecosystem services can lead to reduced site maintenance costs, increase plant community recovery time, and promote vegetative biodiversity.

Chapter 5

Management Recommendations for Grassland Restoration in Post–Extraction Sandpits

Restoring a grassland plant community is challenging when attempting to recreate natural habitat in post–mine sandpits. Native plant growth in sandpits is hampered by stressful abiotic conditions and disrupted connections among plants–microbes–soil animals attributed to severe disturbance and low organic matter. As shown in this study, only C_4 grasses and N–fixing forbs responded positively in sandpit substrate during the plant plug and seed application trials. Composite forbs and C_3 grasses exhibited poor plant response in the study, regardless of treatment. Thus, recreating highly diverse prairie ecosystems remains a challenge, even after addressing the harsh conditions of sandpit substrate using soil amendments and AM fungal inoculum.

My results show that soil amelioration can benefit plant response when restoring grassland vegetation as plugs or seeds. When directly seeding in sandpits, significant increases in plant response were achieved by concurrently amending soils with high rates of biochar and compost ($20\,\mathrm{T\,ha^{-1}}$ and $40\,\mathrm{T\,ha^{-1}}$) and the recommended rate of the commercial AM fungal inoculum, *Rhizophagus irregularis*. In the plant plug trial, no significant differences in total plant biomass were detected in plots adding $20\,\mathrm{T\,ha^{-1}}$ of compost + $10\,\mathrm{T\,ha^{-1}}$ of biochar although a positive trend was indicated. In this case, the rate of $20\,\mathrm{T\,ha^{-1}}$ of compost + $10\,\mathrm{T\,ha^{-1}}$ of biochar may have been too low to create a strong positive plant growth response in the plug experiment. But, in terms of soil food web development, the large biotic response of fungi, bacteria, and soil animals to $20\,\mathrm{T\,ha^{-1}}$ of compost + $10\,\mathrm{T\,ha^{-1}}$ of biochar in the soil food web indicates more favorable conditions for soil microbial and animal growth after two growing seasons. Therefore, the addition of high rates of biochar and compost improved soil conditions to accelerate soil food web development and increase plant response in the field compared to non–amended and non–inoculated control plots.

5.1 Plant species selection and sourcing

Incorporating plant material as seed or plugs is essential when recreating prairie–like habitat in post–mine sandpits to reach plant community targets. When left to natural plant recolonization, control areas (i.e. no plant plugs added) were sporadically colonized by weedy, ephemeral plants with low biomass. Native seed recruitment was minimal in these control areas even with the incorporation of soil amendments in non–vegetated plots. Therefore, incorporating native plant material as seed or plant plugs is essential when restoring grassland habitat in post–mine aggregate sites.

The decision to rehabilitate prairies with native plant seeds or plugs will be determined by desired speed of recovery and future maintenance considerations. Seeding the landscape incorporates drawback such as:

- slower and less successful plant establishment
- possible increased time to achieve rehabilitation certification
- increased site maintenance requirements (i.e. reseeding applications)
- increased influence of weedy, invasive plant species (i.e. herbicide applications may be necessary)

The upfront cost of sowing native plant plugs with soil amendments is initially more cost prohibitive than direct seeding (Table 5.1). The advantage of restoring with plant plugs is high plant biomass production by the C_4 grass and N–fixing forbs compared to seed growth over a similar growing period. Accelerated growth rates in the plant plug trial can increase soil stabilization by binding substrate with native plant roots and reducing wind scour. From personal observation, plant plug addition reduced surface erosion by wind energy immediately at the time of plug installation. Compared to the plant plug trial, plant growth in the seed application trial was stunted after three growing seasons compared to plants starting as plugs. Thus, integrating plug installation with native seeded may be a cost effective hybrid technique to minimize seed loss and stabilize the mine substrate at a restoration site.

This restoration project used locally–collected seed mixtures which were adapted to regional growing conditions. Locally–sourced plant material has been suggested to positively influence plant response in restoration projects with greatest success in soils experiencing lower disturbance (Lesica and Allendorf 1999, Buisson et al. 2006). In southern Ontario, high diversity seed mixes can range from 10 - 30 plant species to include a mixture of warm season (C_4) grasses, cool season (C_3) grasses, legumes (i.e. nitrogen–fixing forbs), and

composite wild flowers (Delaney et al. 2000). As shown by this study, C_4 grasses and N–fixing forbs will likely dominate the vegetative community in a sandpit restoration project after several years of plant growth. Further research on seed mixture proportions needs to be investigated to achieve the best results for creating a high diversity plant community in post–extraction sandpits.

This study shows that plant response to soil amendments and AM fungal inoculation varies among plant species. Therefore, plant selection must be considered on a case by case basis. The environmental tolerance of each plant species to amended post–mine substrate conditions favored some species, while being a detriment to another. All plant species, except for *Andropogon gerardii*, had a neutral to positive response to the compost or compost + biochar amendment addition. All plant species growing in biochar only treatments had a neutral to negative biomass response compared to control. *Panicum virgatum* and *Lespedeza capitata* biomass was significantly greater in AM inoculated compared to non–inoculated plants while the opposite was true for *Andropogon gerardii*. In general, the addition of compost and biochar amendments benefited the growth of most species in the post–mine substrate by alleviating abiotic stress. Conversely, plant response to the commercial AM inoculum was dependent upon species and planting method thus its use in tallgrass prairie restoration is context–dependent.

The composite forbs, *Symphyotrichum laeve* and *Liatris cylindracea*, did not perform well in either restoration trial. In the seed application trial, composite forb cover was negligible after three growing seasons. In the plug trial, significant reductions in composite forb biomass were exhibited in the field after two growing seasons. The chosen composite forb species in this trial were not ideal candidates for this post–mine restoration. In the following years, I expect that these species will be non–existent in the two trials.

In contrast, the C_4 grasses and N–fixing forbs tolerated the post–mine sandpit environment and were responsive to the applied soil amendments. These species were the largest contributor to total plant biomass in the field. Thus, sand pit restoration should include a mixture of these plant functional groups to increase plant community biomass.

Inconsistent plant species responses highlights the need to have clearly stated goals in restoration management plans when recreating prairies in severely disturbed areas. If total plant response is the key component to determine restoration success, then the application of biochar, compost, and AM inoculum would be an effective tool to assist the restoration of grassland plants in post–mine sandpits. If a practitioner is targeting a specific suite of plant species and managing for species at risk, higher caution must be used when choosing amendments and AM fungal inoculum. Target plants may be adversely affected by the addition of soil amendments in the field and/or choice of mycorrhizal inoculum. Context–dependent abiotic and biotic scenarios ultimately determine the success of each restoration.

5.1.1 Soil amendments and commercial AMF inoculum

Commercial AM fungal inoculum

The arbuscular mycorrhizal inoculum, *Rhizophagus irregularis*, was most effective during seed application when co–amended with high rates of compost and biochar. No significant effects on the total plant biomass and soil food web development were detected in the plant plug experiment. This may be a result of background AM fungal present in unsterilized plant plug soils in the commercial greenhouse.

As shown by other mine land reclamation studies, AM fungal inoculum benefits the growth of plants in severely degraded mine areas (Rao and Tak 2002, Rydlová et al. 2008, Wu et al. 2009) but plant response due to AM fungal inoculation can be enhanced by the addition of soil amendments (Gryndler et al. 2008, Püschel et al. 2008a). My research supports the finding that adding AM fungal inoculum enhances plant response when amended with compost in the seed trial. Co-amending soils with increasing rates of compost and biochar further facilitates the plant growth in this system. In this sandpit restoration, adding a commercial AM inoculum is appropriate when establishing grassland plants from seed when applying soil amendments. When growing plants from plugs in an unsterilzed greenhouse setting, my results show that land managers do not need to apply AM fungal inoculum if biochar and compost are not added as no significant plant response was detected in this trial after two growing seasons.

Biochar as a soil amendment

Biochar has been shown to be a beneficial land management tool to enhance plant production in tropical agriculture (Major et al. 2010). To date, no research has been conducted on the role of biochar in the restoration of grasslands in severely degraded post–mine areas. My results show that the incorporation of biochar as a solitary amendment for grassland restoration in sandpits should not be used. The negative responses detected in plant plug growth and soil food web development indicate that biochar addition further stresses the substrate and restricts the development of biota in the recovering system. Comparatively, high rates of biochar had no effect on total plant cover after three growing seasons in the direct seeding trial. Therefore, solitary biochar addition should be approached with caution as plant response can be hampered in a restoration project. More research needs to be conducted for the most appropriate restoration scenarios to add biochar with the goal of increasing plant response.

Compost as a soil amendment

Compost has been shown to increase plant response in degraded mine areas by alleviating stressful abiotic conditions in many studies (Noyd et al. 1996, Püschel et al. 2011). My study indicates that compost positively influences plant growth in the plant plug trial although not significant compared to non–amended controls. In the seed application trial, high levels of compost addition is the main driver of total plant cover. If a land manager is presented with an amendment choice, compost as a solitary amendment outweighs the use of biochar as a soil amendment. Incorporating $20\,\mathrm{T\,ha^{-1}}$ to $40\,\mathrm{T\,ha^{-1}}$ of compost into sandpit substrate has the largest potential to positively influence soil conditions as shown by increased plant response in grassland vegetation.

Synergism among biochar, compost, and AM fungal inoculum

The results of this study indicate that the concurrent addition of municipal compost, biochar and mycorrhizal inoculum are simple land management tools that improve plant performance and soil food web development in post–extraction aggregate sites. In the plant plug experiment, $20\,\mathrm{T\,ha^{-1}}$ of compost mixed with $10\,\mathrm{T\,ha^{-1}}$ of biochar had the highest positive effect on plant biomass, soil microbial biomass, and soil animal abundance. Thus, the amelioration of stressful abiotic conditions in the sandpit was achieved when compost and biochar were used together in the plant plug trial. AMF inoculation combined with high rates of compost ($20\,\mathrm{T\,ha^{-1}}$ to $40\,\mathrm{T\,ha^{-1}}$) and low rates of biochar ($20\,\mathrm{T\,ha^{-1}}$ to $40\,\mathrm{T\,ha^{-1}}$) resulted in the highest plant cover in the seed experiment.

The rates of biochar and compost need to be optimized to achieve the highest plant response at industrially feasible costs. My results suggest that low rates of biochar ($5\,\mathrm{T\,ha^{-1}}$ and $10\,\mathrm{T\,ha^{-1}}$) combined with higher rates of compost ($20\,\mathrm{T\,ha^{-1}}$ and $40\,\mathrm{T\,ha^{-1}}$) may achieve significant responses in the plant community while being cost effective. As the cost of biochar is substantially higher than that of municipal compost, adding $20\,\mathrm{T\,ha^{-1}}$ and $40\,\mathrm{T\,ha^{-1}}$ of biochar is not a cost effective amendment at this time (Table 5.1). Funds to restore grassland vegetation in a sandpit could be more effectively used by incorporating plant plugs with a high diversity seed mixture.

5.2 Purchasing soil amendments and inoculum for a restoration project

Rhizophagus irregularis can be purchased as a seed coat powder from Myke Pro® (www.usemykepro.com) and applied at the rate suggested by the manufacturer. The in-

oculum recommended for agricultural crops, Myke Pro® PS3, would be the most effective AMF inoculum for grassland restoration in sandpits.

Compost can be purchased locally at most landscape supply locations across Ontario. The approximate cost of compost is $40–$50 per metric ton plus delivery. Compost is typically generated from municipal waste collection streams and is readily available for purchase.

In comparison, the U.S. Biochar Initiative reports the cost of biochar as $500 per ton excluding shipping (http://biochar-us.org/, 2014). Currently, production facilities of biochar are not widespread, making large quantities of biochar less readily available to the land manager.

Although the cost and availability of biochar may be prohibitive in 2014, the soil conditioning effect of this amendment when co–amended the compost may become a viable option in the future. As carbon taxes are on the horizon, landholders may soon be able to generate offset carbon credits from activities that reduce emissions or sequester carbon, including biochar application. These offset carbon credits may defer the cost associated with biochar.

5.3 Site preparation

When preparing the pit floor substrate for a grassland restoration project, the area should be roughly graded flat to allow for ease of planting. Once graded, compost and biochar can be tilled into the upper 10 cm of sandpit substrate before planting occurs. I recommend minimizing the time between compost incorporation and planting to reduce the colonization of unwanted weedy plants. Seeds and/or plant plugs can be sown by hand or with machinery depending upon the scale of the project. Ideally, seeds should be compacted with a seed roller to ensure solid contact with the pit floor. I do not recommend reincorporating long–term storage stock piles into the site. A high density of weedy plants will have developed on the stock–piled topsoil and would potentially out compete the growth of seeded native vegetation.

5.4 Summary

My goal was to optimize cost and effectively establish a tallgrass prairie ecosystem. I suggest that integrating both planting approaches (i.e. plant plugs and seed) will be the most effective strategy for ecosystem establishment. I recommend incorporating $20\,\mathrm{T\,ha^{-1}}$ to $40\,\mathrm{T\,ha^{-1}}$ of compost into the substrate before planting and /or seeding the site. If available, co–amend sandpit substrate with low rates of biochar ($5\,\mathrm{T\,ha^{-1}}$ to $15\,\mathrm{T\,ha^{-1}}$).

Table 5.1: The projected materials cost of land rehabilitation in abandoned sandpits in southern Ontario. Two viable options are available for prairie system rehabilitation: seed addition or plug addition. Note that the cost per ha decreases as the rehabilitation area increases.

Approx. Cost to Establish One Hectare of Prairie Grasses	
Prairie Rehabilitation w/ Seed	
Seed Application / ha (no grading required)	$3,000
Miscellaneous Costs (Transportation, etc.)	$500
Total	$3,500
Prairie Rehabilitation w/ Plugs	
Plug Cost ($1.00 × 20,000 plants / ha [1 plant / 0.5 m^2]	$20,000
Miscellaneous Costs (Transportation, labour, etc.)	$2,750
Total	$22,750
Amendments	
AMF Inoculum (4 kg inoculum = 5.3 ha coverage)	$400
Compost [$45 / metric ton × 20 T ha^{-1}]	$900
Biochar [$500 / metric ton × 10 T ha^{-1}]	$5,000

Incorporate plant plugs composed of legumes and warm season grasses at a rate of one plug per square meter. These plants have a high survivorship and growth success at the site, which will maximize the cost effectiveness of plant plugs. Sow a high diversity plant seed mixture containing warm season grasses, cool season grasses, legumes, and wildflowers among the plant plugs. When planting, incorporating AM fungal inoculum can further promote vegetative establishment and growth. Incorporating all of the investigated amendments is an effective restoration strategy that compliments the desired outcome of grassland plant establishment and soil development.

Chapter 6

Conclusion

Severe land disturbance is pervasive among all ecosystems as a result of anthropogenic activities. As a society, we have a responsibility to repair the destruction that accompanies resource extraction and land–use change. Therefore, it is imperative that we restore viable ecosystems that support plant and animal communities in severely impacted sites to account for regional habitat loss. In this study, I show that the restoration of grassland plants in post–mine aggregate sites is a viable management option in southern Ontario. After resource depletion in mine areas, the substrate conditions that are a legacy of aggregate extraction are a stark contrast to functional soils in natural habitats. The edaphic conditions of sand extraction restricts plant growth and soil food web development as shown by the control plots at my research site. My goal was to recreate functional grassland habitat with ecological characteristics that resemble reference sites. Land management tools (i.e compost, biochar, and AM fungal inoculation) were anticipated to accelerate plant growth and soil recovery, translating into aboveground and belowground biota recovery on marginal lands.

A summary of the support garnered for the three main thesis objectives is addressed below:

Objective #1 *Develop a minimally destructive statistical method to increase measurement accuracy and reduce data collection time when estimating aboveground plant biomass.* The sampling method developed to estimate individual herbaceous plant and small shrub biomass in the field via partial least squares regression was superior to linear regression statistical techniques. Partial least squares regression was shown to be a robust statistical technique that should be used to accurately predict plant biomass in ecological experiments. In comparison to liner regression using a sole predictor variable, partial least squares regression increases prediction confidence and reliability in ecological experiments.

Objective #2 *Determine the multi–year plant response of both planting strategies to soil amendments and the commercial AM fungal isolate in a post-mine sandpit.* Incorporating land management tools to mitigate the harsh abiotic conditions of post–extraction substrate is necessary to increase plant production in target grass-

land communities. As solitary amendments, the incorporation of the commercial arbuscular mycorrhizal fungal isolate and biochar did not significantly improve plant growth in either trial. In both field trials, the addition of compost was the most influential driver of plant production in the post–mine sandpit. In the plant plug trial, a trend was detected when comparing the total plant biomass in plots adding $20\,\mathrm{T\,ha^{-1}}$ compost and $20\,\mathrm{T\,ha^{-1}}$ compost + $10\,\mathrm{T\,ha^{-1}}$ biochar compared to control. When compost was supplemented with biochar and the arbuscular mycorrhizal fungal inoculum (*Rhizophagus irregularis*) in the seed application trial, grassland plant response was largely accentuated. Thus, co–amending sandpit soils with biochar, compost, and mycorrhizal inoculum increased the effectiveness of the restoration protocol in this trial. A single application of high rates $20\,\mathrm{T\,ha^{-1}}$ of biochar and compost at the onset of an industrial–scale restoration project will lessen site maintenance costs, increase plant community recovery time.

Objective #3 *Determine the soil food web response to the addition of soil amendments and a AM fungal isolate in sandpit substrate.* As shown in this study, mining sand strongly reduces soil microbial communities and soil animal abundance even after two years of habitat recovery. Non–amended control plots had low soil food web abundance across all trophic levels. Thus, the natural recovery time of a soil food web would be slow if no management action is implemented. The application of biochar alone added stress to the post–mine substrate and further restricted soil food web development compared to control plots. In contrast, compost and arbuscular mycorrhizal inoculum had a negligible effect compared to control plots. Co–amending soils with compost + biochar led to large increases in soil food web development across all trophic levels indicating improved soil conditions at the site. Increasing the function of soil food webs ultimately drives aboveground plant community production due to the ecosystem services provided. These ecosystem services can lead to reduced site maintenance costs, increase plant community recovery time, and promote vegetative biodiversity.

The restoration of severely disturbed mine areas necessitates interventions that address stressful soil conditions (Séré et al. 2008). Restoration projects have successfully used municipal compost in mine areas with low organic matter content to promote plant growth (Noyd et al. 1996, Gryndler et al. 2008, Püschel et al. 2008b) and soil community development (Ros et al. 2003; 2006, Biederman et al. 2008). My research shows marginal improvement of plant growth and soil food web development due to compost in the plant plug trial. The $20\,\mathrm{T\,ha^{-1}}$ compost rate applied in the plant plug trial may have insufficient stimulate large production changes compared to controls.

Biochar has been shown to improve soil nutrient availability and retention, reduce soil acidity, and adsorb organic matter (Lehmann et al. 2003, Shrestha et al. 2010). As biochar research is sparse in restoration, the successful application of biochar to improve biotic response will ultimately depend upon soil type, source feedstock, pyrolysis conditions, and biochar application rates (Verheijen et al. 2014). Biochar as a soil amendment to increase plant growth has generated promising results within agricultural systems and greenhouse experiments. Research has demonstrated that biochar amended soils have greater crop biomass (Rondon et al. 2007, Major et al. 2010, Biederman and Harpole 2012) and enhanced biological N–fixation in leguminous crops (Rondon et al. 2007). In contrast to these studies, a plant growth effect was not detected in my grassland restoration experiment.

In the field of mine restoration, a dearth of information exists on the effect of biochar in facilitating the growth of plants, development of microbial communities, and soil fauna abundance (Lehmann et al. 2011). My research indicates that biochar as a solitary amendment results in no significant improvement in plant growth. Furthermore, biochar added stress to the system as shown by the reduction of microbial and soil fauna abundance in the plant plug trial. In the seed application trail, increasing rates of biochar were only beneficial for promoting plant cover when soils were co–amended with increasing compost and inoculation with *Rhizophagus irregularis*.

Biotic symbionts such as arbuscular mycorrhizal fungi have been used as inoculum to facilitate plant production in severely degraded habitats (Johnson 1998, Gryndler et al. 2008, Rydlová et al. 2008). The effectiveness of mycorrhizal inoculation on plant growth can vary by the combination of plant species, soil disturbance type, and the selection of an arbuscular mycorrhizal fungal isolate (Taheri and Bever 2010, Püschel et al. 2011, Thorne et al. 2013). My research indicates that the addition of a commercial arbuscular mycorrhizal fungal isolate, *Rhizophagus irregularis*, successfully established in the plant plug trial and persisted over the study period. No arbuscular mycorrhizal effect was detected on total plant biomass in the plant plug trial, but individual plants had a varied response to the addition of the inoculum. This indicates that plants did not respond equally to the addition of the isolate.

Background arbuscular mycorrhizal colonization was also detected in the non–inoculated control plants in the plant plug trial and persisted throughout the study period. Thus, the arbuscular mycorrhizal fungi present in the non–inoculated controls may have benefited growth and nutrient acquisition of plants in these plots. Furthermore, several studies suggest that arbuscular mycorrhizas should be collected directly from similar mine sites as these fungi will be better adapted to field conditions when restoring plants (Noyd et al. 1995, Taheri and Bever 2010). As I used a commercial inoculum in this study, the isolate used in my experiment may not have been well adapted to post–mine conditions in the field.

6.1 Strengths and limitations of the dissertation research

This dissertation research makes several strong contributions to the field of restoration ecology and ecological sampling methodology. The techniques developed in this study are applicable to mine restoration projects across the world. Severely degraded soil conditions are pervasive within the resource extraction industry. This research clearly shows that mine soil conditions need to be altered before the restoration of natural vegetation is attempted, especially when establishing plants from seed. Furthermore, the use of plant plugs in a research setting is a novel technique to restore grasslands in post–mine areas. Growing plants as plugs may have assisted the grassland plants in overcoming initial abiotic soil conditions compared to establishing plants from seed. Thus, the use of plant plugs was shown to be a reliable restoration technique, especially for C_4 grasses.

Selecting the appropriate suite of soil amendments carefully is essential when restoring post–mine sandpits. Land managers should be aware that all soil amendments and arbuscular mycorrhizal inoculum will not ubiquitously produce positive growth responses for plants, soil microbial communities, and soil animals. A single application of high rates ($20\,\mathrm{T\,ha^{-1}}$) of biochar and compost with arbuscular mycorrhizal fungal inoculum at the onset of an industrial–scale grassland restoration project can maximize plant response from seed and improve soil food web development. Thus, this translates to a restoration project that more closely approximates reference site conditions when compared to no management intervention.

During this research, I developed a new statistical technique to more accurately estimate herbaceous plant and small shrub biomass in the field. The improved measurement accuracy can reduce the error in plant biomass estimation and increase plant measurement experimental replication. The increased resolution in plant biomass estimation will prove to be an invaluable tool to experimentally measure plant growth in a variety of ecological scenarios. A limitation during the development of this technique was the low numbers of species measured in the field. Ideally, a large suite of plants would have been measured to test the robustness of this technique across many different types of vegetation in southern Ontario.

A strength of this research was testing biochar as a soil amendment when restoring vegetation in severely degraded habitats. My research clearly shows that biochar as a solitary amendment provided no positive plant growth effects and is detrimental to the development of soil food webs. Comparatively, co–amending soils with biochar, compost, and *Rhizophagus irregularis* significantly improved the biotic response of the soil food web and plant cover in the seed application trial, indicating improved soil conditions belowground. When approaching grassland restoration from an ecosystem perspective, the combination of all

amendments was shown to yield the most interesting prospects as a reclamation tool to restore vegetation and belowground soil food webs.

A limitation of this study was the replication in the seed application trial. As natural variability would be anticipated in field, the single replication of each factor combinations led to high variation in total native plant cover. Plot size and labor was a limiting factor that restricted the replication in the seed application experiment. Ideally, a smaller scale project could have been used to increase replication and statistical resolution.

6.2 Future directions

I would like to continue the investigation of the interactions among biochar, organic amendments, and arbuscular mycorrhizal fungi. My career as a restoration ecologist will focus on the recovery of plants and soils in severely degraded habitats. As biochar is a tool to create sustainable biofuels, improve soils conditions, and increase carbon sequestration, I think its role as in restoration needs to be explored more thoroughly. Thus, I would like to experiment with designing and creating ways to more cost-effectively produce biochar from inexpensive feed stocks and incorporate the resulting char into restoration plans.

In addition, restoration projects are often limited to short–term monitoring. I would like to continue to monitor the long–term research site established during my PhD program. The plots I have established will be available to be monitored indefinitely. I plan to continue to track the plant growth rates and soil food web recovery over multiple time points throughout my career. This long–term monitoring will prove to be an invaluable tool to investigate the influence of the land management tools on the recovery trajectory of the community. As the plant plug and seed application trial was only monitored for two years and three years respectively, this may not have been enough time to show large significant effects on the plant growth dynamics among the treatments. As the effect of solitary biochar addition to mine soils is unknown, there may be a lag in the response of plants and soil food web development early in the establishment of the ecosystem.

Bibliography

Abdi, H. 2010. Partial least squares regression and projection on latent structure regression (PLS Regression). Wiley Interdisciplinary Reviews: Computational Statistics 2(1):97–106.

Abel, S., A. Peters, S. Trinks, H. Schonsky, M. Facklam, and G. Wessolek. 2013. Impact of biochar and hydrochar addition on water retention and water repellency of sandy soil. Geoderma 202:183–191.

Abiven, S., S. Menasseri, and C. Chenu. 2009. The effects of organic inputs over time on soil aggregate stability - A literature analysis. Soil Biology and Biochemistry 41(1):1–12.

Adams, M., T. Benjamin, N. Emery, S. Brouder, and K. D. Gibson. 2013. The effect of biochar on native and invasive prairie plant species. Invasive Plant Science and Management 6(2):197–207.

Aggelides, S. M. and P. A. Londra. 2000. Effects of compost produced from town wastes and sewage sludge on the physical properties of a loamy and a clay soil. Bioresource Technology 71(3):253–259.

Albiach, R., R. Canet, F. Pomares, and F. Ingelmo. 2001. Organic matter components and aggregate stability after the application of different amendments to a horticultural soil. Bioresource Technology 76(2):125–129.

Alguacil, M. M., F. Caravaca, and A. Roldan. 2005. Changes in rhizosphere microbial activity mediated by native or allochthonous AM fungi in the reafforestation of a Mediterranean degraded environment. Biology and Fertility of Soils 41(1):59–68.

Allen, E. B. and M. F. Allen. 1980. Natural re–establishment of vesicular–arbuscular mycorrhizae following stripmine reclamation in Wyoming. Journal of Applied Ecology 17(1):139–147.

Allievi, L., A. Marchesini, C. Salardi, V. Piano, and A. Ferrari. 1993. Plant quality and soil residual fertility six years after a compost treatment. Bioresource Technology 43(1):85–89.

Amora-Lazcano, E., M. M. Vázquez, and R. Azcón. 1998. Response of nitrogen–transforming microorganisms to arbuscular mycorrhizal fungi. Biology and Fertility of Soils 27(1):65–70.

Andersen, C. M. and R. Bro. 2010. Variable selection in regressiona tutorial. Journal of Chemometrics 24(11):728–737.

Anderson, J. R. and D. Westmoreland. 1971. Direct counts of soil organisms using a fluorescent brightener and a europium chelate. Soil Biology and Biochemistry 3(1):85–87.

Anderson, T. 2003. Microbial eco–physiological indicators to asses soil quality. Agriculture, Ecosystems and Environment 98(1):285–293.

Annabi, M., S. Houot, C. Francou, M. Poitrenaud, and Y. L. Bissonnais. 2007. Soil aggregate stability improvement with urban composts of different maturities. Soil Science Society of America Journal 71(2):413–423.

Araújo, A. S. F., N. Eisenhauer, L. A. P. Nunes, L. F. C. Leite, and S. Cesarz. 2013. Soil surface–active fauna in degraded and restored lands of northeast Brazil. Land Degradation & Development doi: 10.1002/ldr.2247.

Ash, H. J., R. P. Gemmell, and A. D. Bradshaw. 1994. The introduction of native plant species on industrial waste heaps: A test of immigration and other factors affecting primary succession. Journal of Applied Ecology 31(1):74–84.

Azcón-Aguilar, C. and J. M. Barea. 1997. Applying mycorrhiza biotechnology to horticulture: Significance and potentials. Scientia Horticulturae 68(1):1–24.

Baer, S., J. Blair, S. Collins, and A. Knapp. 2004. Plant community responses to resource availability and heterogeneity during restoration. Oecologia 139(4):617–629.

Baer, S. G., J. M. Blair, S. L. Collins, and A. K. Knapp. 2003. Soil resources regulate productivity and diversity in newly established tallgrass prairie. Ecology 84(3):724–735.

Bailey, V. L., J. L. Smith, and H. Bolton Jr. 2002. Fungal–to–bacterial ratios in soils investigated for enhanced C sequestration. Soil Biology and Biochemistry 34(7):997–1007.

Bardgett, R. D. and R. Cook. 1998. Functional aspects of soil animal diversity in agricultural grasslands. Applied Soil Ecology 10(3):263–276.

Bardgett, R. D. and E. McAlister. 1999. The measurement of soil fungal:bacterial biomass ratios as an indicator of ecosystem self–regulation in temperate meadow grasslands. Biology and Fertility of Soils 29(3):282–290.

Bardgett, R. D. and D. A. Wardle. 2003. Herbivore–mediated linkages between aboveground and belowground communities. Ecology 84(9):2258–2268.

Bardgett, R. D., S. Keiller, R. Cook, and A. S. Gilburn. 1998. Dynamic interactions between soil animals and microorganisms in upland grassland soils amended with sheep dung: A microcosm experiment. Soil Biology and Biochemistry 30(4):531–539.

Bardgett, R. D., A. Richter, R. Bol, M. H. Garnett, R. Bäumler, X. Xu, E. Lopez-Capel, D. A. C. Manning, P. J. Hobbs, I. R. Hartley, and W. Wanek. 2007. Heterotrophic microbial communities use ancient carbon following glacial retreat. Biology Letters 3(5):487–490.

Bartoń, K. 2013. *MuMIn: Multi-model inference*. R package version 1.9.5. Available from http://CRAN.R–project.org/package=MuMIn.

Bastida, F., E. Kandeler, J. L. Moreno, M. Ros, C. García, and T. Hernández. 2008. Application of fresh and composted organic wastes modifies structure, size and activity of soil microbial community under semiarid climate. Applied Soil Ecology 40(2):318–329.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2014. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.0-6 Available from http://CRAN.R–project.org/package=lme4.

Beare, M. H., R. W. Parmelee, P. F. Hendrix, W. Cheng, D. C. Coleman, and D. A. Crossley Jr. 1992. Microbial and faunal interactions and effects on litter nitrogen and decomposition in agroecosystems. Ecological Monographs 62(4):569–591.

Beesley, L., O. S. Inneh, G. J. Norton, E. Moreno-Jimenez, T. Pardo, R. Clemente, and J. J. C. Dawson. 2014. Assessing the influence of compost and biochar amendments on the mobility and toxicity of metals and arsenic in a naturally contaminated mine soil. Environmental Pollution 186:195–202.

Benayas, J. M. R., A. C. Newton, A. Diaz, and J. M. Bullock. 2009. Enhancement of biodiversity and ecosystem services by ecological restoration: A meta–analysis. Science 325(5944):1121–1124.

Bengtsson, J., H. Setälä, and D. W. Zheng. 1996. Food webs and nutrient cycling in soils: Interactions and positive feedbacks. Pages 30–38 *in* G. A. Polis and K. O. Winemiller, editors. Food webs: Integration of patterns and dynamics. Chapman and Hall, New York, New York, U.S.A.

Bi, Y. L., X. L. Li, P. Christie, Z. Q. Hu, and M. H. Wong. 2003. Growth and nutrient uptake of arbuscular mycorrhizal maize in different depths of soil overlying coal fly ash. Chemosphere 50(6):863–869.

Biederman, L. A. 2013. Directing ecological restoration: Impact of organic amendments on above– and belowground ecosystem characteristics. PhD Thesis. Available at http://repository.tamu.edu/bitstream/handle/1969.1/ETD–TAMU–1371/BIEDERMAN–DISSERTATION.pdf?sequence=1.

Biederman, L. A. and W. S. Harpole. 2012. Biochar and its effects on plant productivity and nutrient cycling: a metaanalysis. Global Change Biology Bioenergy 5(2):202–214.

Biederman, L. A. and S. W. Whisenant. 2009. Organic amendments direct grass population dynamics in a landfill prairie restoration. Ecological Engineering 35(5):678–686.

Biederman, L. A., T. W. Boutton, and S. G. Whisenant. 2008. Nematode community development early in ecological restoration: The role of organic amendments. Soil Biology and Biochemistry 40(9):2366–2374.

Birk, J. J., C. Steiner, W. C. Teixiera, W. Zech, and B. Glaser. 2009. Microbial response to charcoal amendments and fertilization of a highly weathered tropical soil. Pages 309–324 *in* W. I. Woods, W. G. Teixeira, J. Lehmann, C. Steiner, A. WinklerPrins, and L. Rebellato, editors. Amazonian dark earths: Wim Sombroek's vision. Springer Netherlands.

Blackwell, P., G. Reithmuller, and M. Collins. 2009. Biochar application to soil. Pages 207–226 *in* C. Lehman and S. Joseph, editors. Biochar for environmental management: Science and technology. Earthscan, London.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. M. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: A practical guide for ecology and evolution. Trends in Ecology and Evolution 24(3):127–135.

Bongers, T. 1990. The maturity index: An ecological measure of environmental disturbance based on nematode species composition. Oecologia 83(1):14–19.

Bongers, T. and M. Bongers. 1998. Functional diversity of nematodes. Applied Soil Ecology 10(3):239–251.

Bongers, T. and H. Ferris. 1999. Nematode community structure as a bioindicator in environmental monitoring. Trends in Ecology and Evolution 14(6):224–228.

Bonham, C. D. 1982. Optimum allocation in multivariate double sampling for biomass estimation. Journal of Range Management 35(6):777–779.

Booth, D. T. and S. E. Cox. 2008. Image–based monitoring to measure ecological change in rangeland. Frontiers in Ecology and the Environment 6(4):185–190.

Bowen, G. D. and A. D. Rovira. 1999. The rhizosphere and its management to improve plant growth. Advances in Agronomy 66:1–102.

Bradshaw, A. 1997. Restoration of mined lands–using natural processes. Ecological Engineering 8(4):255–269.

Bradshaw, A. 2000. The use of natural processes in reclamation–advantages and difficulties. Landscape and Urban Planning 51(2):89–100.

Bresson, L. M., C. Koch, Y. Le Bissonnais, E. Barriuso, and V. Lecomte. 2001. Soil surface structure stabilization by municipal waste compost application. Soil Science Society of America Journal 65(6):1804–1811.

Briggs, J. M. and A. K. Knapp. 1995. Interannual variability in primary production in tallgrass prairie: Climate, soil moisture, topographic position, and fire as determinants of aboveground biomass. American Journal of Botany 82(8):1024–1030.

Brodowski, S., A. Rodionov, L. Haumaier, B. Glaser, and W. Amelung. 2005. Revised black carbon assessment using benzene polycarboxylic acids. Organic Geochemistry 36(9):1299–1310.

Brundrett, M. C., Y. Piche, and R. L. Peterson. 1984. A new method for observing the morphology of vesicular–arbuscular mycorrhizae. Canadian Journal of Botany 62(10):2128–2134.

Bruun, S., E. S. Jensen, and L. S. Jensen. 2008. Microbial mineralization and assimilation of black carbon: Dependency on degree of thermal alteration. Organic Geochemistry 39(7):839–845.

Buisson, E., K. D. Holl, S. Anderson, E. Corcket, G. F. Hayes, F. Torre, A. Peteers, and T. Dutoit. 2006. Effect of seed source, topsoil removal, and plant neighbor removal on restoring California coastal prairies. Restoration Ecology 14(4):569–577.

Burnham, K. P. and D. R. Anderson. 2004. Multimodel inference understanding AIC and BIC in model selection. Sociological Methods and Research 33(2):261–304.

Busby, R. R., H. Allen Torbert, and D. L. Gebhart. 2007. Carbon and nitrogen mineralization of non–composted and composted municipal solid waste in sandy soils. Soil Biology and Biochemistry 39(6):1277–1283.

Byrne, B. M. 2013. Structural equation modeling with AMOS: Basic concepts, applications, and programming. Mahwah, NJ: Lawrence Erlbaum.

Byrne, K. M., W. K. Lauenroth, P. B. Adler, and C. M. Byrne. 2011. Estimating aboveground net primary production in grasslands: A comparison of nondestructive methods. Rangeland Ecology and Management 64(5):498–505.

Cairns, J. 2000. Setting ecological restoration goals for technical feasibility and scientific validity. Ecological Engineering 15(3):171–180.

Callaham Jr., M. A., C. C. Rhoades, and L. Heneghan. 2008. A striking profile: Soil ecological knowledge in restoration management and science. Restoration Ecology 16(4):604–607.

Caravaca, F., J. M. Barea, and A. Roldán. 2002a. Synergistic influence of an arbuscular mycorrhizal fungus and organic amendment on *Pistacia lentiscus* L. seedlings afforested in a degraded semiarid soil. Soil Biology and Biochemistry 34(8):1139–1145.

Caravaca, F., T. Hernández, C. García, and A. Roldán. 2002b. Improvement of rhizosphere aggregate stability of afforested semiarid plant species subjected to mycorrhizal inoculation and compost addition. Geoderma 108(1):133–144.

Caravaca, F., D. Figueroa, C. Azcón-Aguilar, J. M. Barea, and A. Roldán. 2003a. Medium–term effects of mycorrhizal inoculation and composted municipal waste addition on the establishment of two Mediterranean shrub species under semiarid field conditions. Agriculture, Ecosystems and Environment 97(1):95–105.

Caravaca, F., D. Figueroa, A. Roldán, and C. Azcón-Aguilar. 2003b. Alteration in rhizosphere soil properties of afforested *Rhamnus lycioides* seedlings in short–term response to mycorrhizal inoculation with *Glomus intraradices* and organic amendment. Environmental Management 31(3):412–420.

Cardinale, B. J., J. P. Wright, M. W. Cadotte, I. T. Carroll, A. Hector, D. S. Srivastava, M. Loreau, and J. J. Weis. 2007. Impacts of plant diversity on biomass production increase through time because of species complementarity. Proceedings of the National Academy of Sciences 104(46):18123–18128.

Carrascal, L. M., I. Galván, and O. Gordo. 2009. Partial least squares regression as an alternative to current regression methods used in ecology. Oikos 118(5):681–690.

Catchpole, W. R. and C. J. Wheeler. 1992. Estimating plant biomass: A review of techniques. Australian Journal of Ecology 17(2):121–131.

Celik, I., I. Ortas, and S. Kilic. 2004. Effects of compost, mycorrhiza, manure and fertilizer on some physical properties of a Chromoxerert soil. Soil and Tillage Research 78(1):59–67.

Cernova, N. M. 1970. Regularities in the distribution of microarthropods in compost heaps. Pedobiologia 10:365–372.

Chan, K. Y. and Z. Xu. 2009. Biochar: Nutrient properties and their enhancement. Pages 67–84 *in* J. Lehmann and S. Joseph, editors. Biochar for environmental management: Science and technology. Earthscan, London.

Chave, J., R. Condit, S. Aguilar, A. Hernandez, S. Lao, and R. Perez. 2004. Error propagation and scaling for tropical forest biomass estimates. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 359(1443):409–420.

Clements, F. E. 1916. Plant succession: An analysis of the development of vegetation. Carnegie Institution, Washington, D.C.

Coleman, D. C. and W. B. Whitman. 2005. Linking species richness, biodiversity and ecosystem function in soil systems. Pedobiologia 49(6):479–497.

Courtney, R., A. M. Keith, and T. Harrington. 2011. Nematode assemblages in bauxite residue with different restoration histories. Restoration Ecology 19(6):758–764.

Craine, J. M. and R. Dybzinski. 2013. Mechanisms of plant competition for nutrients, water and light. Functional Ecology 27(4):833–840.

Crecchio, C., M. Curci, M. D. R. Pizzigallo, P. Ricciuti, and P. Ruggiero. 2004. Effects of municipal solid waste compost amendments on soil enzyme activities and bacterial genetic diversity. Soil Biology and Biochemistry 36(10):1595–1605.

Curry, J. P. and J. A. Good. 1992. Soil faunal degradation and restoration. Pages 171–215 *in* L. Rattan and B. A. Stewart, editors. Soil restoration. Springer New York.

de Souza, R. G., B. T. Goto, D. K. Alves da Silva, F. S. Barbosa da Silva, E. V. S. B. Sampaio, and L. C. Maia. 2010. The role of arbuscular mycorrhizal fungi and cattle manure in the establishment of *Tocoyena selloana* Schum. in mined dune areas. European Journal of Soil Biology 46(3):237–242.

de Vries, F. T., M. E. Liiri, L. Bjørnlund, M. A. Bowker, S. Christensen, H. M. Setälä, and R. D. Bardgett. 2012. Land use alters the resistance and resilience of soil food webs to drought. Nature Climate Change 2(4):276–280.

de Vries, F. T., E. Thébault, M. Liiri, K. Birkhofer, M. A. Tsiafouli, L. Bjørnlund, H. B. Jørgensen, M. V. Brady, S. Christensen, and P. C. de Ruiter. 2013. Soil food web properties explain ecosystem services across European land use systems. Proceedings of the National Academy of Sciences 110(35):14296–14301.

del Moral, R., L. Walker, and J. Bakker. 2007. Insights gained from succession for the restoration of landscape structure and function. Pages 19–44 *in* L. R. Walker, J. Walker, and R. J. Hobbs, editors. Linking restoration and ecological succession. Springer, New York, New York.

Delaney, K., L. Rodger, P. A. Woodlife, G. Rhynard, and P. Morris. 2000. Planting the seed – A guide to establishing prairie and meadow communities in southern Ontario. Environment Canada, Government of Canada. 61pp.

Diacono, M. and F. Montemurro. 2010. Long-term effects of organic amendments on soil fertility. A review. Agronomy for Sustainable Development 30(2):401–422.

Diaz, G. and M. Honrubia. 1994. A mycorrhizal survey of plants growing on mine wastes in Southeast Spain. Arid Land Research and Management 8(1):59–68.

Ding, Y., Y. Liu, W. Wu, D. Shi, M. Yang, and Z. Zhong. 2010. Evaluation of biochar effects on nitrogen retention and leaching in multi–layered soil columns. Water, Air, and Soil Pollution 213(1):47–55.

Doran, J. W. and M. R. Zeiss. 2000. Soil health and sustainability: Managing the biotic component of soil quality. Applied Soil Ecology 15(1):3–11.

Elkins, N. Z., L. W. Parker, E. Aldon, and W. G. Whitford. 1984. Responses of soil biota to organic amendments in stripmine spoils in northwestern New Mexico. Journal of Environmental Quality 13(2):215–219.

Elliott, K. J. and B. D. Clinton. 1993. Equations for estimating biomass of herbaceous and woody vegetation in early–successional southern Appalachian pine–hardwood forests. USDA Forest Service Research Note SE–365. 7p.

Enkhtuya, B., M. Pöschl, and M. Vosátka. 2005. Native grass facilitates mycorrhizal colonisation and P uptake of tree seedlings in two anthropogenic substrates. Water, Air, and Soil Pollution 166(1):217–236.

Eriksen, G. N., F. J. Coale, and G. A. Bollero. 1999. Soil nitrogen dynamics and maize production in municipal solid waste amended soil. Agronomy Journal 91(6):1009–1016.

Fay, P. A., J. D. Carlisle, A. K. Knapp, J. M. Blair, and S. L. Collins. 2003. Productivity responses to altered rainfall patterns in a C_4–dominated grassland. Oecologia 137(2):245–251.

Fellet, G., L. Marchiol, G. Delle Vedove, and A. Peressotti. 2011. Application of biochar on mine tailings: Effects and perspectives for land reclamation. Chemosphere 83(9):1262–1267.

Fellet, G., M. Marmiroli, and L. Marchiol. 2014. Elements uptake by metal accumulator species grown on mine tailings amended with three types of biochar. Science of the Total Environment 468:598–608.

Ferris, H. and T. Bongers. 2006. Nematode indicators of organic enrichment. Journal of Nematology 38(1):3–12.

Ferris, H., T. Bongers, and R. G. M. De Goede. 2001. A framework for soil food web diagnostics: Extension of the nematode faunal analysis concept. Applied Soil Ecology 18(1):13–29.

Filiberto, D. M. and J. L. Gaunt. 2013. Practicality of biochar additions to enhance soil and crop productivity. Agriculture 3(4):715–725.

Fischer, D. and B. Glaser. 2012. Synergisms between compost and biochar for sustainable soil amelioration. Management of Organic Waste 10:167–198.

Frank, L. E. and J. H. Friedman. 1993. A statistical view of some chemometrics regression tools. Technometrics 35(2):109–135.

Frankenberger, W. T. and W. A. Dick. 1983. Relationships between enzyme activities and microbial growth and activity indices in soil. Soil Science Society of America Journal 47(5):945–951.

Frenot, Y., J. C. Gloaguen, M. Cannavacciuolo, and A. Bellido. 1998. Primary succession on glacier forelands in the subantarctic Kerguelen Islands. Journal of Vegetation Science 9(1):75–84.

Frey, S. D., E. T. Elliott, and K. Paustian. 1999. Bacterial and fungal abundance and biomass in conventional and no–tillage agroecosystems along two climatic gradients. Soil Biology and Biochemistry 31(4):573–585.

Frouz, J., D. Elhottová, V. Kuráž, and M. Šourková. 2006. Effects of soil macrofauna on other soil biota and soil formation in reclaimed and unreclaimed post mining sites: Results of a field microcosm experiment. Applied Soil Ecology 33(3):308–320.

Gabrielle, B., J. Da-Silveira, S. Houot, and J. Michelin. 2005. Field–scale modelling of carbon and nitrogen dynamics in soils amended with urban waste composts. Agriculture, Ecosystems and Environment 110(3):289–299.

Ganesan, V., S. Ragupathy, B. Parthipan, D. B. R. Rani, and A. Mahadevan. 1991. Distribution of vesicular–arbuscular mycorrhizal fungi in coal, lignite, and calcite mine spoils of India. Biology and Fertility of Soils 12(2):131–136.

Garbaye, J. 1991. Biological interactions in the mycorrhizosphere. Experientia 47(4):370–375.

García-Gil, J. C., S. B. Ceppi, M. I. Velasco, A. Polo, and N. Senesi. 2004. Long–term effects of amendment with municipal solid waste compost on the elemental and acidic functional group composition and pH–buffer capacity of soil humic acids. Geoderma 121(1):135–142.

Gartshore, M. E., D. A. Sutherland, and J. D. McCracken. 1987. The natural areas inventory of the regional municipality of Haldimand–Norfolk. The Norfolk Field Naturalists. Volume 1: 112pp.

Gaur, A. and A. Adholeya. 2005. Diverse response of five ornamental plant species to mixed indigenous and single isolate arbuscular–mycorrhizal inocula in marginal soil amended with organic matter. Journal of Plant Nutrition 28(4):707–723.

Gholz, H. L., C. C. Grier, A. G. Campbell, and A. T. Brown. 1979. Equations for estimating biomass and leaf area of plants in the Pacific Northwest. Forest Research Lab, Oregon State University, Corvallis, Oregon, U.S.A. 43pp.

Ghosh, S., L. F. Ow, and B. Wilson. 2014. Influence of biochar and compost on soil properties and tree growth in a tropical urban environment. International Journal of Environmental Science and Technology doi: 10.1007/s13762–014–0508–0.

Giusquiani, P. L., M. Pagliai, G. Gigliotti, D. Businelli, and A. Benetti. 1995. Urban waste compost: Effects on physical, chemical, and biochemical soil properties. Journal of Environmental Quality 24(1):175–182.

Glaser, B. 2007. Prehistorically modified soils of central Amazonia: A model for sustainable agriculture in the twenty–first century. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences 362(1478):187–196.

Glaser, B., J. Lehmann, and W. Zech. 2002. Ameliorating physical and chemical properties of highly weathered soils in the tropics with charcoal–a review. Biology and Fertility of Soils 35(4):219–230.

Glatzle, A., A. Mechel, and M. E. V. Lourenco. 1993. Botanical components of annual Mediterranean grassland as determined by point–intercept and clipping methods. Journal of Range Management 46(3):271–274.

Glenn-Lewin, D. C. and E. van der Maarel. 1992. Patterns and processes of vegetation dynamics. Pages 11–59 *in* D. C. Glenn-Lewis, R. K. Peet, and T. T. Veblen, editors. Plant succession: Theory and prediction. Chapmann and Hall, London.

Govindarajulu, M., P. E. Pfeffer, H. Jin, J. Abubaker, D. D. Douds, J. W. Allen, H. Bücking, P. J. Lammers, and Y. Shachar-Hill. 2005. Nitrogen transfer in the arbuscular mycorrhizal symbiosis. Nature 435(7043):819–823.

Graber, E. R., Y. M. Harel, M. Kolton, E. Cytryn, A. Silber, D. R. David, L. Tsechansky, M. Borenshtein, and Y. Elad. 2010. Biochar impact on development and productivity of pepper and tomato grown in fertigated soilless media. Plant and Soil 337(1):481–496.

Grace, J. B. 2006. Structural equation modeling and natural systems. Cambridge University Press.

Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84(11):2809–2815.

Grueber, C. E., S. Nakagawa, R. J. Laws, and I. G. Jamieson. 2011. Multimodel inference in ecology and evolution: Challenges and solutions. Journal of Evolutionary Biology 24(4):699–711.

Gryndler, M., R. Sudová, D. Püschel, J. Rydlova, M. Janouskova, and M. Vosatka. 2008. Cultivation of high-biomass crops on coal mine spoil banks: Can microbial inoculation compensate for high doses of organic matter? Bioresource Technology 99(14):6391–6399.

Gupta, M. C., R. S. Singh, and K. Sitaramaiah. 1979. A new endoparasitic fungus on *Xiphinema* and cultivation of *Rhabditis spp.* and *Aphelenchus avenae* on same fungus. Nematologica 25(1):142.

Hadas, A. and R. Portnoy. 1997. Rates of decomposition in soil and release of available nitrogen from cattle manure and municipal waste composts. Compost Science and Utilization 5(3):48–54.

Hannah, L., J. L. Carr, and A. Lankerani. 1995. Human disturbance and natural habitat: A biome level analysis of a global data set. Biodiversity and Conservation 4(2):128–155.

Háněl, L. 2008. Nematode assemblages indicate soil restoration on colliery spoils afforested by planting different tree species and by natural succession. Applied Soil Ecology 40(1):86–99.

Hargreaves, J. C., M. S. Adl, and P. R. Warman. 2008. A review of the use of composted municipal solid waste in agriculture. Agriculture, Ecosystems and Environment 123(1):1–14.

Harinikumar, K. M., D. J. Bagyaraj, and B. C. Mallesha. 1990. Effect of intercropping and organic soil amendments on native VA mycorrhizal fungi in an oxisol. Arid Land Research and Management 4(3):193–197.

Harris, J. 2009. Soil microbial communities and restoration ecology: Facilitators or followers? Science 325(5940):573–574.

Harte, J. and A. P. Kinzig. 1993. Mutualism and competition between plants and decomposers: Implications for nutrient allocation in ecosystems. The American Naturalist 141(6):829–846.

Hastie, T., R. Tibshirani, and J. J. H. Friedman. 2001. The elements of statistical learning, vol. 1. Springer New York.

Hättenschwiler, S., A. V. Tiunov, and S. Scheu. 2005. Biodiversity and litter decomposition in terrestrial ecosystems. Annual Review of Ecology, Evolution, and Systematics 36:191–218.

Hector, A., B. Schmid, C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, H. Freitas, P. S. Giller, J. Good, R. Harris, P. Högberg, K. Huss-Danell, J. Joshi, A. Jumpponen, C. Körner, P. W. Leadley, M. Loreau, A. Minns, C. P. H. Mulder, G. O'Donovan, S. J. Otway, J. S. Pereira, A. Prinz, D. J. Read, M. Scherer-Lorenzen, E.-D. Schulze, A.-S. D. Siamantziouras, E. M. Spehn, A. C. Terry, A. Y. Troumbis, F. I. Woodward, S. Yachi, and J. H. Lawton. 1999. Plant diversity and productivity experiments in european grasslands. Science 286(5442):1123–1127.

Hedlund, K. and M. S. Öhrn. 2000. Tritrophic interactions in a soil community enhance decomposition rates. Oikos 88(3):585–591.

Hendrix, P. F., R. W. Parmelee, D. A. Crossley, D. C. Coleman, E. P. Odum, and P. M. Groffman. 1986. Detritus food webs in conventional and no–tillage agroecosystems. Bioscience 374–380.

Heneghan, L., S. P. Miller, S. Baer, M. A. Callaham, J. Montgomery, M. Pavao-Zuckerman, C. C. Rhoades, and S. Richardson. 2008. Integrating soil ecological knowledge into restoration management. Restoration Ecology 16(4):608–617.

Hodge, A. and A. H. Fitter. 2010. Substantial nitrogen acquisition by arbuscular mycorrhizal fungi from organic material has implications for N cycling. Proceedings of the National Academy of Sciences 107(31):13754–13759.

Hodge, A., T. Helgason, and A. H. Fitter. 2010. Nutritional ecology of arbuscular mycorrhizal fungi. Fungal Ecology 3(4):267–273.

Hodkinson, I. D., N. R. Webb, and S. J. Coulson. 2002. Primary community assembly on land–the missing stages: Why are the heterotrophic organisms always there first? Journal of Ecology 90(3):569–577.

Holtkamp, R., P. Kardol, A. van der Wal, S. C. Dekker, W. H. van der Putten, and P. C. de Ruiter. 2008. Soil food web structure during ecosystem development after land abandonment. Applied Soil Ecology 39(1):23–34.

Hooper, D. U., F. S. Chapin Iii, J. J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. M. Lodge, M. Loreau, and S. Naeem. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. Ecological Monographs 75(1):3–35.

Hortenstine, C. C. and D. F. Rothwell. 1972. Use of municipal compost in reclamation of phosphate–mining sand tailings. Journal of Environmental Quality 1(4):415–418.

Hunt, G. A. and R. Fogel. 1983. Fungal hyphal dynamics in a western Oregon Douglas–fir stand. Soil Biology and Biochemistry 15(6):641–649.

Hunt, H. W., D. C. Coleman, E. R. Ingham, R. E. Ingham, E. T. Elliott, J. C. Moore, S. L. Rose, C. P. P. Reid, and C. R. Morley. 1987. The detrital food web in a shortgrass prairie. Biology and Fertility of Soils 3(1):57–68.

Iglesias-Jiménez, E. 2001. Nitrogen availability from a mature urban compost determined by the ^{15}N isotope dilution method. Soil Biology and Biochemistry 33(3):409–412.

Ingham, R. E., J. A. Trofymow, E. R. Ingham, and D. C. Coleman. 1985. Interactions of bacteria, fungi, and their nematode grazers: Effects on nutrient cycling and plant growth. Ecological Monographs 55(1):119–140.

Ishii, T. and K. Kadoya. 1994. Effects of charcoal as a soil conditioner on citrus growth and vesicular–arbuscular mycorrhizal development. Journal of the Japanese Society for Horticultural Science (Japan) 63(3):529–535.

Jackson, S. T. and R. J. Hobbs. 2009. Ecological restoration in the light of ecological history. Science 325(5940):567–569.

Jakobsen, S. T. 1995. Aerobic decomposition of organic wastes 2. Value of compost as a fertilizer. Resources, Conservation and Recycling 13(1):57–71.

Jansa, J., A. Mozafar, T. Anken, R. Ruh, I. Sanders, and E. Frossard. 2002. Diversity and structure of AMF communities as affected by tillage in a temperate soil. Mycorrhiza 12(5):225–234.

Jansa, J., A. Mozafar, G. Kuhn, T. Anken, R. Ruh, I. R. Sanders, and E. Frossard. 2003. Soil tillage affects the community structure of mycorrhizal fungi in maize roots. Ecological Applications 13(4):1164–1176.

Jasper, D. A., L. K. Abbott, and A. D. Robson. 1989. Soil disturbance reduces the infectivity of external hyphae of vesicular–arbuscular mycorrhizal fungi. New Phytologist 112(1):93–99.

Jastrow, J. D., R. M. Miller, and J. Lussenhop. 1998. Contributions of interacting biological mechanisms to soil aggregate stabilization in restored prairie. Soil Biology and Biochemistry 30(7):905–916.

Jastrow, J. D., J. E. Amonette, and V. L. Bailey. 2007. Mechanisms controlling soil carbon turnover and their potential application for enhancing carbon sequestration. Climatic Change 80(1):5–23.

Jeffery, S., F. G. A. Verheijen, M. Van der Velde, and A. C. Bastos. 2011. A quantitative review of the effects of biochar application to soils on crop productivity using meta–analysis. Agriculture, Ecosystems and Environment 144(1):175–187.

Johansson, J. F., L. R. Paul, and R. D. Finlay. 2004. Microbial interactions in the mycorrhizosphere and their significance for sustainable agriculture. FEMS Microbiology Ecology 48(1):1–13.

Johnson, J. B. and K. S. Omland. 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution 19(2):101–108.

Johnson, N. C. 1998. Responses of *Salsola kali* and *Panicum virgatum* to mycorrhizal fungi, phosphorus and soil organic matter: Implications for reclamation. Journal of Applied Ecology 35(1):86–94.

Johnson, N. C., J.-H. Graham, and F. A. Smith. 1997. Functioning of mycorrhizal associations along the mutualism–parasitism continuum. New Phytologist 135(4):575–585.

Jonasson, S. 1988. Evaluation of the point intercept method for the estimation of plant biomass. Oikos 52(1):101–106.

Jones, B. E. H., R. J. Haynes, and I. R. Phillips. 2010. Effect of amendment of bauxite processing sand with organic materials on its chemical, physical and microbial properties. Journal of Environmental Management 91(11):2281–2288.

Jones, D. L., J. Rousk, G. Edwards-Jones, T. H. DeLuca, and D. V. Murphy. 2012. Biochar–mediated changes in soil quality and plant growth in a three year field trial. Soil Biology and Biochemistry 45:113–124.

Jørgensen, H. B. and K. Hedlund. 2013. Organic amendment and fungal species in combination can alter collembolan fitness. Soil Biology and Biochemistry 65:316–321.

Kardol, P. and D. A. Wardle. 2010. How understanding aboveground–belowground linkages can assist restoration ecology. Trends in Ecology and Evolution 25(11):670–679.

Karlen, D. L., M. J. Mausbach, J. W. Doran, R. G. Cline, R. F. Harris, and G. E. Schuman. 1997. Soil quality: A concept, definition, and framework for evaluation (A guest editorial). Soil Science Society of America Journal 61(1):4–10.

Kelly, C. N., C. D. Peltz, M. Stanton, D. W. Rutherford, and C. E. Rostad. 2014. Biochar application to hardrock mine tailings: Soil quality, microbial activity, and toxic element sorption. Applied Geochemistry 43:35–48.

Khan, A. G. 1981. Growth responses of endomycorrhizal onions in unsterilized coal waste. New Phytologist 87(2):363–370.

Khater, C., A. Martin, and J. Maillet. 2003. Spontaneous vegetation dynamics and restoration prospects for limestone quarries in Lebanon. Applied Vegetation Science 6(2):199–204.

Kinney, T. J., C. A. Masiello, B. Dugan, W. C. Hockaday, M. R. Dean, K. Zygourakis, and R. T. Barnes. 2012. Hydrologic properties of biochars produced at different temperatures. Biomass and Bioenergy 41:34–43.

Kjøller, A. and S. Struwe. 1982. Microfungi in ecosystems: Fungal occurrence and activity in litter and soil. Oikos 391–422.

Klironomos, J. N. 2002. Feedback with soil biota contributes to plant rarity and invasiveness in communities. Nature 417(6884):67–70.

Klironomos, J. N. 2003. Variation in plant response to native and exotic arbuscular mycorrhizal fungi. Ecology 84(9):2292–2301.

Klironomos, J. N., P. Moutoglis, B. Kendrick, and P. Widden. 1993. A comparison of spatial heterogeneity of vesicular–arbuscular mycorrhizal fungi in two maple–forest soils. Canadian Journal of Botany 71(11):1472–1480.

Klironomos, J. N., M. C. Rillig, and M. F. Allen. 1996. Below–ground microbial and microfaunal responses to *Artemisia tridentata* grown under elevated atmospheric CO_2. Functional Ecology 10(4):527–534.

Knapp, A. K. and T. R. Seastedt. 1986. Detritus accumulation limits productivity of tallgrass prairie. BioScience 36(10):662–668.

Knapp, A. K. and M. D. Smith. 2001. Variation among biomes in temporal dynamics of aboveground primary production. Science 291(5503):481–484.

Knapp, A. K., J. M. Blair, J. M. Briggs, S. L. Collins, D. C. Hartnett, L. C. Johnson, and E. G. Towne. 1999. The keystone role of bison in North American tallgrass prairie. BioScience 49(1):39–50.

Korthals, G. W., P. Smilauer, C. Van Dijk, and W. H. Van der Putten. 2001. Linking above– and below–ground biodiversity: Abundance and trophic complexity in soil as a response to experimental plant communities on abandoned arable land. Functional Ecology 15(4):506–514.

Kuzyakov, Y. 2002. Review: Factors affecting rhizosphere priming effects. Journal of Plant Nutrition and Soil Science 165(4):382.

Laakso, J. and H. Setälä. 1999. Sensitivity of primary production to changes in the architecture of belowground food webs. Oikos 87:57–64.

Laakso, J., H. Setälä, and A. Palojärvi. 2000. Influence of decomposer food web structure and nitrogen availability on plant growth. Plant and Soil 225(1):153–165.

Laird, D. A. 2008. The charcoal vision: A win–win–win scenario for simultaneously producing bioenergy, permanently sequestering carbon, while improving soil and water quality. Agronomy Journal 100(1):178–181.

Lal, R. 2009. Challenges and opportunities in soil organic matter research. European Journal of Soil Science 60(2):158–169.

Langley, J. A. and B. A. Hungate. 2003. Mycorrhizal controls on belowground litter quality. Ecology 84(9):2302–2312.

Larney, F. J. and D. A. Angers. 2012. The role of organic amendments in soil reclamation: A review. Canadian Journal of Soil Science 92(1):19–38.

Lavelle, P., T. Decans, M. Aubert, S. Barot, M. Blouin, F. Bureau, P. Margerie, P. Mora, and J. P. Rossi. 2006. Soil invertebrates and ecosystem services. European Journal of Soil Biology 42:S3–S15.

Lee, Q. and P. Widden. 1996. *Folsomia candida*, a "fungivorous" collembolan, feeds preferentially on nematodes rather than soil fungi. Soil Biology and Biochemistry 28(4):689–690.

Lehmann, J., J. Pereira da Silva, C. Steiner, T. Nehls, W. Zech, and B. Glaser. 2003. Nutrient availability and leaching in an archaeological Anthrosol and a Ferralsol of the Central Amazon basin: Fertilizer, manure and charcoal amendments. Plant and Soil 249(2):343–357.

Lehmann, J., J. Gaunt, and M. Rondon. 2006. Bio–char sequestration in terrestrial ecosystems–a review. Mitigation and Adaptation Strategies for Global Change 11(2):395–419.

Lehmann, J., C. Czimczik, D. A. Laird, and S. Sohi. 2009. Stability of Biochar in Soil. Pages 183–205 *in* J. Lehmann and S. Joseph, editors. Biochar for environmental management: Science and technology. Earthscan, London.

Lehmann, J., M. Rillig, J. Thies, C. A. Masiello, W. C. Hockaday, and D. Crowley. 2011. Biochar effects on soil biota–a review. Soil Biology and Biochemistry 43(9):1812–1836.

Lenoir, L., T. Persson, J. Bengtsson, H. Wallander, and A. Wiren. 2007. Bottom–up or top–down control in forest soil microcosms? Effects of soil fauna on fungal biomass and C/N mineralisation. Biology and Fertility of Soils 43(3):281–294.

Lesica, P. and F. W. Allendorf. 1999. Ecological genetics and the restoration of plant communities: Mix or match? Restoration Ecology 7(1):42–50.

Li, X., Q. Liu, Z. Liu, W. Shi, D. Yang, and E. Tarasco. 2014. Effects of organic and other management practices on soil nematode communities in tea plantation: A case study in southern China. Journal of Plant Nutrition and Soil Science 177(4):604–612.

Liang, B., J. Lehmann, D. Solomon, J. Kinyangi, J. Grossman, B. O'neill, J. O. Skjemstad, J. Thies, F. J. Luizao, and J. Petersen. 2006. Black carbon increases cation exchange capacity in soils. Soil Science Society of America Journal 70(5):1719–1730.

Lou, Y., S. A. Clay, A. S. Davis, A. Dille, J. Felix, A. H. M. Ramirez, C. L. Sprague, and A. C. Yannarell. 2014. An affinity–effect relationship for microbial communities in plant–soil feedback loops. Microbial Ecology 67(4):866–876.

Lucas, S. T., E. M. D'Angelo, and M. A. Williams. 2014. Improving soil structure by promoting fungal abundance with organic soil amendments. Applied Soil Ecology 75:13–23.

Lussenhop, J. 1971. A simplified canister–type soil arthropod extractor. Pedobiologia 11:40–45.

Maherali, H. and J. N. Klironomos. 2007. Influence of phylogeny on fungal community assembly and ecosystem functioning. Science 316(5832):1746–1748.

Maiti, S. K. 2013. Ecology and ecosystem in mine–degraded land. Pages 21–37 *in* Ecorestoration of the coalmine degraded lands. Springer India.

Major, J., M. Rondon, D. Molina, S. Riha, and J. Lehmann. 2010. Maize yield and nutrition during 4 years after biochar application to a Colombian savanna oxisol. Plant and Soil 333(1):117–128.

Matias, S. R., M. C. Pagano, F. C. Muzzi, C. A. Oliveira, A. A. Carneiro, S. N. Horta, and M. R. Scotti. 2009. Effect of rhizobia, mycorrhizal fungi and phosphate-solubilizing microorganisms in the rhizosphere of native plants used to recover an iron ore area in Brazil. European Journal of Soil Biology 45(3):259–266.

McCann, K. S. 2000. The diversity–stability debate. Nature 405(6783):228–233.

McConnell, D., A. Shiralipour, and W. Smith. 1994. Compost impact on soil/plant properties. Pages 89–91 *in* Composting source separated organics JG Press, Emmaus, PA.

McCormack, S. A., N. Ostle, R. D. Bardgett, D. W. Hopkins, and A. J. Vanbergen. 2013. Biochar in bioenergy cropping systems: impacts on soil faunal communities and linked ecosystem processes. Global Change Biology Bioenergy 5(2):81–95.

McGonigle, T. P., M. H. Miller, D. G. Evans, G. L. Fairchild, and J. A. Swan. 1990. A new method which gives an objective measure of colonization of roots by vesicular–arbuscular mycorrhizal fungi. New Phytologist 115(3):495–501.

McKinley, V. L., A. D. Peacock, and D. C. White. 2005. Microbial community PLFA and PHB responses to ecosystem restoration in tallgrass prairie soils. Soil Biology and Biochemistry 37(10):1946–1958.

McLauchlan, K. 2006. The nature and longevity of agricultural impacts on soil carbon and nutrients: A review. Ecosystems 9(8):1364–1382.

Mehmood, T., H. Martens, S. Sæbø, J. Warringer, and L. Snipen. 2011. A partial least squares based algorithm for parsimonious variable selection. Algorithms for Molecular Biology 6(1):27.

Menta, C. 2012. Soil fauna diversity–function, soil degradation, biological indices, soil restoration. Pages 59–94 *in* G. A. Lameed, editor. Biodiversity conservation in a diverse world. InTech, Rijeka, Croatia.

Mevik, B. H. and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. Journal of Statistical Software 18(2):1–24.

Mikola, J. and H. Setälä. 1998. Productivity and trophic-level biomasses in a microbial-based soil food web. Oikos 82(1):158–168.

Miller, R. M. and J. D. Jastrow. 1992. The application of VA mycorrhizae to ecosystem restoration and reclamation. Chapman and Hall, New York, New York.

Mkhabela, M. S. and P. R. Warman. 2005. The influence of municipal solid waste compost on yield, soil phosphorus availability and uptake by two vegetable crops grown in a Pugwash sandy loam soil in Nova Scotia. Agriculture, Ecosystems and Environment 106(1):57–67.

Montemurro, F., M. Maiorana, G. Convertini, and D. Ferri. 2006. Compost organic amendments in fodder crops: Effects on yield, nitrogen utilization and soil characteristics. Compost Science and Utilization 14(2):114–123.

Movahedi-Naeini, S. A. R. and H. F. Cook. 2000. Influence of municipal compost on temperature, water, nutrient status and the yield of maize in a temperate soil. Soil Use and Management 16(3):215–221.

Moya-Larao, J. and G. Corcobado. 2008. Plotting partial correlation and regression in ecological studies. Web Ecology 8:35–46.

Mukherjee, A. and A. R. Zimmerman. 2013. Organic carbon and nutrient release from a range of laboratory–produced biochars and biochar–soil mixtures. Geoderma 193:122–130.

Mummey, D. L., P. D. Stahl, and J. S. Buyer. 2002. Microbial biomarkers as an indicator of ecosystem recovery following surface mine reclamation. Applied Soil Ecology 21(3):251–259.

Mummey, D. L., P. M. Antunes, and M. C. Rillig. 2009. Arbuscular mycorrhizal fungi pre–inoculant identity determines community composition in roots. Soil Biology and Biochemistry 41(6):1173–1179.

Nemergut, D. R., S. P. Anderson, C. C. Cleveland, A. P. Martin, A. E. Miller, A. Seimon, and S. K. Schmidt. 2007. Microbial community succession in an unvegetated, recently deglaciated soil. Microbial Ecology 53(1):110–122.

Nguyen, B. T., J. Lehmann, J. Kinyangi, R. Smernik, S. J. Riha, and M. H. Engelhard. 2008. Long–term black carbon dynamics in cultivated soil. Biogeochemistry 89(3):295–308.

Niklas, K. J. and B. J. Enquist. 2002. On the vegetative biomass partitioning of seed plant leaves, stems, and roots. The American Naturalist 159(5):482–497.

Nishio, M. and S. Okano. 1991. Stimulation of the growth of alfalfa and infection of roots with indigenous vesicular–arbuscular mycorrhizal fungi by the application of charcoal. Bull. Natl. Grassi. Res. Inst 45:61–71.

Norland, M. R. and D. L. Veith. 1995. Revegetation of coarse taconite iron ore tailing using municipal solid waste compost. Journal of Hazardous Materials 41(2):123–134.

Novak, J. M., I. Lima, B. Xing, J. W. Gaskin, C. Steiner, K. C. Das, M. Ahmedna, D. Rehrah, D. W. Watts, and W. J. Busscher. 2009. Characterization of designer biochar produced at different temperatures and their effects on a loamy sand. Annals of Environmental Science 3:195–206.

Noyd, R. K., F. L. Pfleger, and M. P. Russelle. 1995. Interactions between native prairie grasses and indigenous arbuscular mycorrhizal fungi: Implications for reclamation of taconite iron ore tailing. New Phytologist 129(4):651–660.

Noyd, R. K., F. L. Pfleger, and M. R. Norland. 1996. Field responses to added organic matter, arbuscular mycorrhizal fungi, and fertilizer in reclamation of taconite iron ore tailing. Plant and Soil 179(1):89–97.

Odum, E. P. 1969. The strategy of ecosystem development. Science 164(877):262–270.

Oelmann, Y., W. Wilcke, V. M. Temperton, N. Buchmann, C. Roscher, J. Schumacher, E. D. Schulze, and W. W. Weisser. 2007. Soil and plant nitrogen pools as related to plant diversity in an experimental grassland. Soil Science Society of America Journal 71(3):720–729.

Ojala, R. and V. Huhta. 2001. Dispersal of microarthropods in forest soil. Pedobiologia 45(5):443–450.

Ojima, D. S., B. O. M. Dirks, E. P. Glenn, C. E. Owensby, and J. O. Scurlock. 1993. Assessment of C budget for grasslands and drylands of the world. Water, Air, and Soil Pollution 70(1):95–109.

Okada, H., H. Harada, and I. Kadota. 2005. Fungal–feeding habits of six nematode isolates in the genus *Filenchus*. Soil Biology and Biochemistry 37(6):1113–1120.

Ontario-Biodiversity-Council. 2010. State of Ontario's Biodiversity 2010. A report of the Ontario Biodiversity Council, Peterborough, ON.

Ouédraogo, E., A. Mando, and N. P. Zombré. 2001. Use of compost to improve soil properties and crop productivity under low input agricultural system in West Africa. Agriculture, Ecosystems & Environment 84(3):259–266.

Parisi, V., C. Menta, C. Gardi, C. Jacomini, and E. Mozzanica. 2005. Microarthropod communities as a tool to assess soil quality and biodiversity: a new approach in Italy. Agriculture, Ecosystems & Environment 105(1):323–333.

Parton, W. J., J. M. O. Scurlock, D. S. Ojima, T. G. Gilmanov, R. J. Scholes, D. S. Schimel, T. Kirchner, J. C. Menaut, T. Seastedt, and E. G. Moya. 1993. Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. Global Biogeochemical Cycles 7(4):785–809.

Pascual, J. A., C. García, and T. Hernandez. 1999. Lasting microbiological and biochemical effects of the addition of municipal solid waste to an arid soil. Biology and Fertility of Soils 30(1):1–6.

Picard, N., L. Saint-André, and M. Henry. 2012. Manual for building tree volume and biomass allometric equations: From field measurement to prediction. Food and Agricultural Organization of the United Nations, Rome and and Centre de Coopration Internationale en Recherche Agronomique pour le Dveloppement , Montpellier.

Pimm, S. L. 1991. The balance of nature?: Ecological issues in the conservation of species and communities. University of Chicago Press.

Piper, J. K. and S. L. Pimm. 2002. The creation of diverse prairie–like communities. Community Ecology 3(2):205–216.

Pokorny, M. L., R. L. Sheley, T. J. Svejcar, and R. E. Engel. 2004. Plant species diversity in a grassland plant community: Evidence for forbs as a critical management consideration. Western North American Naturalist 64(2):219–230.

Poorter, H. and O. Nagel. 2000. The role of biomass allocation in the growth response of plants to different levels of light, CO_2, nutrients and water: A quantitative review. Functional Plant Biology 27(12):1191–1191.

Prach, K. and R. J. Hobbs. 2008. Spontaneous succession versus technical reclamation in the restoration of disturbed sites. Restoration Ecology 16(3):363–366.

Püschel, D., J. Rydlová, R. Sudová, and M. Gryndler. 2008a. Cultivation of flax in spoil bank clay: Mycorrhizal inoculation vs. high organic amendments. Journal of Plant Nutrition and Soil Science 171(6):872–877.

Püschel, D., J. Rydlová, and M. Vosátka. 2008b. Does the sequence of plant dominants affect mycorrhiza development in simulated succession on spoil banks? Plant and Soil 302(1):273–282.

Püschel, D., J. Rydlová, R. Sudová, M. Gryndler, and M. Vosátka. 2011. The potential of mycorrhizal inoculation and organic amendment to increase yields of *Galega orientalis* and *Helianthus tuberosus* in a spoil bank substrate. Journal of Plant Nutrition and Soil Science 174(4):664–672.

Pywell, R. F., J. M. Bullock, A. Hopkins, K. J. Walker, T. H. Sparks, M. J. W. Burke, and S. Peel. 2002. Restoration of species rich grassland on arable land: Assessing the limiting processes using a multi site experiment. Journal of Applied Ecology 39(2):294–309.

R-Core-Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at http://www.R-project.org/.

Rao, A. V. and R. Tak. 2002. Growth of different tree species and their nutrient uptake in limestone mine spoil as influenced by arbuscular mycorrhizal (AM)–fungi in Indian arid zone. Journal of Arid Environments 51(1):113–119.

Rayburn, E. and J. Lozier. 2003. A falling plate meter for estimating pasture forage mass. West Virginia University, Morgantown, West Virginia Extension. Fact Sheet. Nov. 2003 .

Rayburn, E. B. and S. B. Rayburn. 1998. A standardized plate meter for estimating pasture mass in on–farm research trials. Agronomy Journal 90(2):238–241.

Reynolds, H. L., A. Packer, J. D. Bever, and K. Clay. 2003. Grassroots ecology: Plant–microbe–soil interactions as drivers of plant community structure and dynamics. Ecology 84(9):2281–2291.

Richter, B. S. and J. C. Stutz. 2002. Mycorrhizal inoculation of big sacaton: Implications for grassland restoration of abandoned agricultural fields. Restoration Ecology 10(4):607–616.

Rillig, M. C. 2004. Arbuscular mycorrhizae and terrestrial ecosystem processes. Ecology Letters 7(8):740–754.

Risser, P. G. and W. J. Parton. 1982. Ecosystem analysis of the tallgrass prairie: Nitrogen cycle. Ecology 63(5):1342–1351.

Rodger, L. 1998. Tallgrass communities of southern Ontario: A recovery plan. World Wildlife Fund and Ontario Ministry of Natural Resources.

Rondon, M., J. Lehmann, J. Ramirez, and M. Hurtado. 2007. Biological nitrogen fixation by common beans (*Phaseolus vulgaris* L.) increases with bio–char additions. Biology and Fertility of Soils 43(6):699–708.

Ros, M., M. T. Hernandez, and C. Garcia. 2003. Soil microbial activity after restoration of a semiarid soil by organic amendments. Soil Biology and Biochemistry 35(3):463–469.

Ros, M., S. Klammer, B. Knapp, K. Aichberger, and H. Insam. 2006. Long–term effects of compost amendment of soil on functional and structural diversity and microbial activity. Soil Use and Management 22(2):209–218.

Roubíčková, A., O. Mudrák, and J. Frouz. 2013. Soil fauna plant interactions during succession at post–mining sites. Pages 250–264 *in* J. Frouz, editor. Soil biota and ecosystem development in post mining sites. CRC Press, Boca Raton, FL.

Rowe, H. I., C. S. Brown, and V. P. Claassen. 2007. Comparisons of mycorrhizal responsiveness with field soil and commercial inoculum for six native montane species and *Bromus tectorum*. Restoration Ecology 15(1):44–52.

Ruiz-Jaen, M. C. and M. Aide. 2005. Restoration success: How is it being measured? Restoration Ecology 13(3):569–577.

Rydlová, J., D. Püschel, M. Vosátka, and K. Charvátová. 2008. Different effect of mycorrhizal inoculation in direct and indirect reclamation of spoil banks. Journal of Applied Botany and Food Quality 82(1):15–20.

Sala, O. E., W. J. Parton, L. A. Joyce, and W. K. Lauenroth. 1988. Primary production of the central grassland region of the United States. Ecology 69(1):40–45.

Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, and B. Walker. 2001. Catastrophic shifts in ecosystems. Nature 413(6856):591–596.

Scherber, C., N. Eisenhauer, W. W. Weisser, B. Schmid, W. Voigt, M. Fischer, E.-D. Schulze, C. Roscher, A. Weigelt, and E. Allan. 2010. Bottom–up effects of plant diversity on multitrophic interactions in a biodiversity experiment. Nature 468(7323):553–556.

Scheu, S. and E. Schulz. 1996. Secondary succession, soil formation and development of a diverse community of oribatids and saprophagous soil macro–invertebrates. Biodiversity and Conservation 5(2):235–250.

Schielzeth, H. 2010. Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution 1(2):103–113.

Schimmelpfennig, S., C. Müller, L. Grünhage, C. Koch, and C. Kammann. 2014. Biochar, hydrochar and uncarbonized feedstock application to permanent grassland – effects on greenhouse gas emissions and plant growth. Agriculture, Ecosystems & Environment 191:39–52.

Schmidt, H.-P., C. Kammann, C. Niggli, M. W. H. Evangelou, K. A. Mackie, and S. Abiven. 2014. Biochar and biochar-compost as soil amendments to a vineyard soil: Influences on plant growth, nutrient uptake, plant health and grape quality. Agriculture, Ecosystems & Environment 191:117–123.

Schmidt, M. W. I. and A. G. Noack. 2000. Black carbon in soils and sediments: Analysis, distribution, implications, and current challenges. Global Biogeochemical Cycles 14(3):777–793.

Schulz, H. and B. Glaser. 2012. Effects of biochar compared to organic and inorganic fertilizers on soil quality and plant growth in a greenhouse experiment. Journal of Plant Nutrition and Soil Science 175(3):410–422.

Schwartz, M. W., J. D. Hoeksema, C. A. Gehring, N. C. Johnson, J. N. Klironomos, L. K. Abbott, and A. Pringle. 2006. The promise and the potential consequences of the global transport of mycorrhizal fungal inoculum. Ecology Letters 9(5):501–515.

Seastedt, T. R. 1984. The role of microarthropods in decomposition and mineralization processes. Annual Review of Entomology 29(1):25–46.

Séré, G., C. Schwartz, S. Ouvrard, C. Sauvage, J. Renat, and J. Morel. 2008. Soil construction: A step for ecological reclamation of derelict lands. Journal of Soils and Sediments 8(2):130–136.

Shiralipour, A., D. B. McConnell, and W. H. Smith. 1992. Physical and chemical properties of soils as affected by municipal solid waste compost application. Biomass and Bioenergy 3(3):261–266.

Shrestha, G., S. J. Traina, and C. W. Swanston. 2010. Black carbon's properties and role in the environment: A comprehensive review. Sustainability 2:294–320.

Singh, B., B. P. Singh, and A. L. Cowie. 2010. Characterisation and evaluation of biochars for their application as a soil amendment. Soil Research 48(7):516–525.

Singh, J. S., W. K. Lauenroth, and R. K. Steinhorst. 1975. Review and assessment of various techniques for estimating net aerial primary production in grasslands from harvest data. Botanical Review 41(2):181–232.

Six, J., H. Bossuyt, S. Degryze, and K. Denef. 2004. A history of research on the link between (micro)aggregates, soil biota, and soil organic matter dynamics. Soil and Tillage Research 79(1):7–31.

Skjemstad, J. O., D. C. Reicosky, A. R. Wilts, and J. A. McGowan. 2002. Charcoal carbon in US agricultural soils. Soil Science Society of America Journal 66(4):1249–1255.

Smith, R. S., R. S. Shiel, R. D. Bardgett, D. Millward, P. Corkhill, G. Rolph, P. J. Hobbs, and S. Peacock. 2003. Soil microbial community, fertility, vegetation and diversity as targets in the restoration management of a meadow grassland. Journal of Applied Ecology 40(1):51–64.

Smith, S. E. and D. J. Read. 2008. Mycorrhizal Symbiosis, 3^{rd} edition. Academic Press, Elsevier, Amsterdam.

Smith, V. C. and M. A. Bradford. 2003. Litter quality impacts on grassland litter decomposition are differently dependent on soil fauna across time. Applied Soil Ecology 24(2):197–203.

Sohi, S., E. Lopez-Capel, E. Krull, and R. Bol. 2009. Biochar, climate change and soil: A review to guide future research. CSIRO Land and Water Science Report 05/09. 64pp.

Stahl, P. D., S. E. Williams, and M. Christensen. 1988. Efficacy of native vesicular–arbuscular mycorrhizal fungi after severe soil disturbance. New Phytologist 110(3):347–354.

Steel, H., B. Vandecasteele, K. Willekens, K. Sabbe, T. Moens, and W. Bert. 2012. Nematode communities and macronutrients in composts and compost–amended soils as affected by feedstock composition. Applied Soil Ecology 61:100–112.

Steel, H., D. Buchan, S. De Neve, M. Couvreur, T. Moens, and W. Bert. 2013a. Nematode and microbial communities in a rapidly changing compost environment: How nematode assemblages reflect composting phases. European Journal of Soil Biology 56:1–10.

Steel, H., F. Verdoodt, A. Čerevková, M. Couvreur, P. Fonderie, T. Moens, and W. Bert. 2013b. Survival and colonization of nematodes in a composting process. Invertebrate Biology 132(2):108–119.

Steinbeiss, S., G. Gleixner, and M. Antonietti. 2009. Effect of biochar amendment on soil carbon balance and soil microbial activity. Soil Biology and Biochemistry 41(6):1301–1310.

Steiner, C., W. Teixeira, J. Lehmann, T. Nehls, J. Luis, J. de Macêdo, W. E. H. Blum, and W. Zech. 2007. Long term effects of manure, charcoal and mineral fertilization on crop production and fertility on a highly weathered Central Amazonian upland soil. Plant and Soil 291(1):275–290.

Streit, B., A. Buehlmann, and P. Reutimann. 1985. Mite succession in compost communities: Studies with Oribatei, Gamasina, and Uropodina. Pedobiologia 28(1):1–12.

Suding, K. N., K. L. Gross, and G. R. Houseman. 2004. Alternative states and positive feedbacks in restoration ecology. Trends in Ecology and Evolution 19(1):46–53.

Swift, M. J., O. W. Heal, and J. M. Anderson. 1979. Decomposition in terrestrial ecosystems, vol. 5. Blackwell Scientific Publication, Oxford, UK.

Symonds, J. E. and J. Pither. 2012. Multi–scale analyses of exotic and native plant species diversity within Canada's endangered antelope–brush ecosystem. Plant Ecology 213(8):1263–1275.

Taheri, W. I. and J. D. Bever. 2010. Adaptation of plants and arbuscular mycorrhizal fungi to coal tailings in Indiana. Applied Soil Ecology 45(3):138–143.

Termorshuizen, A. J., S. W. Moolenaar, A. H. M. Veeken, and W. J. Blok. 2004. The value of compost. Reviews in Environmental Science and Biotechnology 3(4):343–347.

Thies, J. E. and M. C. Rillig. 2009. Characteristics of biochar: Biological properties. Pages 85–105 *in* J. Lehmann, and S. Joseph, editors. Biochar for environmental management: Science and technology. Earthscan, London.

Thorne, M., L. Rhodes, and J. Cardina. 2013. Effectivity of arbuscular mycorrhizal fungi collected from reclaimed mine soil and tallgrass prairie. Open Journal of Ecology 3(3):224–233.

Tiessen, H., J. O. Moir, and M. R. Carter. 1993. Characterization of available P by sequential extraction. Pages 75–86 *in* M. R. Carter, editor. Soil sampling and methods of analysis. Lewis Publishers, Boca Raton, FL.

Tilman, D. and J. A. Downing. 1994. Biodiversity and stability in grasslands. Nature 367(6461):363–365.

Tilman, D., J. Knops, D. Wedin, P. Reich, M. Ritchie, and E. Siemann. 1997. The influence of functional diversity and composition on ecosystem processes. Science 277:1300–1302.

Tremblay, A. and J. Ransijn. 2013. LMERConvenienceFunctions: A suite of functions to back–fit fixed effects and forward–fit random effects, as well as other miscellaneous functions. R package version 2.5. Available at http://CRAN.R-project.org/package=LMERConvenienceFunctions .

Tscherko, D., J. Rustemeier, A. Richter, W. Wanek, and E. Kandeler. 2003. Functional diversity of the soil microflora in primary succession across two glacier forelands in the Central Alps. European Journal of Soil Science 54(4):685–696.

van der Heijden, M. G., R. D. Bardgett, and N. M. van Straalen. 2008. The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecology Letters 11(3):296–310.

van der Heijden, M. G. A., J. N. Klironomos, M. Ursic, P. Moutoglis, R. Streitwolf-Engel, T. Boller, A. Wiemken, and I. R. Sanders. 1998. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. Nature 396:69–72.

van Veen, J. A. and E. A. Paul. 1979. Conversion of biovolume measurements of soil organisms, grown under various moisture tensions, to biomass and their nutrient content. Applied and Environmental Microbiology 37(4):686–692.

Verheijen, F. G. A., E. R. Graber, N. Ameloot, A. C. Bastos, S. Sohi, and H. Knicker. 2014. Biochars in soils: New insights and emerging research needs. European Journal of Soil Science 65(1):22–27.

Visser, S. and D. Parkinson. 1992. Soil biological criteria as indicators of soil quality: Soil microorganisms. American Journal of Alternative Agriculture 7(1):33–37.

Vittoz, P. and A. Guisan. 2007. How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. Journal of Vegetation Science 18(3):413–422.

Wahba, M. M. and K. M. Darwish. 2008. Micro–morphological changes of sandy soils through the application of compost manure. Journal of Applied Biological Sciences 2(3):95–98.

Wali, M. K. 1999. Ecological succession and the rehabilitation of disturbed terrestrial ecosystems. Plant and Soil 213(1):195–220.

Walker, L. R., J. Walker, and R. del Moral. 2007. Forging a new alliance between succession and restoration. Pages 1–18 *in* L. R. Walker, J. Walker, and R. J. Hobbs, editors. Linking restoration and ecological succession. Springer New York, New York.

Wallwork, J. A. 1976. The distribution and diversity of soil fauna. Academic Press New York.

Walter, I., F. Martinez, and G. Cuevas. 2006. Plant and soil responses to the application of composted MSW in a degraded, semiarid shrubland in central Spain. Compost Science and Utilization 14(2):147–154.

Wardle, D. A. 1999. How soil food webs make plants grow. Trends in Ecology and Evolution 14(11):418–420.

Wardle, D. A. 2006. The influence of biotic interactions on soil biodiversity. Ecology Letters 9(7):870–886.

Wardle, D. A., G. W. Yeates, R. N. Watson, and K. S. Nicholson. 1995. The detritus food–web and the diversity of soil fauna as indicators of disturbance regimes in agro–ecosystems. Pages 35–43 *in* H. P. Collins, G. P. Robertson, and M. J. Klug, editors. The significance and regulation of soil biodiversity. Springer Netherlands.

Wardle, D. A., O. Zackrisson, and M. C. Nilsson. 1998. The charcoal effect in boreal forests: Mechanisms and ecological consequences. Oecologia 115(3):419–426.

Wardle, D. A., K. I. Bonner, and G. M. Barker. 2002. Linkages between plant litter decomposition, litter quality, and vegetation responses to herbivores. Functional Ecology 16(5):585–595.

Wardle, D. A., R. D. Bardgett, J. N. Klironomos, H. Setälä, W. H. van der Putten, and D. H. Wall. 2004. Ecological linkages between aboveground and belowground biota. Science 304(5677):1629–1633.

Warnock, D. D., J. Lehmann, T. W. Kuyper, and M. C. Rillig. 2007. Mycorrhizal responses to biochar in soil – concepts and mechanisms. Plant and Soil 300(1):9–20.

Warnock, D. D., D. L. Mummey, B. McBride, J. Major, J. Lehmann, and M. C. Rillig. 2010. Influences of non–herbaceous biochar on arbuscular mycorrhizal fungal abundances in roots and soils: Results from growth–chamber and field experiments. Applied Soil Ecology 46(3):450–456.

Weber, J., A. Karczewska, J. Drozd, M. Licznar, S. Licznar, E. Jamroz, and A. Kocowicz. 2007. Agricultural and ecological aspects of a sandy soil as affected by the application of municipal solid waste composts. Soil Biology and Biochemistry 39(6):1294–1302.

Wold, S., M. Sjóström, and L. Eriksson. 2001. PLS–regression: A basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58(2):109–130.

Wolkowski, R. P. 2003. Nitrogen management considerations for landspreading municipal solid waste compost. Journal of Environmental Quality 32(5):1844–1850.

Wortmann, C. S. and D. T. Walters. 2007. Residual effects of compost and plowing on phosphorus and sediment in runoff. Journal of Environmental Quality 36(5):1521–1527.

Wu, F. Y., Y. L. Bi, and M. H. Wong. 2009. Dual inoculation with an arbuscular mycorrhizal fungus and *Rhizobium* to facilitate the growth of Alfalfa on coal mine substrates. Journal of Plant Nutrition 32(5):755–771.

Xu, G., L. L. Wei, J. N. Sun, H. B. Shao, and S. X. Chang. 2013. What is more important for enhancing nutrient bioavailability with biochar application into a sandy soil: Direct or indirect mechanism? Ecological Engineering 52:119–124.

Yeates, G. W., T. Bongers, R. G. M. De Goede, D. W. Freckman, and S. S. Georgieva. 1993. Feeding habits in soil nematode families and generaan outline for soil ecologists. Journal of Nematology 25(3):315.

Yeniay, O. and A. Goktas. 2002. A comparison of partial least squares regression with other prediction methods. Hacettepe Journal of Mathematics and Statistics 31(99):99–111.

Zhao, J. and D. A. Neher. 2013. Soil nematode genera that predict specific types of disturbance. Applied Soil Ecology 64:135–141.

Zuur, A. F., E. N. Ieno, and G. M. Smith. 2007. Analysing ecological data. Springer New York, New York.

Appendices

Appendix A

R Code for AM Fungal Plant Plug Root Colonization in the Plant Plug Trial

```
library(ggplot2)

col.df <- read.csv("C:\\Users\\Ohsowski\\Documents\\PhD\\Dissertation\\Data\\
    AMF Colonization\\Plant Plug AMF Colonization.csv")
names(col.df) <- c("abb", "fun", "rep", "myco", "ac", "vc")


col.df$tot <- col.df$ac + col.df$vc

colMn.df <- aggregate(col.df$tot, by=list(col.df$abb, col.df$myco, col.df$fun)
    , FUN = mean, na.rm = TRUE)

names(colMn.df) <- c("abb", "myco", "fun", "totMn")


colTotSD.df <- aggregate(col.df$tot, by=list(col.df$abb, col.df$myco, col.df$
    fun), FUN = sd, na.rm = TRUE)

names(colTotSD.df) <- c("abb", "myco", "fun", "totSD")

colTotSD.df$totSE <- colTotSD.df$totSD / sqrt(10)

final.df <- merge(colMn.df, colTotSD.df, by = c("abb", "myco", "fun"))
```

```
# Graphs
graph <- ggplot(final.df, aes(x = factor(abb), y = totMn, fill = myco,))


limits <- aes(ymax = totMn + totSE, ymin = totMn - totSE)

dodge <- position_dodge(width = 0.9)


legendLabels <- c("No Inoc", "Inoc")
legendBreaks <- c("N", "Y")


graph_output <- graph +
  theme_bw() +
  geom_bar(position = dodge, stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge) +

  labs(x = "Plant Species",
       y = "Total % Colonization") +
  scale_fill_grey(breaks = legendBreaks, labels = legendLabels, start = 0.3,
                  end = 0.75) +
  facet_grid(~ fun) +
  theme(panel.grid.major = element_line(colour = 'grey85')) +
  theme(panel.grid.minor = element_line(colour = 'grey85')) +
  theme(legend.position = c(0.88, 0.88)) +
  theme(axis.text.x = element_text(size = 13)) +
  theme(axis.text.y = element_text(size = 13)) +
  theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
  theme(axis.title.y = element_text(size = 12, angle = 90)) +
  theme(legend.title = element_blank()) +
  theme(strip.background = element_rect(fill = 'grey85')) +
  theme(strip.text.x = element_text(face = 'bold', size = 14))


graph_output
```

Appendix B

R Code for AM Fungal Root Colonization in the Plant Plug Trial Field Plots

```
library(RODBC)
library(ggplot2)
library(glmmADMB)

# Data Transformation Section

Cube.Tns <- function (x) { x ^ 3 }

Square.Tns <- function (x) { x ^ 2 }

Raw.Tns <- function (x) { x }

Sqrt.Tns <- function (x) { sqrt(x) }

Log.Tns <- function (x) { log10(x + 1) }

RecipRoot.Tns <- function (x) { -1 / sqrt(x) }

Recip.Tns <- function (x) { -1 / (x) }

InvSquare.Tns <- function (x) { -1 / (x ^ 2) }

# Back Transformations

unscale.fn <- function(x){ as.data.frame(unscale(x, unscale.object)) }

backLog.Tns <- function (x){ 10 ^ (x) - 1 }

backSqrt.Tns <- function (x){ (x) ^ 2 }

backRaw.Tns <- function (x){ (x) }
```

```
# Prediction Data Centered and Scaled to Unity

## This unity function compares the collected test plants to the plsr data
    analyzed
## x = test variable data; y = main data; required to have same mean /
    variance

PredU <- function (x,y){
center <- x - mean(y)
sdCenter <- sd(y - mean(y))
center / sdCenter
}

# Pearson's Method (Parametric Test)

PearsonsMethod <- function (x) {
cor(x, use = "complete.obs", method = "pearson")
}

# Spearman's Method (Non-Parametric Test)

SpearmansMethod <- function (x,y) {
cor(x, use = "complete.obs", method = "spearman")
}

# Plot Information Data Frame

channel <- odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/
    Data/dissertation_data_exp_1_13aug2")
plotInfoCol.df <- sqlQuery(channel, "SELECT plot, biochar_rate, compost_rate,
    amf, hgt, treatment FROM DATA_Q WHERE harvestOne = FALSE")
close(channel)


# Training Data, Sum and Mean Stem Length Organization

channel <- odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/
    Data/dissertation_data_exp_1_13aug2")
ac.df <- sqlQuery(channel, "SELECT plot, data, dataType, season FROM AMF WHERE
     dataType = 'acCol' AND harvestOne = FALSE")
vc.df <- sqlQuery(channel, "SELECT plot, data, dataType, season FROM AMF WHERE
     dataType = 'vcCol' AND harvestOne = FALSE")
close(channel)

totCol.df <- merge(ac.df, vc.df, by = c("plot", "season"))
totCol.df$tot <- totCol.df$data.x + totCol.df$data.y
```

```
finalCol.df <- unique(merge(plotInfoCol.df, totCol.df, by = c("plot")))

finalCol.df <- data.frame(finalCol.df$plot, finalCol.df$treatment, finalCol.df
    $hgt, finalCol.df$compost_rate, finalCol.df$biochar_rate, finalCol.df$amf,
     finalCol.df$season)

finalCol.df$data.x, finalCol.df$data.y, finalCol.df$tot)
names(finalCol.df) <- c("plot", "treatment", "hgt", "compost_rate", "biochar_
    rate", "amf", "season", "arb", "ves", "tot")

finalCol.df$season <- ifelse(finalCol.df$season == "2011", "A", "B")

finalCol.df$plot <- as.factor(finalCol.df$plot)
finalCol.df$treatment <- as.factor(finalCol.df$treatment)
finalCol.df$amf <- as.factor(finalCol.df$amf)
finalCol.df$season <- as.factor(finalCol.df$season)
finalCol.df$rate <- finalCol.df$compost_rate + finalCol.df$biochar_rate

mean(finalCol.df$tot)
var(finalCol.df$tot)

finalCol.df$totT <- Raw.Tns(finalCol.df$tot)
qqnorm(finalCol.df$totT)
qqline(finalCol.df$totT)

TC.null <-  lmer(totT ~ 1 + (1|plot), data = finalCol.df)

TC1 <-  lmer(totT ~ treatment * season * hgt * amf + (1|plot), data = finalCol
    .df)
summary(TC1)
anova(TC.null, TC1)
relLik(TC.null, TC1)
mcp.fnc(TC1)

pamer.fnc(TC1)

TC2 <- update(TC1,.~. -treatment:season:amf:hgt)
summary(TC2)
anova(TC2)
anova(TC1, TC2)
relLik(TC1, TC2)

TC3 <- update(TC2,.~. -season:hgt:amf )
summary(TC3)
anova(TC3)
```

```
anova(TC2, TC3)
relLik(TC2, TC3)

TC4 <- update(TC3,.~. -treatment:season:hgt )
summary(TC4)
anova(TC4)
anova(TC3, TC4)
relLik(TC3, TC4)

# FINAL MODEL
TC5 <- update(TC4,.~. -treatment:hgt:amf  )
summary(TC5)
anova(TC5)
anova(TC4, TC5)
relLik(TC4, TC5)

anova(TC5)
summary(TC5)
pamer.fnc(TC5)
TC5.ph <- mcposthoc.fnc(model = TC5, var = list(ph1 = "treatment"))
summary(TC5.ph)


#Wireframe Graph
scatter3d(finalCol.df$totT ~ finalCol.df$hgt + finalCol.df$rate,
           bg="white", axis.scales=TRUE, grid=TRUE, id.method="identify",
          ellipsoid=FALSE, xlab="biochar_rate", ylab="cover", zlab="compost_
    rate",
          groups = as.factor(finalCol.df$amf))

wireframe(totT ~ rate + hgt, data=finalCol.df, xlab = "Compost Rate", ylab = "
    Biochar Rate)",
  drape = TRUE,
  colorkey = TRUE)

p <- wireframe(variable ~ rate * hgt, data=final.df)
npanel <- c(4, 2)
rotx <- c(-50, -80)
rotz <- seq(30, 300, length = npanel[1]+1)
update(p[rep(1, prod(npanel))], layout = npanel,
    panel = function(..., screen) {
        panel.wireframe(..., screen = list(z = rotz[current.column()],
                                           x = rotx[current.row()]))

# Graph Set-up
```

```
totCol <- aggregate(finalCol.df$tot, by = list(finalCol.df$season, finalCol.df
    $treatment, finalCol.df$amf), FUN = mean, na.rm = TRUE)

names(totCol) <- c("season", "treatment", "amf", "data")

totColSD <- aggregate(finalCol.df$tot, by = list(finalCol.df$season, finalCol.
    df$treatment, finalCol.df$amf), FUN = sd, na.rm = TRUE)

names(totColSD) <- c("season", "treatment", "amf", "sd")

totCol.gr <- merge(totCol, totColSD, by = c("season", "treatment", "amf"))


# GGPLOT Graphing

graph <- ggplot(data = totCol.gr, aes(x = treatment, y = data, fill = amf))

limits <- aes(ymax = totCol.gr$data + totCol.gr$sd, ymin = totCol.gr$data -
    totCol.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
  theme_bw() +

  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +
  facet_grid(. ~ season) +

  labs(x = "Carbon Amendment",
       y = "Estimated Biomass\n dry mass (g)") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output
```

Appendix C

R Code for Plant Plug Trial Biomass Predictions and Plot Mass Calculations

```
library(RODBC)
library(reshape)
library(tcltk)
library(tcltk2)
library(Hmisc)
library(pls)
library(DMwR)
library(ggplot2)
library(plyr)
library(scatterplot3d)
library(rgl, pos=4)
library(mgcv, pos=4)
library(MuMIn)


# Data Transformation Section

Cube.Tns <- function (x) { x ^ 3 }

Square.Tns <- function (x) { x ^ 2 }

Raw.Tns <- function (x) { x }

Sqrt.Tns <- function (x) { sqrt(x) }

Log.Tns <- function (x) { log10(x + 1) }

RecipRoot.Tns <- function (x) { -1 / sqrt(x) }

Recip.Tns <- function (x) { -1 / (x) }

InvSquare.Tns <- function (x) { -1 / (x ^ 2) }
```

```
# Back Transformations

unscale.fn <- function(x){ as.data.frame(unscale(x, unscale.object)) }

backLog.Tns <- function (x){ 10 ^ (x) - 1 }

backSqrt.Tns <- function (x){ (x) ^ 2 }

backRaw.Tns <- function (x){ (x) }

# Prediction Data Centered and Scaled to Unity

## This unity function compares the collected test plants to the plsr data
    analyzed
## x = test variable data; y = main data; required to have same mean /
    variance

PredU <- function (x,y){
  center <- x - mean(y)
  sdCenter <- sd(y - mean(y))
  center / sdCenter
  }

# Pearson's Method (Parametric Test)

PearsonsMethod <- function (x) {
  cor(x, use = "complete.obs", method = "pearson")
}

# Spearman's Method (Non-Parametric Test)

SpearmansMethod <- function (x,y) {
  cor(x, use = "complete.obs", method = "spearman")
}

################################
## 2011 Plant Prediction Data ##
################################

# Plot Information Data Frame
channel <- odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/
    Data/dissertation_data_exp_1_13aug2")
```

```
plot11Info.df <- sqlQuery(channel, "SELECT plot, position, funGroup, spAbbr,
    amf, treatment, biochar_rate, compost_rate, season, hgt FROM DATA_Q WHERE
    season = '2011'")
close(channel)


# Training Data (Variable Creation)

channel <-
odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/Data/
    dissertation_data_exp_1_13aug2")


mainTrain11.df <- sqlQuery(channel, "SELECT plot, position, data, dataType,
    spAbbr, plsr, test FROM DATA_Q WHERE season = '2011' AND test = 0 AND plsr
     = 1")


stemLenTrain11.df <- sqlQuery(channel, "SELECT plot, position, data, spAbbr,
    plsr FROM DATA_Q WHERE season = '2011' AND dataType = 'stemLen' AND plsr =
     1 AND test = 0")

## Calculates the Mean of stem length for the training plant

stemLenMeanTrain11.df <- aggregate(stemLenTrain11.df$data, by=list(
    stemLenTrain11.df$plot, stemLenTrain11.df$position, stemLenTrain11.df$
    spAbbr, stemLenTrain11.df$plsr), FUN = mean, na.rm = TRUE)

names(stemLenMeanTrain11.df) <- c("plot", "position", "spAbbr", "plsr", "
    stemLenMean")


## Sum of stem lengths for prediction plants

stemLenSumTrain11.df <- aggregate(stemLenTrain11.df$data, by=list(
    stemLenTrain11.df$plot, stemLenTrain11.df$position, stemLenTrain11.df$
    spAbbr, stemLenTrain11.df$plsr), FUN = sum, na.rm = TRUE)

names(stemLenSumTrain11.df) <- c("plot","position", "spAbbr","plsr", "
    stemLenSum")

## Sum of stem lengths for prediction plants

stemLenCountTrain11.df <- aggregate(stemLenTrain11.df$data, by=list(
    stemLenTrain11.df$plot, stemLenTrain11.df$position, stemLenTrain11.df$
    spAbbr, stemLenTrain11.df$plsr), FUN = function(x) c(count = length(x)))
```

```
names(stemLenCountTrain11.df) <- c("plot","position", "spAbbr","plsr", "
    stemLenCount")


## Mean of plant biomass for prediction plants

massTrain11.df <- sqlQuery(channel, "SELECT plot, position, data, spAbbr, plsr
     FROM DATA_Q WHERE season = '2011' AND dataType = 'mass' AND plsr = 1 AND
    test = 0")

massMeanTrain11.df <-aggregate(massTrain11.df$data, by=list(massTrain11.df$
    plot, massTrain11.df$position, massTrain11.df$spAbbr, massTrain11.df$plsr)
    , FUN = mean, na.rm = TRUE)

names(massMeanTrain11.df) <- c("plot","position", "spAbbr","plsr", "mass")

close(channel)

# Data (Variable Creation) -

channel <-
  odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/Data/
    dissertation_data_exp_1_13aug2")


data11.df <- sqlQuery(channel, "SELECT plot, position, data, dataType, spAbbr,
     core FROM DATA_Q WHERE season = '2011' AND core = 1 AND harvestOne = 0")


stemLenData11.df <- sqlQuery(channel, "SELECT plot, position, data, spAbbr,
    core FROM DATA_Q WHERE season = '2011' AND dataType = 'stemLen' AND core =
     1 AND harvestOne = 0")
close(channel)

stemLenMeanData11.df <-aggregate(stemLenData11.df$data, by=list(stemLenData11.
    df$plot, stemLenData11.df$position, stemLenData11.df$spAbbr, stemLenData11
    .df$core), FUN = mean, na.rm = TRUE)

names(stemLenMeanData11.df) <- c("plot","position", "spAbbr","core", "
    stemLenMean")


stemLenSumData11.df <-aggregate(stemLenData11.df$data, by=list(stemLenData11.
    df$plot, stemLenData11.df$position, stemLenData11.df$spAbbr, stemLenData11
    .df$core), FUN = sum)
```

```
names(stemLenSumData11.df) <- c("plot","position", "spAbbr","core", "
    stemLenSum")


## Sum of stem lengths for prediction plants

stemLenCountData11.df <- aggregate(stemLenData11.df$data, by=list(
    stemLenData11.df$plot, stemLenData11.df$position, stemLenData11.df$spAbbr,
     stemLenData11.df$core), FUN = function(x) c(count = length(x)))

names(stemLenCountData11.df) <- c("plot","position", "spAbbr","core", "
    stemLenCount")




# AG Training Data Import
visAG <- subset(mainTrain11.df, spAbbr == 'AG')

agMassMean <- subset(massMeanTrain11.df, spAbbr == 'AG')


agWPHgt <- subset(mainTrain11.df, spAbbr == 'AG'& dataType == 'sWPHgt', select
     = c(plot, position, spAbbr, plsr, data))

names(agWPHgt) <- c("plot","position", "spAbbr", "plsr", "WPHgt")


agLv4Hgt <- subset(mainTrain11.df, spAbbr == 'AG'& dataType == 's4LvHgt',
    select = c(plot, position, spAbbr, plsr, data))

names(agLv4Hgt) <- c("plot","position", "spAbbr", "plsr", "lv4Hgt")


agTrain11.tmp <- merge(merge(agWPHgt, agLv4Hgt, by = c("plot", "position", "
    spAbbr","plsr")), agMassMean, by = c("plot", "position", "spAbbr", "plsr")
    )

# AG Training Transformation Section

agTrain11.tmp$massT <- Log.Tns(agTrain11.tmp$mass)
agTrain11.tmp$lv4HgtT <- Log.Tns(agTrain11.tmp$lv4Hgt)
agTrain11.tmp$WPHgtT <- Sqrt.Tns(agTrain11.tmp$WPHgt)
```

```
agTrain11.tmp$massUT <-  scale(agTrain11.tmp$massT)
agTrain11.tmp$lv4HgtUT <- scale(agTrain11.tmp$lv4HgtT)
agTrain11.tmp$WPHgtUT <- scale(agTrain11.tmp$WPHgtT)

agTrain11.sc <- agTrain11.tmp [,-1:-10]

agTrain11UTSq <- as.data.frame(
apply(agTrain11.sc[,-1], 2,
function(x){ x^2 })
); names(agTrain11UTSq) <- paste(
names(agTrain11UTSq), "sq",
sep = "."
)


agTrain11UTCb <- as.data.frame(
apply(agTrain11.sc[,-1], 2,
function(x){ x^3 })
); names(agTrain11UTCb) <- paste(
names(agTrain11UTCb), "cb",
sep = "."
)

ag11.df <- as.data.frame(cbind(agTrain11.sc, agTrain11UTSq, agTrain11UTCb))


# AG AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

# ag11LMF <- lm(massUT ~
# lv4HgtUT + lv4HgtUT.sq + lv4HgtUT.cb +
# WPHgtUT + WPHgtUT.sq + WPHgtUT.cb,
# data = ag11.df
# )

# ag11Subset <- expression(
# (lv4HgtUT | !lv4HgtUT.sq) & (lv4HgtUT & lv4HgtUT.sq | !lv4HgtUT.cb) &
# (WPHgtUT | !WPHgtUT.sq) & (WPHgtUT & WPHgtUT.sq | !WPHgtUT.cb)
# )


# AIC Dredge
# ag11.AIC.dredge <-dredge(
# ag11LMF,
```

```
# subset = ag11Subset,
# rank = AIC
# )

# ag11.AIC.top <- subset(
# ag11.AIC.dredge,
# delta < 2
# )

# BIC Dredge
# ag11.BIC.dredge <-dredge(
# ag11LMF,
# subset = ag11Subset,
# rank = BIC
# )

# ag11.BIC.top <- subset(
# ag11.BIC.dredge,
# delta < 2
# )


# Global model call: lm(formula = massUT ~ lv4HgtUT + lv4HgtUT.sq + lv4HgtUT.
    cb +
# WPHgtUT + WPHgtUT.sq + WPHgtUT.cb, data = ag11.df)
# -
# Model selection table
# (Int)  l4H  WPH df  logLik  BIC delta weight
# 10 -3.978e-17 0.1875 0.7806  4 -17.427 49.2  0.00  0.567
# 9 1.462e-16  0.9067  3 -19.488 49.7  0.54  0.433

# AG Training PLSR Section -

agUT11.PLSR <- plsr(ag11.df$massUT ~ lv4HgtUT + ag11.df$WPHgtUT,
  ncomp = 2, data = ag11.df, validation = "CV",
  method = "oscorespls")

summary(agUT11.PLSR)

# AG Data Import --

agWPHgt <- subset(data11.df, spAbbr == 'AG'& dataType == 'sWPHgt',
  select = c(plot, position, spAbbr, core, data))

names(agWPHgt) <- c("plot","position", "spAbbr", "core", "WPHgt")
```

```
agLv4Hgt <- subset(data11.df, spAbbr == 'AG'& dataType == 's4LvHgt',
 select = c(plot, position, spAbbr, core, data))

names(agLv4Hgt) <- c("plot","position", "spAbbr", "core", "lv4Hgt")


agAll11.tmp <- merge(agWPHgt, agLv4Hgt, by = c("plot", "position", "spAbbr", "
    core"))

# AG Data Transformation -

agUTAll11.df <- agAll11.tmp

agUTAll11.df$lv4HgtT <- Log.Tns(agUTAll11.df$lv4Hgt)

agUTAll11.df$WPHgtT <- Sqrt.Tns(agUTAll11.df$WPHgt)


agUTAll11.df$lv4HgtUT <-PredU(agUTAll11.df$lv4HgtT, agTrain11.tmp$lv4HgtT)

agUTAll11.df$WPHgtUT <-PredU(agUTAll11.df$WPHgtT, agTrain11.tmp$WPHgtT)

# AG Data Prediction -

agUTAll11.df$mx <- as.matrix(agUTAll11.df [c(9:10)])

agUTAll11.df$predictMass <- predict(agUT11.PLSR, ncomp = 2, newdata =
    agUTAll11.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- agTrain11.tmp$massUT
agUTAll11.df$massFinalT <- unscale.fn(agUTAll11.df$predictMass)

agUTAll11.df$massFinal <-backLog.Tns(agUTAll11.df$massFinalT)

agUTMass11.df <- data.frame(agUTAll11.df$plot, agUTAll11.df$spAbbr, agUTAll11.
    df$position, agUTAll11.df$massFinal)
names(agUTMass11.df) <- c("plot", "spAbbr", "position", "massFinal")

ag11Test.df <- merge(agTrain11.tmp, agUTMass11.df, by = c("plot", "position"))


##########################
```

```
## PREDICTION Accuracy ##
##########################

agCV <- 0.4359
RMSEag <- data.frame(agCV)

RMSEag$agBICUn  <- unscale(agCV, agTrain11.tmp$massUT)
RMSEag$agBIC <- backLog.Tns(RMSEag$agBICUn)

ag11TEST.df <- data.frame(predict(agUT11.PLSR, ncomp = 2, newdata = ag11.df))
names(ag11TEST.df) <- c("massUT")

unscale.object <- agTrain11.tmp$massUT

ag11TEST.df$massFinalT <- unscale(ag11TEST.df$massUT, agTrain11.tmp$massUT)

ag11TEST.df$massFinal <- backLog.Tns(ag11TEST.df$massFinalT)

ag11TEST.df$massREAL <- agTrain11.tmp$mass

ag11TEST.df$subtract <- ag11TEST.df$massFinal - ag11TEST.df$massREAL
mean(ag11TEST.df$subtract)
sd(ag11TEST.df$subtract)

ag11LM <- lm(ag11TEST.df$massFinal ~ ag11TEST.df$massREAL -1 ,  offset = 1.00
    * ag11TEST.df$massREAL)
summary (ag11LM)

mean(ag11TEST.df$massREAL)
sd(ag11TEST.df$massREAL)

# PV Training Data Import -

pvMassMean <- subset(massMeanTrain11.df, spAbbr == 'PV')

pvWPHgt <- subset(mainTrain11.df, spAbbr == 'PV'& dataType == 'sWPHgt',
  select = c(plot, position, spAbbr, plsr, data))

names(pvWPHgt) <- c("plot","position", "spAbbr", "plsr", "WPHgt")


pvBCirc <- subset(mainTrain11.df, spAbbr == 'PV'& dataType == 'bCirc',
 select = c(plot, position, spAbbr, plsr, data))

names(pvBCirc) <- c("plot","position", "spAbbr", "plsr", "bCirc")
```

```
pvTrain11.tmp <-
  merge(
  merge(pvWPHgt, pvBCirc, by = c("plot", "position", "spAbbr","plsr")),
  pvMassMean, by = c("plot", "position", "spAbbr", "plsr"))

# PV Training Transformation --

pvTrain11.tmp$massT <- Log.Tns(pvTrain11.tmp$mass)
pvTrain11.tmp$bCircT <- Raw.Tns(pvTrain11.tmp$bCirc)
pvTrain11.tmp$WPHgtT <- Sqrt.Tns(pvTrain11.tmp$WPHgt)


pvTrain11.tmp$massUT <-  scale(pvTrain11.tmp$massT)
pvTrain11.tmp$bCircUT <- scale(pvTrain11.tmp$bCircT)
pvTrain11.tmp$WPHgtUT <- scale(pvTrain11.tmp$WPHgtT)

pvTrain11.sc <- pvTrain11.tmp [,-1:-10]

pvTrain11UTSq <- as.data.frame(
apply(pvTrain11.sc[,-1], 2,
function(x){ x^2 })
); names(pvTrain11UTSq) <- paste(
names(pvTrain11UTSq), "sq",
sep = "."
)


pvTrain11UTCb <- as.data.frame(
apply(pvTrain11.sc[,-1], 2,
function(x){ x^3 })
); names(pvTrain11UTCb) <- paste(
names(pvTrain11UTCb), "cb",
sep = "."
)

pv11.df <- as.data.frame(cbind(pvTrain11.sc, pvTrain11UTSq, pvTrain11UTCb))



# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

pv11LMF <- lm(massUT ~
bCircUT + bCircUT.sq + bCircUT.cb +
```

```
WPHgtUT + WPHgtUT.sq + WPHgtUT.cb,
data = pv11.df
)

pv11Subset <- expression(
(bCircUT | !bCircUT.sq) & (bCircUT & bCircUT.sq | !bCircUT.cb) &
(WPHgtUT | !WPHgtUT.sq) & (WPHgtUT & WPHgtUT.sq | !WPHgtUT.cb)
)

# AIC Dredge
# pv11.AIC.dredge <-dredge(
# pv11LMF,
# subset = pv11Subset,
# rank = AIC
# )

# pv11.AIC.top <- subset(
# pv11.AIC.dredge,
# delta < 2
# )

# BIC Dredge
# pv11.BIC.dredge <-dredge(
# pv11LMF,
# subset = pv11Subset,
# rank = BIC
# )

# pv11.BIC.top <- subset(
# pv11.BIC.dredge,
# delta < 2
# )

# Global model call: lm(formula = massUT ~ bCircUT + bCircUT.sq + bCircUT.cb +
    WPHgtUT +
  # WPHgtUT.sq + WPHgtUT.cb, data = pv11.df)
# -
# Model selection table
  # (Int)  bCU  WPH df  logLik  BIC delta weight
# 10 -3.927e-16 0.3363 0.6606  4 -20.805 55.7 0  1

# PV Training PLSR  --

pvUT11.PLSR <- plsr(pv11.df$massUT ~ pv11.df$bCircUT + pv11.df$WPHgtUT,
  ncomp = 2, data = pv11.df, validation = "CV",
  method = "oscorespls")
```

```
summary(pvUT11.PLSR)
# PV Data Import --

pvWPHgt <- subset(data11.df, spAbbr == 'PV'& dataType == 'sWPHgt',
  select = c(plot, position, spAbbr, core, data))

names(pvWPHgt) <- c("plot","position", "spAbbr", "core", "WPHgt")


pvBCirc <- subset(data11.df, spAbbr == 'PV'& dataType == 'bCirc',
 select = c(plot, position, spAbbr, core, data))

names(pvBCirc) <- c("plot","position", "spAbbr", "core", "bCirc")


pvAll11.tmp <- merge(pvWPHgt, pvBCirc, by = c("plot", "position", "spAbbr", "
    core"))

# PV Data Transformation --

pvUTAll11.df <- pvAll11.tmp

pvUTAll11.df$bCircT <- Raw.Tns(pvUTAll11.df$bCirc)
pvUTAll11.df$WPHgtT <- Sqrt.Tns(pvUTAll11.df$WPHgt)


pvUTAll11.df$bCircUT <-PredU(pvUTAll11.df$bCircT, pvTrain11.tmp$bCircT)
pvUTAll11.df$WPHgtUT <-PredU(pvUTAll11.df$WPHgtT, pvTrain11.tmp$WPHgtT)

# PV Data Prediction -

pvUTAll11.df$mx <- as.matrix(pvUTAll11.df [c(9:10)])

pvUTAll11.df$predictMass <- predict(pvUT11.PLSR, ncomp = 2, newdata =
    pvUTAll11.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- pvTrain11.tmp$massUT
pvUTAll11.df$massFinalT <- unscale.fn(pvUTAll11.df$predictMass)

pvUTAll11.df$massFinal <-backLog.Tns(pvUTAll11.df$massFinalT)
```

```
pvUTMass11.df <- data.frame(pvUTAll11.df$plot, pvUTAll11.df$spAbbr, pvUTAll11.
    df$position, pvUTAll11.df$massFinal)
names(pvUTMass11.df) <- c("plot", "spAbbr", "position", "massFinal")

#########################
## PREDICTION Accuracy ##
#########################

pvCV <- 0.4870
RMSEpv <- data.frame(pvCV)

RMSEpv$pvBICUn  <- unscale(pvCV, pvTrain11.tmp$massUT)
RMSEpv$pvBIC <- backLog.Tns(RMSEpv$pvBICUn)

pv11TEST.df <- data.frame(predict(pvUT11.PLSR, ncomp = 2, newdata = pv11.df))
names(pv11TEST.df) <- c("massUT")

unscale.object <- pvTrain11.tmp$massUT

pv11TEST.df$massFinalT <- unscale(pv11TEST.df$massUT, pvTrain11.tmp$massUT)

pv11TEST.df$massFinal <- backLog.Tns(pv11TEST.df$massFinalT)

pv11TEST.df$massREAL <- pvTrain11.tmp$mass

pv11TEST.df$subtract <- pv11TEST.df$massFinal - pv11TEST.df$massREAL
mean(pv11TEST.df$subtract)
sd(pv11TEST.df$subtract)

pv11LM <- lm(pv11TEST.df$massFinal ~ pv11TEST.df$massREAL -1 ,  offset = 1.00
    * pv11TEST.df$massREAL)
summary (pv11LM)

mean(pv11TEST.df$massREAL)
sd(pv11TEST.df$massREAL)


# LC Training Data Import -

lcMassMean <- subset(massMeanTrain11.df, spAbbr == 'LC')

lcLv4Hgt <- subset(mainTrain11.df, spAbbr == 'LC'& dataType == 's4LvHgt',
 select = c(plot, position, spAbbr, plsr, data))

names(lcLv4Hgt) <- c("plot","position", "spAbbr", "plsr", "lv4Hgt")
```

```
lcLeafNo <- subset(mainTrain11.df, spAbbr == 'LC'& dataType == 'leafNo',
 select = c(plot, position, spAbbr, plsr, data))

names(lcLeafNo) <- c("plot","position", "spAbbr", "plsr", "leafNo")


lcTrain11.tmp <-
  merge(
  merge(lcLv4Hgt, lcLeafNo, by = c("plot", "position", "spAbbr","plsr")),
  lcMassMean, by = c("plot", "position", "spAbbr", "plsr"))


# LC Training Data Transformation --

lcTrain11.tmp$massT <- Log.Tns(lcTrain11.tmp$mass)
lcTrain11.tmp$leafNoT <- Sqrt.Tns(lcTrain11.tmp$leafNo)
lcTrain11.tmp$lv4HgtT <- Raw.Tns(lcTrain11.tmp$lv4Hgt)

lcTrain11.tmp$massUT <-  scale(lcTrain11.tmp$massT)
lcTrain11.tmp$leafNoUT <- scale(lcTrain11.tmp$leafNoT)
lcTrain11.tmp$lv4HgtUT <- scale(lcTrain11.tmp$lv4HgtT)

lcTrain11.sc <- lcTrain11.tmp [,-1:-10]

lcTrain11UTSq <- as.data.frame(
apply(lcTrain11.sc[,-1], 2,
function(x){ x^2 })
); names(lcTrain11UTSq) <- paste(
names(lcTrain11UTSq), "sq",
sep = "."
)


lcTrain11UTCb <- as.data.frame(
apply(lcTrain11.sc[,-1], 2,
function(x){ x^3 })
); names(lcTrain11UTCb) <- paste(
names(lcTrain11UTCb), "cb",
sep = "."
)

lc11.df <- as.data.frame(cbind(lcTrain11.sc, lcTrain11UTSq, lcTrain11UTCb))
```

```
# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

lc11LMF <- lm(massUT ~
leafNoUT + leafNoUT.sq + leafNoUT.cb +
lv4HgtUT + lv4HgtUT.sq + lv4HgtUT.cb,
data = lc11.df
)

lc11Subset <- expression(
(leafNoUT | !leafNoUT.sq) & (leafNoUT & leafNoUT.sq | !leafNoUT.cb) &
(lv4HgtUT | !lv4HgtUT.sq) & (lv4HgtUT & lv4HgtUT.sq | !lv4HgtUT.cb)
)

# AIC Dredge
# lc11.AIC.dredge <-dredge(
# lc11LMF,
# subset = lc11Subset,
# rank = AIC
# )

# lc11.AIC.top <- subset(
# lc11.AIC.dredge,
# delta < 2
# )

# BIC Dredge
# lc11.BIC.dredge <-dredge(
# lc11LMF,
# subset = lc11Subset,
# rank = BIC
# )

# lc11.BIC.top <- subset(
# lc11.BIC.dredge,
# delta < 2
# )

# Global model call: lm(formula = massUT ~ leafNoUT + leafNoUT.sq + leafNoUT.
    cb +
  # lv4HgtUT + lv4HgtUT.sq + lv4HgtUT.cb, data = lc11.df)
# -
# Model selection table
 # (Int)  lNU lNU.cb  lNU.sq  l4H l4H.sq df  logLik  BIC delta weight
```

```
# 48 0.08204 0.3404 0.2421 -0.2994 0.1593 0.1884  7 -16.498 57.9 0  1

# LC Training PLSR --

lcUT11.PLSR <- plsr(massUT ~ leafNoUT + leafNoUT.sq + leafNoUT.cb +
          lv4HgtUT + lv4HgtUT.sq,
  ncomp = 5, data = lc11.df, validation = "CV",
  method = "oscorespls")

summary(lcUT11.PLSR)

# LC Data Import --

lcLv4Hgt <- subset(data11.df, spAbbr == 'LC'& dataType == 's4LvHgt',
 select = c(plot, position, spAbbr, core, data))

names(lcLv4Hgt) <- c("plot","position", "spAbbr", "core", "lv4Hgt")


lcLeafNo <- subset(data11.df, spAbbr == 'LC'& dataType == 'leafNo',
 select = c(plot, position, spAbbr, core, data))

names(lcLeafNo) <- c("plot","position", "spAbbr", "core", "leafNo")


lcAll11.tmp <- merge(lcLv4Hgt, lcLeafNo, by = c("plot", "position", "spAbbr",
    "core"))

# LC Data Transformation --

lcUTAll11.df <- lcAll11.tmp

lcUTAll11.df$leafNoT <- Sqrt.Tns(lcUTAll11.df$leafNo)
lcUTAll11.df$lv4HgtT <- Raw.Tns(lcUTAll11.df$lv4Hgt)

lcUTAll11.df$leafNoUT <-PredU(lcUTAll11.df$leafNoT, lcTrain11.tmp$leafNoT)
lcUTAll11.df$lv4HgtUT <-PredU(lcUTAll11.df$lv4HgtT, lcTrain11.tmp$lv4HgtT)

lcUTAll11.df$leafNoUT.sq <- (lcUTAll11.df$leafNoUT)^2
lcUTAll11.df$leafNoUT.cb <- (lcUTAll11.df$leafNoUT)^3

lcUTAll11.df$lv4HgtUT.sq <- (lcUTAll11.df$lv4HgtUT)^2

lcUTAll11.df <- data.frame(lcUTAll11.df$plot, lcUTAll11.df$spAbbr, lcUTAll11.
    df$position, lcUTAll11.df$leafNoUT, lcUTAll11.df$leafNoUT.sq , lcUTAll11.
    df$leafNoUT.cb,
```

```
                lcUTAll11.df$lv4HgtUT, lcUTAll11.df$lv4HgtUT.sq)

names(lcUTAll11.df) <- c("plot", "spAbbr", "position", "leafNoUT", "leafNoUT.
    sq" , "leafNoUT.cb", "lv4HgtUT", "lv4HgtUT.sq")

# LC Prediction --

lcUTAll11.df$mx <- as.matrix(lcUTAll11.df [,-1:-3])

lcUTAll11.df$predictMass <- predict(lcUT11.PLSR, ncomp = 2, newdata =
    lcUTAll11.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- lcTrain11.tmp$massUT
lcUTAll11.df$massFinalT <- unscale.fn(lcUTAll11.df$predictMass)

lcUTAll11.df$massFinal <-backLog.Tns(lcUTAll11.df$massFinalT)

lcUTMass11.df <- data.frame(lcUTAll11.df$plot, lcUTAll11.df$spAbbr, lcUTAll11.
    df$position, lcUTAll11.df$massFinal)
names(lcUTMass11.df) <- c("plot", "spAbbr", "position", "massFinal")



##########################
## PREDICTION Accuracy ##
##########################

lcCV <- 0.5734
RMSElc <- data.frame(lcCV)

RMSElc$lcBICUn  <- unscale(lcCV, lcTrain11.tmp$massUT)
RMSElc$lcBIC <- backLog.Tns(RMSElc$lcBICUn)

lc11TEST.df <- data.frame(predict(lcUT11.PLSR, ncomp = 2, newdata = lc11.df))
names(lc11TEST.df) <- c("massUT")

unscale.object <- lcTrain11.tmp$massUT

lc11TEST.df$massFinalT <- unscale(lc11TEST.df$massUT, lcTrain11.tmp$massUT)

lc11TEST.df$massFinal <- backLog.Tns(lc11TEST.df$massFinalT)

lc11TEST.df$massREAL <- lcTrain11.tmp$mass
```

```
lc11TEST.df$subtract <- lc11TEST.df$massFinal - lc11TEST.df$massREAL
mean(lc11TEST.df$subtract)
sd(lc11TEST.df$subtract)

lc11LM <- lm(lc11TEST.df$massFinal ~ lc11TEST.df$massREAL -1 ,  offset = 1.00
    * lc11TEST.df$massREAL)
summary (lc11LM)

mean(lc11TEST.df$massREAL)
sd(lc11TEST.df$massREAL)




# SL Training Data Import -

slStemLenMean <- subset(stemLenMeanTrain11.df, spAbbr == 'SL')

slStemLenSum <- subset(stemLenSumTrain11.df, spAbbr == 'SL')

slMassMean <- subset(massMeanTrain11.df, spAbbr == 'SL')



slTrain11.tmp <-
  merge(
    merge(slStemLenSum, slStemLenMean, by = c("plot", "position", "spAbbr","
    plsr")),
    slMassMean, by = c("plot", "position", "spAbbr", "plsr"))


# SL Training Data Transformation -

slTrain11.tmp$massT <- Log.Tns(slTrain11.tmp$mass)
slTrain11.tmp$stemLenSumT <- Log.Tns(slTrain11.tmp$stemLenSum)
slTrain11.tmp$stemLenMeanT <- Log.Tns(slTrain11.tmp$stemLenMean)


slTrain11.tmp$massUT <-  scale(slTrain11.tmp$massT)
slTrain11.tmp$stemLenSumUT <- scale(slTrain11.tmp$stemLenSumT)
slTrain11.tmp$stemLenMeanUT <- scale(slTrain11.tmp$stemLenMeanT)
```

```
slTrain11.sc <- slTrain11.tmp [,11:13]

slTrain11UTSq <- as.data.frame(
apply(slTrain11.sc[,-1], 2,
function(x){ x^2 })
); names(slTrain11UTSq) <- paste(
names(slTrain11UTSq), "sq",
sep = "."
)


slTrain11UTCb <- as.data.frame(
apply(slTrain11.sc[,-1], 2,
function(x){ x^3 })
); names(slTrain11UTCb) <- paste(
names(slTrain11UTCb), "cb",
sep = "."
)

sl11.df <- as.data.frame(cbind(slTrain11.sc, slTrain11UTSq, slTrain11UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

sl11LMF <- lm(massUT ~
stemLenSumUT + stemLenSumUT.sq + stemLenSumUT.cb +
stemLenMeanUT + stemLenMeanUT.sq + stemLenMeanUT.cb,
data = sl11.df
)

sl11Subset <- expression(
(stemLenSumUT | !stemLenSumUT.sq) & (stemLenSumUT & stemLenSumUT.sq | !
    stemLenSumUT.cb) &
(stemLenMeanUT | !stemLenMeanUT.sq) & (stemLenMeanUT & stemLenMeanUT.sq | !
    stemLenMeanUT.cb)
)

# AIC Dredge
# sl11.AIC.dredge <-dredge(
# sl11LMF,
# subset = sl11Subset,
# rank = AIC
```

```
# )

# sl11.AIC.top <- subset(
# sl11.AIC.dredge,
# delta < 2
# )

# BIC Dredge
# sl11.BIC.dredge <-dredge(
# sl11LMF,
# subset = sl11Subset,
# rank = BIC
# )

# sl11.BIC.top <- subset(
# sl11.BIC.dredge,
# delta < 2
# )

# Global model call: lm(formula = massUT ~ stemLenSumUT + stemLenSumUT.sq +
    stemLenSumUT.cb +
  # stemLenMeanUT + stemLenMeanUT.sq + stemLenMeanUT.cb, data = sl11.df)
# -
# Model selection table
 # (Int)  sLS df  logLik  BIC delta weight
# 9 -5.337e-17 0.7303  3 -42.047 95.2 0  1

# SL Training PLSR -

slUT11.PLSR <- plsr(sl11.df$massUT ~ sl11.df$stemLenSumUT + sl11.df$
    stemLenMeanUT +
          sl11.df$stemLenMeanUT.sq,
 ncomp = 3, data = sl11.df, validation = "CV",
 method = "oscorespls")
summary(slUT11.PLSR)

# SL Data Import -

slStemLenMean <- subset(stemLenMeanData11.df, spAbbr == 'SL')

slStemLenSum <- subset(stemLenSumData11.df, spAbbr == 'SL')

slAll11.tmp <- merge(slStemLenSum, slStemLenMean, by = c("plot", "position", "
    spAbbr", "core"))
# SL Data Transformation -
```

```
slUTAll11.df <- slAll11.tmp

slUTAll11.df$stemLenSumT <- Log.Tns(slUTAll11.df$stemLenSum)
slUTAll11.df$stemLenMeanT <- Log.Tns(slUTAll11.df$stemLenMean)

slUTAll11.df$stemLenSumUT <-PredU(slUTAll11.df$stemLenSumT, slTrain11.tmp$
    stemLenSumT)
slUTAll11.df$stemLenMeanUT <-PredU(slUTAll11.df$stemLenMeanT, slTrain11.tmp$
    stemLenMeanT)
slUTAll11.df$stemLenMeanUT.sq <- (slUTAll11.df$stemLenMeanUT)^2

# SL Prediction -


slUTAll11.df$mx <- as.matrix(slUTAll11.df [c(9:11)])

slUTAll11.df$predictMass <- predict(slUT11.PLSR, ncomp = 3, newdata =
    slUTAll11.df$mx)


## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- slTrain11.tmp$massUT
slUTAll11.df$massFinalT <- unscale.fn(slUTAll11.df$predictMass)

slUTAll11.df$massFinal <-backLog.Tns(slUTAll11.df$massFinalT)

slUTMass11.df <- data.frame(slUTAll11.df$plot, slUTAll11.df$spAbbr, slUTAll11.
    df$position, slUTAll11.df$massFinal)
names(slUTMass11.df) <- c("plot", "spAbbr", "position", "massFinal")




##########################
## PREDICTION Accuracy ##
##########################

slCV <- 0.7437
RMSEsl <- data.frame(slCV)

RMSEsl$slBICUn  <- unscale(slCV, slTrain11.tmp$massUT)
RMSEsl$slBIC <- backLog.Tns(RMSEsl$slBICUn)

sl11TEST.df <- data.frame(predict(slUT11.PLSR, ncomp = 3, newdata = sl11.df))
```

```
names(sl11TEST.df) <- c("massUT")

unscale.object <- slTrain11.tmp$massUT

sl11TEST.df$massFinalT <- unscale(sl11TEST.df$massUT, slTrain11.tmp$massUT)

sl11TEST.df$massFinal <- backLog.Tns(sl11TEST.df$massFinalT)

sl11TEST.df$massREAL <- slTrain11.tmp$mass

sl11TEST.df$subtract <- sl11TEST.df$massFinal - sl11TEST.df$massREAL
mean(sl11TEST.df$subtract)
sd(sl11TEST.df$subtract)

sl11LM <- lm(sl11TEST.df$massFinal ~ sl11TEST.df$massREAL -1 ,  offset = 1.00
    * sl11TEST.df$massREAL)
summary (sl11LM)

mean(sl11TEST.df$massREAL)
sd(sl11TEST.df$massREAL)




# LZCA Training Data Import -

lzcaStemLenMean <- subset(stemLenMeanTrain11.df, spAbbr == 'LZCA')

lzcaStemLenSum <- subset(stemLenSumTrain11.df, spAbbr == 'LZCA')

lzcaMassMean <- subset(massMeanTrain11.df, spAbbr == 'LZCA')

lzcaTrain11.tmp <-
  merge(
  merge(lzcaStemLenSum, lzcaStemLenMean, by = c("plot", "position", "spAbbr","
    plsr")),
  lzcaMassMean, by = c("plot", "position", "spAbbr", "plsr"))

# LZCA Training Data Transformation -

lzcaTrain11.tmp$massT <- Log.Tns(lzcaTrain11.tmp$mass)
lzcaTrain11.tmp$stemLenSumT <- Sqrt.Tns(lzcaTrain11.tmp$stemLenSum)
lzcaTrain11.tmp$stemLenMeanT <- Raw.Tns(lzcaTrain11.tmp$stemLenMean)
```

```
lzcaTrain11.tmp$massUT <-  scale(lzcaTrain11.tmp$massT)
lzcaTrain11.tmp$stemLenSumUT <- scale(lzcaTrain11.tmp$stemLenSumT)
lzcaTrain11.tmp$stemLenMeanUT <- scale(lzcaTrain11.tmp$stemLenMeanT)

lzcaTrain11.sc <- lzcaTrain11.tmp [,11:13]

lzcaTrain11UTSq <- as.data.frame(
apply(lzcaTrain11.sc[,-1], 2,
function(x){ x^2 })
); names(lzcaTrain11UTSq) <- paste(
names(lzcaTrain11UTSq), "sq",
sep = "."
)


lzcaTrain11UTCb <- as.data.frame(
apply(lzcaTrain11.sc[,-1], 2,
function(x){ x^3 })
); names(lzcaTrain11UTCb) <- paste(
names(lzcaTrain11UTCb), "cb",
sep = "."
)

lzca11.df <- as.data.frame(cbind(lzcaTrain11.sc, lzcaTrain11UTSq,
    lzcaTrain11UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

lzca11LMF <- lm(massUT ~
stemLenSumUT + stemLenSumUT.sq + stemLenSumUT.cb +
stemLenMeanUT + stemLenMeanUT.sq + stemLenMeanUT.cb,
data = lzca11.df
)

lzca11Subset <- expression(
(stemLenSumUT | !stemLenSumUT.sq) & (stemLenSumUT & stemLenSumUT.sq | !
    stemLenSumUT.cb) &
(stemLenMeanUT | !stemLenMeanUT.sq) & (stemLenMeanUT & stemLenMeanUT.sq | !
    stemLenMeanUT.cb)
)
```

```
# AIC Dredge
# lzca11.AIC.dredge <-dredge(
# lzca11LMF,
# subset = lzca11Subset,
# rank = AIC
# )

# lzca11.AIC.top <- subset(
# lzca11.AIC.dredge,
# delta < 2
# )

# BIC Dredge
# lzca11.BIC.dredge <-dredge(
# lzca11LMF,
# subset = lzca11Subset,
# rank = BIC
# )

# lzca11.BIC.top <- subset(
# lzca11.BIC.dredge,
# delta < 2
# )

# LZCA Training PLSR -

lzcaUT11.PLSR <- plsr(lzca11.df$massUT ~ lzca11.df$stemLenSumUT + lzca11.df$
    stemLenMeanUT,
  ncomp = 2, data = lzca11.df, validation = "CV",
  method = "oscorespls")
summary(lzcaUT11.PLSR)
# LZCA Data Import --

lzcaStemLenMean <- subset(stemLenMeanData11.df, spAbbr == 'LZCA')

lzcaStemLenSum <- subset(stemLenSumData11.df, spAbbr == 'LZCA')

lzcaAll11.tmp <- merge(lzcaStemLenSum, lzcaStemLenMean, by = c("plot", "
    position",
 "spAbbr", "core"))

# LZCA Data Transformation --

lzcaUTAll11.df <- lzcaAll11.tmp

lzcaUTAll11.df$stemLenSumT <- Sqrt.Tns(lzcaUTAll11.df$stemLenSum)
```

```
lzcaUTAll11.df$stemLenMeanT <- Raw.Tns(lzcaUTAll11.df$stemLenMean)


lzcaUTAll11.df$stemLenSumUT <-PredU(lzcaUTAll11.df$stemLenSumT, lzcaTrain11.
    tmp$stemLenSumT)
lzcaUTAll11.df$stemLenMeanUT <-PredU(lzcaUTAll11.df$stemLenMeanT, lzcaTrain11.
    tmp$stemLenMeanT)

# LZCA Prediction -

lzcaUTAll11.df$mx <- as.matrix(lzcaUTAll11.df [c(9:10)])

lzcaUTAll11.df$predictMass <- predict(lzcaUT11.PLSR, ncomp = 2, newdata =
    lzcaUTAll11.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- lzcaTrain11.tmp$massUT
lzcaUTAll11.df$massFinalT <- unscale.fn(lzcaUTAll11.df$predictMass)

lzcaUTAll11.df$massFinal <-backLog.Tns(lzcaUTAll11.df$massFinalT)

lzcaUTMass11.df <- data.frame(lzcaUTAll11.df$plot, lzcaUTAll11.df$spAbbr,
    lzcaUTAll11.df$position, lzcaUTAll11.df$massFinal)
names(lzcaUTMass11.df) <- c("plot", "spAbbr", "position", "massFinal")


##########################
## PREDICTION Accuracy ##
##########################

lzcaCV <- 0.3948
RMSElzca <- data.frame(lzcaCV)

RMSElzca$lzcaBICUn  <- unscale(lzcaCV, lzcaTrain11.tmp$massUT)
RMSElzca$lzcaBIC <- backLog.Tns(RMSElzca$lzcaBICUn)

lzca11TEST.df <- data.frame(predict(lzcaUT11.PLSR, ncomp = 2, newdata = lzca11
    .df))
names(lzca11TEST.df) <- c("massUT")

unscale.object <- lzcaTrain11.tmp$massUT

lzca11TEST.df$massFinalT <- unscale(lzca11TEST.df$massUT, lzcaTrain11.tmp$
    massUT)
```

```
lzca11TEST.df$massFinal <- backLog.Tns(lzca11TEST.df$massFinalT)

lzca11TEST.df$massREAL <- lzcaTrain11.tmp$mass

lzca11TEST.df$subtract <- lzca11TEST.df$massFinal - lzca11TEST.df$massREAL
mean(lzca11TEST.df$subtract)
sd(lzca11TEST.df$subtract)

lzca11LM <- lm(lzca11TEST.df$massFinal ~ lzca11TEST.df$massREAL -1 ,  offset =
     1.00 * lzca11TEST.df$massREAL)
summary (lzca11LM)

mean(lzca11TEST.df$massREAL)
sd(lzca11TEST.df$massREAL)



# DC Training Data Import -

dcStemLenMean <- subset(stemLenMeanTrain11.df, spAbbr == 'DC')

dcStemLenSum <- subset(stemLenSumTrain11.df, spAbbr == 'DC')

dcMassMean <- subset(massMeanTrain11.df, spAbbr == 'DC')

dcBCirc <- subset(mainTrain11.df, spAbbr == 'DC'& dataType == 'bCirc',
 select = c(plot, position, spAbbr, plsr, data))

names(dcBCirc) <- c("plot","position", "spAbbr", "plsr", "bCirc")

dcTrain11.tmp <-
  merge(
  merge(
  merge(dcStemLenSum, dcStemLenMean, by = c("plot", "position", "spAbbr","plsr
    ")),
  dcBCirc, by = c("plot", "position", "spAbbr", "plsr")),
  dcMassMean, by = c("plot", "position", "spAbbr", "plsr"))

# DC Training Data Transformation -

dcTrain11.tmp$massT <- Log.Tns(dcTrain11.tmp$mass)
dcTrain11.tmp$stemLenSumT <- Log.Tns(dcTrain11.tmp$stemLenSum)
dcTrain11.tmp$stemLenMeanT <- Sqrt.Tns(dcTrain11.tmp$stemLenMean)
dcTrain11.tmp$bCircT <- Log.Tns(dcTrain11.tmp$bCirc)
```

```
dcTrain11.tmp$massUT <-  scale(dcTrain11.tmp$massT)
dcTrain11.tmp$stemLenSumUT <- scale(dcTrain11.tmp$stemLenSumT)
dcTrain11.tmp$stemLenMeanUT <- scale(dcTrain11.tmp$stemLenMeanT)
dcTrain11.tmp$bCircUT <- scale(dcTrain11.tmp$bCircT)

dcTrain11.sc <- dcTrain11.tmp [,13:16]

dcTrain11UTSq <- as.data.frame(
apply(dcTrain11.sc[,-1], 2,
function(x){ x^2 })
); names(dcTrain11UTSq) <- paste(
names(dcTrain11UTSq), "sq",
sep = "."
)


dcTrain11UTCb <- as.data.frame(
apply(dcTrain11.sc[,-1], 2,
function(x){ x^3 })
); names(dcTrain11UTCb) <- paste(
names(dcTrain11UTCb), "cb",
sep = "."
)

dc11.df <- as.data.frame(cbind(dcTrain11.sc, dcTrain11UTSq, dcTrain11UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

dc11LMF <- lm(massUT ~
stemLenSumUT + stemLenSumUT.sq + stemLenSumUT.cb +
stemLenMeanUT + stemLenMeanUT.sq + stemLenMeanUT.cb +
bCircUT + bCircUT.sq + bCircUT.cb,
data = dc11.df
)

dc11Subset <- expression(
(stemLenSumUT | !stemLenSumUT.sq) & (stemLenSumUT & stemLenSumUT.sq | !
    stemLenSumUT.cb) &
(stemLenMeanUT | !stemLenMeanUT.sq) & (stemLenMeanUT & stemLenMeanUT.sq | !
    stemLenMeanUT.cb) &
(bCircUT | !bCircUT.sq) & (bCircUT & bCircUT.sq | !bCircUT.cb)
)
```

```
# AIC Dredge
# dc11.AIC.dredge <-dredge(
# dc11LMF,
# subset = dc11Subset,
# rank = AIC
# )

# dc11.AIC.top <- subset(
# dc11.AIC.dredge,
# delta < 2
# )

# BIC Dredge
# dc11.BIC.dredge <-dredge(
# dc11LMF,
# subset = dc11Subset,
# rank = BIC
# )

# dc11.BIC.top <- subset(
# dc11.BIC.dredge,
# delta < 2
# )

# DC Training PLSR --


dcUT11.PLSR <- plsr(dc11.df$massUT ~ dc11.df$stemLenMeanUT + dc11.df$
    stemLenSumUT +
          dc11.df$stemLenSumUT.sq,
  ncomp = 3, data = dc11.df, validation = "CV",
  method = "oscorespls")

summary(dcUT11.PLSR)

# DC Data Import --

dcStemLenMean <- subset(stemLenMeanData11.df, spAbbr == 'DC')

dcStemLenSum <- subset(stemLenSumData11.df, spAbbr == 'DC')

dcBCirc <- subset(data11.df, spAbbr == 'DC'& dataType == 'bCirc',
 select = c(plot, position, spAbbr, core, data))

names(dcBCirc) <- c("plot","position", "spAbbr", "core", "bCirc")
```

```
dcAll11.tmp <-
  merge(
  merge(dcStemLenSum, dcStemLenMean, by = c("plot", "position", "spAbbr","core
    ")),
  dcBCirc, by = c("plot", "position", "spAbbr", "core"))

# DC Data Transformation --

dcUTAll11.df <- dcAll11.tmp

dcUTAll11.df$stemLenSumT <- Log.Tns(dcUTAll11.df$stemLenSum)
dcUTAll11.df$stemLenMeanT <- Sqrt.Tns(dcUTAll11.df$stemLenMean)
dcUTAll11.df$stemLenMeanUT <-PredU(dcUTAll11.df$stemLenMeanT, dcTrain11.tmp$
    stemLenMeanT)


dcUTAll11.df$stemLenSumT <- Log.Tns(dcUTAll11.df$stemLenSum)
dcUTAll11.df$stemLenSumUT <-PredU(dcUTAll11.df$stemLenSumT, dcTrain11.tmp$
    stemLenSumT)
dcUTAll11.df$stemLenSumUT.sq <- (dcUTAll11.df$stemLenSumUT)^2
dcUTAll11.df$stemLenSumUT.cb <- (dcUTAll11.df$stemLenSumUT)^3


# DC Prediction -

dcUTAll11.df$mx <- as.matrix(dcUTAll11.df [c(10:12)])

dcUTAll11.df$predictMass <- predict(dcUT11.PLSR, ncomp = 2, newdata =
    dcUTAll11.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- dcTrain11.tmp$massUT
dcUTAll11.df$massFinalT <- unscale.fn(dcUTAll11.df$predictMass)

dcUTAll11.df$massFinal <-backLog.Tns(dcUTAll11.df$massFinalT)

dcUTMass11.df <- data.frame(dcUTAll11.df$plot, dcUTAll11.df$spAbbr, dcUTAll11.
    df$position, dcUTAll11.df$massFinal)
names(dcUTMass11.df) <- c("plot", "spAbbr", "position", "massFinal")


##########################
## PREDICTION Accuracy ##
```

```
#########################

dcCV <- 0.4064
RMSEdc <- data.frame(dcCV)

RMSEdc$dcBICUn  <- unscale(dcCV, dcTrain11.tmp$massUT)
RMSEdc$dcBIC <- backLog.Tns(RMSEdc$dcBICUn)

dc11TEST.df <- data.frame(predict(dcUT11.PLSR, ncomp = 2, newdata = dc11.df))
names(dc11TEST.df) <- c("massUT")

unscale.object <- dcTrain11.tmp$massUT

dc11TEST.df$massFinalT <- unscale(dc11TEST.df$massUT, dcTrain11.tmp$massUT)

dc11TEST.df$massFinal <- backLog.Tns(dc11TEST.df$massFinalT)

dc11TEST.df$massREAL <- dcTrain11.tmp$mass

dc11TEST.df$subtract <- dc11TEST.df$massFinal - dc11TEST.df$massREAL
mean(dc11TEST.df$subtract)
sd(dc11TEST.df$subtract)

dc11LM <- lm(dc11TEST.df$massFinal ~ dc11TEST.df$massREAL -1 ,  offset = 1.00
    * dc11TEST.df$massREAL)
summary (dc11LM)

mean(dc11TEST.df$massREAL)
sd(dc11TEST.df$massREAL)




# Plot Level Biomass Prediction -

bindUT11.df <- rbind(agUTMass11.df, pvUTMass11.df, dcUTMass11.df, lzcaUTMass11
    .df,
  slUTMass11.df, lcUTMass11.df)

plotUT11.df <- aggregate(bindUT11.df$massFinal, by=list(bindUT11.df$plot), FUN
     = sum, na.rm = TRUE)
names(plotUT11.df) <- c("plot", "biomassPred")

plotUT11.final <- unique(merge(plotUT11.df, plot11Info.df[,-2:-4], by = "plot"
    ))
```

```
biomass11 <- aov(Log.Tns(plotUT11.final$biomassPred) ~ treatment * amf, data =
    plotUT11.final)

TukeyHSD(biomass11, ordered = TRUE)


# Species Dataframe --
agMass11.df <- aggregate(agUTMass11.df$massFinal, by=list(agUTMass11.df$plot,
    agUTMass11.df$spAbbr), FUN = sum, na.rm = TRUE)

pvMass11.df <- aggregate(pvUTMass11.df$massFinal, by=list(pvUTMass11.df$plot,
    pvUTMass11.df$spAbbr), FUN = sum, na.rm = TRUE)

lcMass11.df <- aggregate(lcUTMass11.df$massFinal, by=list(lcUTMass11.df$plot,
    lcUTMass11.df$spAbbr), FUN = sum, na.rm = TRUE)

slMass11.df <- aggregate(slUTMass11.df$massFinal, by=list(slUTMass11.df$plot,
    slUTMass11.df$spAbbr), FUN = sum, na.rm = TRUE)

lzcaMass11.df <- aggregate(lzcaUTMass11.df$massFinal, by=list(lzcaUTMass11.df$
    plot, lzcaUTMass11.df$spAbbr), FUN = sum, na.rm = TRUE)

dcMass11.df <- aggregate(dcUTMass11.df$massFinal, by=list(dcUTMass11.df$plot,
    dcUTMass11.df$spAbbr), FUN = sum, na.rm = TRUE)


speciesUT11.df <- rbind(agMass11.df, pvMass11.df, lcMass11.df, lzcaMass11.df,
    slMass11.df, dcMass11.df)
names(speciesUT11.df) <- c("plot", "spAbbr", "mass")

plot11Info.df1 <- plot11Info.df[,-2:-4]
species11.final <- unique(merge(speciesUT11.df, plot11Info.df1, by = c("plot")
    ))

qqnorm(Log.Tns(dcUTMass11.df$massFinal))
qqline(Log.Tns(dcUTMass11.df$massFinal))

#Plot Level Biomass Prediction Grass ONLY

grassUT11.df <- rbind(agUTMass11.df, pvUTMass11.df)

grassUT11.df <- aggregate(grassUT11.df$massFinal, by=list(grassUT11.df$plot),
    FUN = sum, na.rm = TRUE)

names(grassUT11.df) <- c("plot", "biomassPred")
```

```
grassUT11.final <- unique(merge(grassUT11.df, plot11Info.df [,-2:-4], by = "
    plot"))

grass11 <- aov(Log.Tns(grassUT11.final$biomassPred) ~ treatment * amf, data =
    grassUT11.final)

summary(grass11)
TukeyHSD(grass11, ordered = TRUE)


#Plot Level Biomass Prediction NF ONLY-

nfUT11.df <- rbind(dcUTMass11.df, lzcaUTMass11.df)

nfUT11.df <- aggregate(nfUT11.df$massFinal, by=list(nfUT11.df$plot), FUN = sum
    , na.rm = TRUE)

names(nfUT11.df) <- c("plot", "biomassPred")

nfUT11.final <- unique(merge(nfUT11.df, plot11Info.df[,-2:-4], by = "plot"))

nf11 <- aov(Log.Tns(nfUT11.final$biomassPred) ~ treatment * amf, data = nfUT11
    .final)

summary(nf11)
TukeyHSD(nf11, ordered = TRUE)

################################
## 2012 Plant Prediction Data ##
################################

# Plot Information Data Frame -

channel <-
odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/Data/
    dissertation_data_exp_1_13aug2")

plot12Info.df <- sqlQuery(channel, "SELECT plot, position, funGroup, spAbbr,
    amf, treatment, biochar_rate, compost_rate, season, hgt
FROM DATA_Q WHERE season = '2012'")

close(channel)
```

```
# Training Data (Variable Creation) -

channel <-
odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/Data/
    dissertation_data_exp_1_13aug2")


mainTrain12.df <- sqlQuery(channel, "SELECT plot, position, data, dataType,
    spAbbr, funGroup, plsr, test
FROM DATA_Q WHERE season = '2012' AND plsr = 1")


stemLenTrain12.df <- sqlQuery(channel, "SELECT plot, position, data, spAbbr,
    funGroup, plsr, test
FROM DATA_Q WHERE season = '2012' AND dataType = 'stemLen'
AND plsr = 1")

inflorLenTrain12.df <- sqlQuery(channel, "SELECT plot, position, data, spAbbr,
     funGroup, plsr, test
FROM DATA_Q WHERE season = '2012' AND dataType = 'inflorLen'
AND plsr = 1")

massTrain12.df <- sqlQuery(channel, "SELECT plot, position, data, spAbbr,
    funGroup, plsr, test
FROM DATA_Q WHERE season = '2012' AND dataType = 'mass'
AND plsr = 1")

close(channel)

## Calculates the Mean of stem length for the training plant

stemLenMeanTrain12.df <- aggregate(stemLenTrain12.df$data,
by=list(stemLenTrain12.df$plot, stemLenTrain12.df$position,
stemLenTrain12.df$spAbbr, stemLenTrain12.df$funGroup, stemLenTrain12.df$plsr,
stemLenTrain12.df$test),
FUN = mean, na.rm = TRUE)

names(stemLenMeanTrain12.df) <- c("plot","position", "spAbbr", "funGroup", "
    plsr", "test", "stemLenMean")


## Sum of stem lengths for prediction plants

stemLenSumTrain12.df <- aggregate(stemLenTrain12.df$data,
by=list(stemLenTrain12.df$plot, stemLenTrain12.df$position,
stemLenTrain12.df$spAbbr, stemLenTrain12.df$funGroup, stemLenTrain12.df$plsr,
```

```
stemLenTrain12.df$test),
FUN = sum, na.rm = TRUE)

names(stemLenSumTrain12.df) <- c("plot","position", "spAbbr", "funGroup", "
    plsr", "test", "stemLenSum")

## Sum of stem lengths for prediction plants

stemLenCountTrain12.df <- aggregate(stemLenTrain12.df$data,
by=list(stemLenTrain12.df$plot, stemLenTrain12.df$position,
stemLenTrain12.df$spAbbr, stemLenTrain12.df$funGroup, stemLenTrain12.df$plsr,
stemLenTrain12.df$test),
FUN = function(x) c(count = length(x)))

names(stemLenCountTrain12.df) <- c("plot","position", "spAbbr", "funGroup", "
    plsr", "test", "stemLenCount")


## Calculates the Mean of inflorescence length for the training plant

inflorLenMeanTrain12.df <- aggregate(inflorLenTrain12.df$data,
by=list(inflorLenTrain12.df$plot, inflorLenTrain12.df$position,
inflorLenTrain12.df$spAbbr, inflorLenTrain12.df$funGroup, inflorLenTrain12.df$
    plsr,
inflorLenTrain12.df$test),
FUN = mean, na.rm = TRUE)

names(inflorLenMeanTrain12.df) <- c("plot","position", "spAbbr", "funGroup", "
    plsr", "test", "inflorLenMean")


## Sum of inflorescence lengths for prediction plants

inflorLenSumTrain12.df <- aggregate(inflorLenTrain12.df$data,
by=list(inflorLenTrain12.df$plot, inflorLenTrain12.df$position,
inflorLenTrain12.df$spAbbr, inflorLenTrain12.df$funGroup, inflorLenTrain12.df$
    plsr,
inflorLenTrain12.df$test),
FUN = sum, na.rm = TRUE)

names(inflorLenSumTrain12.df) <- c("plot","position", "spAbbr", "funGroup", "
    plsr", "test", "inflorLenSum")


## Mean of plant biomass for prediction plants
```

```
massMeanTrain12.df <-aggregate(massTrain12.df$data,
by=list(massTrain12.df$plot, massTrain12.df$position,
massTrain12.df$spAbbr, massTrain12.df$funGroup, massTrain12.df$plsr,
massTrain12.df$test),
FUN = mean, na.rm = TRUE)

names(massMeanTrain12.df) <- c("plot","position", "spAbbr", "funGroup", "plsr"
    ,"test", "mass")

# Data (Variable Creation) -

channel <-
odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/Data/
    dissertation_data_exp_1_13aug2")


data12.df <- sqlQuery(channel, "SELECT plot, position, data, dataType, spAbbr,
     funGroup, core
FROM DATA_Q WHERE season = '2012' AND core = 1 AND harvestOne = 0")


stemLenData12.df <- sqlQuery(channel, "SELECT plot, position, data, spAbbr,
    funGroup, core
FROM DATA_Q WHERE season = '2012' AND dataType = 'stemLen'
AND core = 1 AND harvestOne = 0")

inflorLenData12.df <- sqlQuery(channel, "SELECT plot, position, data, spAbbr,
    funGroup, core
FROM DATA_Q WHERE season = '2012' AND dataType = 'inflorLen'
AND core = 1 AND harvestOne = 0")

close(channel)

## Stem Length Mean
stemLenMeanData12.df <-aggregate(stemLenData12.df$data,
by=list(stemLenData12.df$plot, stemLenData12.df$position,
stemLenData12.df$spAbbr, stemLenData12.df$funGroup, stemLenData12.df$core),
FUN = mean, na.rm = TRUE)

names(stemLenMeanData12.df) <- c("plot","position", "spAbbr", "funGroup", "
    core", "stemLenMean")

## Stem Length Sum
stemLenSumData12.df <-aggregate(stemLenData12.df$data,
by=list(stemLenData12.df$plot, stemLenData12.df$position,
stemLenData12.df$spAbbr, stemLenData12.df$funGroup, stemLenData12.df$core),
```

```
FUN = sum)

names(stemLenSumData12.df) <- c("plot","position", "spAbbr", "funGroup", "core
    ", "stemLenSum")


## Stem Length Count
stemLenCountData12.df <- aggregate(stemLenData12.df$data,
by=list(stemLenData12.df$plot, stemLenData12.df$position,
stemLenData12.df$spAbbr, stemLenData12.df$funGroup, stemLenData12.df$core),
FUN = function(x) c(count = length(x)))

names(stemLenCountData12.df) <- c("plot","position", "spAbbr", "funGroup", "
    core", "stemLenCount")

## Inflorescence Length Mean
inflorLenMeanData12.df <-aggregate(inflorLenData12.df$data,
by=list(inflorLenData12.df$plot, inflorLenData12.df$position,
inflorLenData12.df$spAbbr, inflorLenData12.df$funGroup, inflorLenData12.df$
    core),
FUN = mean, na.rm = TRUE)

names(inflorLenMeanData12.df) <- c("plot","position", "spAbbr", "funGroup", "
    core", "inflorLenMean")

## Inflorescence Length Sum
inflorLenSumData12.df <-aggregate(inflorLenData12.df$data,
by=list(inflorLenData12.df$plot, inflorLenData12.df$position,
inflorLenData12.df$spAbbr, inflorLenData12.df$funGroup, inflorLenData12.df$
    core),
FUN = sum, na.rm = TRUE)

names(inflorLenSumData12.df) <- c("plot","position", "spAbbr", "funGroup", "
    core", "inflorLenSum")


# AG Training Data Import --

agCirc30 <- subset(mainTrain12.df, spAbbr == 'AG'& dataType == 'circ30',
select = c(plot, position, spAbbr, funGroup, plsr, test, data))

names(agCirc30) <- c("plot","position", "spAbbr", "funGroup", "plsr", "test",
    "circ30")


agLWPHgt <- subset(mainTrain12.df, spAbbr == 'AG'& dataType == 'lWPHgt',
```

```
select = c(plot, position, spAbbr, funGroup, plsr, test, data))

names(agLWPHgt) <- c("plot","position", "spAbbr", "funGroup", "plsr", "test",
    "lWPHgt")

agSeedCulms <- subset(mainTrain12.df, spAbbr == 'AG'& dataType == 'noSeedCulms
    ',
select = c(plot, position, spAbbr, funGroup, plsr, test,  data))

names(agSeedCulms) <- c("plot","position", "spAbbr", "funGroup", "plsr", "test
    ", "seedCulms")

agMassMean <- subset(massMeanTrain12.df, spAbbr == 'AG')


agTrain12.tmp <-
merge(
merge(
merge(agCirc30, agLWPHgt, by = c("plot", "position", "spAbbr", "funGroup", "
    plsr", "test")),
agSeedCulms, by = c("plot", "position", "spAbbr", "funGroup", "plsr", "test"))
    ,
agMassMean, by = c("plot", "position", "spAbbr", "funGroup", "plsr", "test"))
# AG Training Transformation Section --

agTrain12.tmp$massT <- Log.Tns(agTrain12.tmp$mass)
agTrain12.tmp$lWPHgtT <- Log.Tns(agTrain12.tmp$lWPHgt)
agTrain12.tmp$circ30T <- Raw.Tns(agTrain12.tmp$circ30)
agTrain12.tmp$seedCulmsT <- Sqrt.Tns(agTrain12.tmp$seedCulms)


agTrain12.tmp$massUT <- scale(agTrain12.tmp$massT)
agTrain12.tmp$lWPHgtUT <- scale(agTrain12.tmp$lWPHgtT)
agTrain12.tmp$circ30UT <- scale(agTrain12.tmp$circ30T)
agTrain12.tmp$seedCulmsUT <- scale(agTrain12.tmp$seedCulmsT)


agTrain12.tmp <- subset(agTrain12.tmp, test == '0')

agTrain12.sc <- agTrain12.tmp [,15:18]

agTrain12UTSq <- as.data.frame(
apply(agTrain12.sc[,-1], 2,
function(x){ x^2 })
); names(agTrain12UTSq) <- paste(
names(agTrain12UTSq), "sq",
```

```
sep = "."
)


agTrain12UTCb <- as.data.frame(
apply(agTrain12.sc[,-1], 2,
function(x){ x^3 })
); names(agTrain12UTCb) <- paste(
names(agTrain12UTCb), "cb",
sep = "."
)

ag12.df <- as.data.frame(cbind(agTrain12.sc, agTrain12UTSq, agTrain12UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

ag12LMF <- lm(massUT ~
lWPHgtUT  + lWPHgtUT.sq + lWPHgtUT.cb +
circ30UT + circ30UT.sq + circ30UT.cb +
seedCulmsUT + seedCulmsUT.sq + seedCulmsUT.cb,
data = ag12.df
)

ag12Subset <- expression(
(lWPHgtUT | !lWPHgtUT.sq) & (lWPHgtUT & lWPHgtUT.sq | !lWPHgtUT.cb) &
(circ30UT | !circ30UT.sq) & (circ30UT & circ30UT.sq | !circ30UT.cb) &
(seedCulmsUT | !seedCulmsUT.sq) & (seedCulmsUT & seedCulmsUT.sq | !seedCulmsUT
    .cb)
)

#AIC Dredge
# ag12.AIC.dredge <-dredge(
# ag12LMF,
# subset = ag12Subset,
# rank = AIC
# )

# ag12.AIC.top <- subset(
# ag12.AIC.dredge,
# delta < 2
# )
```

```
#BIC Dredge
# ag12.BIC.dredge <-dredge(
# ag12LMF,
# subset = ag12Subset,
# rank = BIC
# )

# ag12.BIC.top <- subset(
# ag12.BIC.dredge,
# delta < 2
# )

# confset.95p <- get.models(ag12.BIC.dredge, cumsum(weight) <= .95)
# avgmod.95p <- model.avg(confset.95p)
# summary(avgmod.95p)


# AG Training PLSR Section -

agUT12.PLSR <- plsr(ag12.df$massUT ~ ag12.df$lWPHgtUT + ag12.df$circ30UT +
    ag12.df$circ30UT.sq +
ag12.df$seedCulmsUT + ag12.df$seedCulmsUT.sq, ncomp = 5, data = ag12.df,
    validation = "CV",
method = "oscorespls")

summary(agUT12.PLSR)



# AG Data Import --

agCirc30 <- subset(data12.df, spAbbr == 'AG'& dataType == 'circ30',
select = c(plot, position, spAbbr, funGroup, core, data))

names(agCirc30) <- c("plot","position", "spAbbr", "funGroup", "core", "circ30"
    )


agLWPHgt <- subset(data12.df, spAbbr == 'AG'& dataType == 'lWPHgt',
select = c(plot, position, spAbbr, funGroup, core, data))

names(agLWPHgt) <- c("plot","position", "spAbbr", "funGroup", "core", "lWPHgt"
    )

agSeedCulms <- subset(data12.df, spAbbr == 'AG'& dataType == 'noSeedCulms',
select = c(plot, position, spAbbr, funGroup, core, data))
```

```
names(agSeedCulms) <- c("plot","position", "spAbbr", "funGroup", "core", "
    seedCulms")


agAll12.tmp <-
merge(
merge(agCirc30, agLWPHgt, by = c("plot", "position", "spAbbr", "funGroup", "
    core")),
agSeedCulms, by = c("plot", "position", "spAbbr", "funGroup", "core"))

# AG Data Transformation -

agUTAll12.df <- agAll12.tmp

agUTAll12.df$lWPHgtT <- Log.Tns(agUTAll12.df$lWPHgt)
agUTAll12.df$circ30T <- Raw.Tns(agUTAll12.df$circ30)
agUTAll12.df$seedCulmsT <- Sqrt.Tns(agUTAll12.df$seedCulms)

agUTAll12.df$lWPHgtUT <-PredU(agUTAll12.df$lWPHgtT, agTrain12.tmp$lWPHgtT)
agUTAll12.df$circ30UT <-PredU(agUTAll12.df$circ30T, agTrain12.tmp$circ30T)
agUTAll12.df$circ30UT.sq <- (agUTAll12.df$circ30UT)^2

agUTAll12.df$seedCulmsUT <-PredU(agUTAll12.df$seedCulmsT, agTrain12.tmp$
    seedCulmsT)
agUTAll12.df$seedCulmsUT.sq <- (agUTAll12.df$seedCulmsUT)^2


# AG Data Prediction -

agUTAll12.df$mx <- as.matrix(agUTAll12.df [c(12:16)])

agUTAll12.df$predictMass <- predict(agUT12.PLSR, ncomp = 3, newdata =
    agUTAll12.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- agTrain12.tmp$massUT
agUTAll12.df$massFinalT <- unscale.fn(agUTAll12.df$predictMass)

agUTAll12.df$massFinal <-backLog.Tns(agUTAll12.df$massFinalT)

agUTMass12.df <- data.frame(agUTAll12.df$plot, agUTAll12.df$position,
    agUTAll12.df$spAbbr,
agUTAll12.df$funGroup, agUTAll12.df$massFinal)
```

```
names(agUTMass12.df) <- c("plot", "position", "spAbbr", "funGroup", "massFinal
    ")

ag12Test.df <- merge(agTrain12.tmp, agUTMass12.df, by = c("plot", "position",
    "spAbbr", "funGroup"))
ag12Test.df$subtract <- ag12Test.df$mass - ag12Test.df$massFinal

##########################
## PREDICTION Accuracy ##
##########################

agCV <- 0.2664
RMSEag <- data.frame(agCV)

RMSEag$agBICUn  <- unscale(agCV, agTrain12.tmp$massUT)
RMSEag$agBIC <- backLog.Tns(RMSEag$agBICUn)

ag12TEST.df <- data.frame(predict(agUT12.PLSR, ncomp = 3, newdata = ag12.df))
names(ag12TEST.df) <- c("massUT")

unscale.object <- agTrain12.tmp$massUT

ag12TEST.df$massFinalT <- unscale(ag12TEST.df$massUT, agTrain12.tmp$massUT)

ag12TEST.df$massFinal <-backLog.Tns(ag12TEST.df$massFinalT)

ag12TEST.df$massREAL <- agTrain12.tmp$mass

ag12TEST.df$subtract <- ag12TEST.df$massFinal - ag12TEST.df$massREAL
mean(ag12TEST.df$subtract)
sd(ag12TEST.df$subtract)

ag12LM <- lm(ag12TEST.df$massFinal ~ ag12TEST.df$massREAL -1 ,  offset = 1.00
    * ag12TEST.df$massREAL)
summary (ag12LM)


# PV Training Data Import --

pvCirc30 <- subset(mainTrain12.df, spAbbr == 'PV'& dataType == 'circ30',
select = c(plot, position, spAbbr, funGroup, plsr, data))

names(pvCirc30) <- c("plot","position", "spAbbr", "funGroup", "plsr", "circ30"
    )
```

```
pvLWPHgt <- subset(mainTrain12.df, spAbbr == 'PV'& dataType == 'lWPHgt',
select = c(plot, position, spAbbr, funGroup, plsr, data))

names(pvLWPHgt) <- c("plot","position", "spAbbr", "funGroup", "plsr", "lWPHgt"
    )

pvSeedCulms <- subset(mainTrain12.df, spAbbr == 'PV'& dataType == 'noSeedCulms
    ',
select = c(plot, position, spAbbr, funGroup, plsr, data))

names(pvSeedCulms) <- c("plot","position", "spAbbr", "funGroup", "plsr", "
    seedCulms")


pvMassMean <- subset(massMeanTrain12.df, spAbbr == 'PV')


pvTrain12.tmp <-
merge(
merge(
merge(pvCirc30, pvLWPHgt, by = c("plot", "position", "spAbbr", "funGroup", "
    plsr")),
pvSeedCulms, by = c("plot", "position", "spAbbr", "funGroup", "plsr")),
pvMassMean, by = c("plot", "position", "spAbbr", "funGroup", "plsr"))

# PV Training Transformation --

pvTrain12.tmp$massT <- Sqrt.Tns(pvTrain12.tmp$mass)
pvTrain12.tmp$lWPHgtT <- Sqrt.Tns(pvTrain12.tmp$lWPHgt)
pvTrain12.tmp$circ30T <- Log.Tns(pvTrain12.tmp$circ30)
pvTrain12.tmp$seedCulmsT <- Sqrt.Tns(pvTrain12.tmp$seedCulms)

pvTrain12.tmp$massUT <- scale(pvTrain12.tmp$massT)
pvTrain12.tmp$lWPHgtUT <- scale(pvTrain12.tmp$lWPHgtT)
pvTrain12.tmp$circ30UT <- scale(pvTrain12.tmp$circ30T)
pvTrain12.tmp$seedCulmsUT <- scale(pvTrain12.tmp$seedCulmsT)

pvTrain12.sc <- pvTrain12.tmp [,15:18]

pvTrain12UTSq <- as.data.frame(
apply(pvTrain12.sc[,-1], 2,
function(x){ x^2 })
); names(pvTrain12UTSq) <- paste(
names(pvTrain12UTSq), "sq",
sep = "."
)
```

```
pvTrain12UTCb <- as.data.frame(
apply(pvTrain12.sc[,-1], 2,
function(x){ x^3 })
); names(pvTrain12UTCb) <- paste(
names(pvTrain12UTCb), "cb",
sep = "."
)

pv12.df <- as.data.frame(cbind(pvTrain12.sc, pvTrain12UTSq, pvTrain12UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

pv12LMF <- lm(massUT ~
lWPHgtUT  + lWPHgtUT.sq + lWPHgtUT.cb +
circ30UT + circ30UT.sq + circ30UT.cb +
seedCulmsUT + seedCulmsUT.sq + seedCulmsUT.cb,
data = pv12.df
)

pv12Subset <- expression(
(lWPHgtUT | !lWPHgtUT.sq) & (lWPHgtUT & lWPHgtUT.sq | !lWPHgtUT.cb) &
(circ30UT | !circ30UT.sq) & (circ30UT & circ30UT.sq | !circ30UT.cb) &
(seedCulmsUT | !seedCulmsUT.sq) & (seedCulmsUT & seedCulmsUT.sq | !seedCulmsUT
    .cb)
)

#AIC Dredge
# pv12.AIC.dredge <-dredge(
# pv12LMF,
# subset = pv12Subset,
# rank = AIC
# )

# pv12.AIC.top <- subset(
# pv12.AIC.dredge,
# delta < 2
# )

#BIC Dredge
# pv12.BIC.dredge <-dredge(
```

```
# pv12LMF,
# subset = pv12Subset,
# rank = BIC
# )

# pv12.BIC.top <- subset(
# pv12.BIC.dredge,
# delta < 2
# )

# PV Training PLSR Section --

pvUT12.PLSR <- plsr(pv12.df$massUT ~ pv12.df$circ30UT + pv12.df$lWPHgtUT +
+ pv12.df$lWPHgtUT.sq + pv12.df$seedCulmsUT
, ncomp = 4, data = pv12.df, validation = "CV",
method = "oscorespls")

summary(pvUT12.PLSR)
# PV Data Import --

pvCirc30 <- subset(data12.df, spAbbr == 'PV'& dataType == 'circ30',
select = c(plot, position, spAbbr, funGroup, core, data))

names(pvCirc30) <- c("plot","position", "spAbbr", "funGroup", "core", "circ30"
    )


pvLWPHgt <- subset(data12.df, spAbbr == 'PV'& dataType == 'lWPHgt',
select = c(plot, position, spAbbr, funGroup, core, data))

names(pvLWPHgt) <- c("plot","position", "spAbbr", "funGroup", "core", "lWPHgt"
    )

pvSeedCulms <- subset(data12.df, spAbbr == 'PV'& dataType == 'noSeedCulms',
select = c(plot, position, spAbbr, funGroup, core, data))

names(pvSeedCulms) <- c("plot","position", "spAbbr", "funGroup", "core", "
    seedCulms")


pvAll12.tmp <-
merge(
merge(pvCirc30, pvLWPHgt, by = c("plot", "position", "spAbbr", "funGroup", "
    core")),
pvSeedCulms, by = c("plot", "position", "spAbbr", "funGroup", "core"))
```

```
# PV Data Transformation -

pvUTAll12.df <- pvAll12.tmp

pvUTAll12.df$lWPHgtT <- Sqrt.Tns(pvUTAll12.df$lWPHgt)
pvUTAll12.df$circ30T <- Log.Tns(pvUTAll12.df$circ30)
pvUTAll12.df$seedCulmsT <- Sqrt.Tns(pvUTAll12.df$seedCulms)

pvUTAll12.df$circ30UT <-PredU(pvUTAll12.df$circ30T, pvTrain12.tmp$circ30T)
pvUTAll12.df$lWPHgtUT <-PredU(pvUTAll12.df$lWPHgtT, pvTrain12.tmp$lWPHgtT)
pvUTAll12.df$lWPHgtUT.sq <- (pvUTAll12.df$lWPHgtUT)^2
pvUTAll12.df$seedCulmsUT <-PredU(pvUTAll12.df$seedCulmsT, pvTrain12.tmp$
    seedCulmsT)

# PV Data Prediction -

pvUTAll12.df$mx <- as.matrix(pvUTAll12.df [c(12:15)])

pvUTAll12.df$predictMass <- predict(pvUT12.PLSR, ncomp = 2, newdata =
    pvUTAll12.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- pvTrain12.tmp$massUT
pvUTAll12.df$massFinalT <- unscale.fn(pvUTAll12.df$predictMass)

pvUTAll12.df$massFinal <- backSqrt.Tns(pvUTAll12.df$massFinalT)

pvUTMass12.df <- data.frame(pvUTAll12.df$plot, pvUTAll12.df$position,
    pvUTAll12.df$spAbbr,
pvUTAll12.df$funGroup, pvUTAll12.df$massFinal)
names(pvUTMass12.df) <- c("plot", "position", "spAbbr", "funGroup", "massFinal
    ")

pv12Test.df <- merge(pvTrain12.tmp, pvUTMass12.df, by = c("plot", "position",
    "spAbbr", "funGroup"))
pv12Test.df$subtract <- pv12Test.df$mass - pv12Test.df$massFinal


##########################
## PREDICTION Accuracy ##
##########################

pvCV <- 0.2923
RMSEpv <- data.frame(pvCV)
```

```
RMSEpv$pvBICUn  <- unscale(pvCV, pvTrain12.tmp$massUT)
RMSEpv$pvBIC <- backSqrt.Tns(RMSEpv$pvBICUn)

pv12TEST.df <- data.frame(predict(pvUT12.PLSR, ncomp = 2, newdata = pv12.df))
names(pv12TEST.df) <- c("massUT")

unscale.object <- pvTrain12.tmp$massUT

pv12TEST.df$massFinalT <- unscale(pv12TEST.df$massUT, pvTrain12.tmp$massUT)

pv12TEST.df$massFinal <- backSqrt.Tns(pv12TEST.df$massFinalT)

pv12TEST.df$massREAL <- pvTrain12.tmp$mass

pv12TEST.df$subtract <- pv12TEST.df$massFinal - pv12TEST.df$massREAL
mean(pv12TEST.df$subtract)
sd(pv12TEST.df$subtract)

pv12LM <- lm(pv12TEST.df$massFinal ~ pv12TEST.df$massREAL -1 ,  offset = 1.00
    * pv12TEST.df$massREAL)
summary (pv12LM)

mean(pv12TEST.df$massREAL)
sd(pv12TEST.df$massREAL)



# LC Training Data Import --

lcLeafNo <- subset(mainTrain12.df, spAbbr == 'LC'& dataType == 'leafNo',
select = c(plot, position, spAbbr, funGroup, plsr, data))

names(lcLeafNo) <- c("plot","position", "spAbbr", "funGroup", "plsr", "leafNo"
    )

lcS4LvHgt <- subset(mainTrain12.df, spAbbr == 'LC'& dataType == 's4LvHgt',
select = c(plot, position, spAbbr, funGroup, plsr, data))

names(lcS4LvHgt) <- c("plot","position", "spAbbr", "funGroup", "plsr", "
    s4LvHgt")

lcMassMean <- subset(massMeanTrain12.df, spAbbr == 'LC')

lcTrain12.tmp <-  merge(
```

```
merge(lcMassMean, lcS4LvHgt, by = c("plot", "position", "spAbbr", "funGroup",
    "plsr")),
lcLeafNo, by = c("plot", "position", "spAbbr", "funGroup", "plsr"))

# LC Training Transformation --

lcTrain12.tmp$massT <- Log.Tns(lcTrain12.tmp$mass)
lcTrain12.tmp$leafNoT <- Log.Tns(lcTrain12.tmp$leafNo)
lcTrain12.tmp$s4LvHgtT <- Raw.Tns(lcTrain12.tmp$s4LvHgt)

lcTrain12.tmp$massUT <- scale(lcTrain12.tmp$massT)
lcTrain12.tmp$leafNoUT <- scale(lcTrain12.tmp$leafNoT)
lcTrain12.tmp$s4LvHgtUT <- scale(lcTrain12.tmp$s4LvHgtT)

lcTrain12.sc <- lcTrain12.tmp[,13:15]

lcTrain12UTSq <- as.data.frame(
apply(lcTrain12.sc[,-1], 2,
function(x){ x^2 })
); names(lcTrain12UTSq) <- paste(
names(lcTrain12UTSq), "sq",
sep = "."
)


lcTrain12UTCb <- as.data.frame(
apply(lcTrain12.sc[,-1], 2,
function(x){ x^3 })
); names(lcTrain12UTCb) <- paste(
names(lcTrain12UTCb), "cb",
sep = "."
)

lc12.df <- as.data.frame(cbind(lcTrain12.sc, lcTrain12UTSq, lcTrain12UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

lc12LMF <- lm(massUT ~
leafNoUT  + leafNoUT.sq + leafNoUT.cb +
s4LvHgtUT + s4LvHgtUT.sq + s4LvHgtUT.cb,
data = lc12.df
)
```

```
lc12Subset <- expression(
(leafNoUT | !leafNoUT.sq) & (leafNoUT & leafNoUT.sq | !leafNoUT.cb) &
(s4LvHgtUT | !s4LvHgtUT.sq) & (s4LvHgtUT & s4LvHgtUT.sq | !s4LvHgtUT.cb)
)

#AIC Dredge
# lc12.AIC.dredge <-dredge(
# lc12LMF,
# subset = lc12Subset,
# rank = AIC
# )

# lc12.AIC.top <- subset(
# lc12.AIC.dredge,
# delta < 2
# )

#BIC Dredge
# lc12.BIC.dredge <-dredge(
# lc12LMF,
# subset = lc12Subset,
# rank = BIC
# )

# lc12.BIC.top <- subset(
# lc12.BIC.dredge,
# delta < 2
# )
# LC Training PLSR Section -

lcUT12.PLSR <- plsr(lc12.df$massUT ~ lc12.df$leafNoUT + lc12.df$s4LvHgtUT,
ncomp = 2, data = lc12.df, validation = "CV",
method = "oscorespls")
summary(lcUT12.PLSR)
# LC Data Import --

lcLeafNo <- subset(data12.df, spAbbr == 'LC'& dataType == 'leafNo',
select = c(plot, position, spAbbr, funGroup, core, data))

names(lcLeafNo) <- c("plot","position", "spAbbr", "funGroup", "core", "leafNo"
    )

lcS4LvHgt <- subset(data12.df, spAbbr == 'LC'& dataType == 's4LvHgt',
select = c(plot, position, spAbbr, funGroup, core, data))
```

```
names(lcS4LvHgt) <- c("plot","position", "spAbbr", "funGroup", "core", "
    s4LvHgt")


lcAll12.tmp <- merge(lcLeafNo, lcS4LvHgt, by = c("plot", "position", "spAbbr",
     "funGroup", "core"))

# LC Data Transformation --

lcUTAll12.df <- lcAll12.tmp

lcUTAll12.df$leafNoT <- Log.Tns(lcUTAll12.df$leafNo)
lcUTAll12.df$s4LvHgtT <- Raw.Tns(lcUTAll12.df$s4LvHgt)

lcUTAll12.df$leafNoUT <- PredU(lcUTAll12.df$leafNoT, lcTrain12.tmp$leafNoT)
lcUTAll12.df$s4LvHgtUT <- PredU(lcUTAll12.df$s4LvHgtT, lcTrain12.tmp$s4LvHgtT)

# LC Data Prediction -

lcUTAll12.df$mx <- as.matrix(lcUTAll12.df [c(10:11)])

lcUTAll12.df$predictMass <- predict(lcUT12.PLSR, ncomp = 1, newdata =
    lcUTAll12.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- lcTrain12.tmp$massUT
lcUTAll12.df$massFinalT <- unscale.fn(lcUTAll12.df$predictMass)

lcUTAll12.df$massFinal <- backLog.Tns(lcUTAll12.df$massFinalT)

lcUTMass12.df <- data.frame(lcUTAll12.df$plot, lcUTAll12.df$position,
    lcUTAll12.df$spAbbr,
lcUTAll12.df$funGroup, lcUTAll12.df$massFinal)
names(lcUTMass12.df) <- c("plot", "position", "spAbbr", "funGroup", "massFinal
    ")

lc12Test.df <- merge(lcTrain12.tmp, lcUTMass12.df, by = c("plot", "position"))
lc12Test.df$subtract <- lc12Test.df$mass - lc12Test.df$massFinal
lc12Test.df



##########################
## PREDICTION Accuracy ##
```

```
#########################

lcCV <- 0.596
RMSElc <- data.frame(lcCV)

RMSElc$lcBICUn  <- unscale(lcCV, lcTrain12.tmp$massUT)
RMSElc$lcBIC <- backLog.Tns(RMSElc$lcBICUn)

lc12TEST.df <- data.frame(predict(lcUT12.PLSR, ncomp = 1, newdata = lc12.df))
names(lc12TEST.df) <- c("massUT")

unscale.object <- lcTrain12.tmp$massUT

lc12TEST.df$massFinalT <- unscale(lc12TEST.df$massUT, lcTrain12.tmp$massUT)

lc12TEST.df$massFinal <- backLog.Tns(lc12TEST.df$massFinalT)

lc12TEST.df$massREAL <- lcTrain12.tmp$mass

lc12TEST.df$subtract <- lc12TEST.df$massFinal - lc12TEST.df$massREAL
mean(lc12TEST.df$subtract)
sd(lc12TEST.df$subtract)

lc12LM <- lm(lc12TEST.df$massFinal ~ lc12TEST.df$massREAL -1 ,  offset = 1.00
    * lc12TEST.df$massREAL)
summary (lc12LM)

mean(lc12TEST.df$massREAL)
sd(lc12TEST.df$massREAL)




# SL Training Data Import --

slStemLenMean <- subset(stemLenMeanTrain12.df, spAbbr == 'SL')

slStemLenSum <- subset(stemLenSumTrain12.df, spAbbr == 'SL')

slInflorLenMean <- subset(inflorLenMeanTrain12.df, spAbbr == 'SL')

slInflorLenSum <- subset(inflorLenSumTrain12.df, spAbbr == 'SL')


slMassMean <- subset(massMeanTrain12.df, spAbbr == 'SL')
```

```
slTrain12.tmp <-
merge(
merge(
merge(
merge(slMassMean, slStemLenMean, by = c("plot", "position", "spAbbr", "
    funGroup", "plsr", "test")),
slStemLenSum, by = c("plot", "position", "spAbbr", "funGroup", "plsr", "test")
    ),
slInflorLenMean, by = c("plot", "position", "spAbbr", "funGroup", "plsr", "
    test")),
slInflorLenSum, by = c("plot", "position", "spAbbr", "funGroup", "plsr", "test
    "))

# SL Training Transformation --

slTrain12.tmp$massT <- Log.Tns(slTrain12.tmp$mass)
slTrain12.tmp$stemLenMeanT <- Sqrt.Tns(slTrain12.tmp$stemLenMean)
slTrain12.tmp$stemLenSumT <- Log.Tns(slTrain12.tmp$stemLenSum)
slTrain12.tmp$inflorLenMeanT <- Sqrt.Tns(slTrain12.tmp$inflorLenMean)
slTrain12.tmp$inflorLenSumT <- Sqrt.Tns(slTrain12.tmp$inflorLenSum)

slTrain12.tmp$massUT <- scale(slTrain12.tmp$massT)
slTrain12.tmp$stemLenMeanUT <- scale(slTrain12.tmp$stemLenMeanT)
slTrain12.tmp$stemLenSumUT <- scale(slTrain12.tmp$stemLenSumT)
slTrain12.tmp$inflorLenMeanUT <- scale(slTrain12.tmp$inflorLenMeanT)
slTrain12.tmp$inflorLenSumUT <- scale(slTrain12.tmp$inflorLenSumT)

slTrain12.sc <- slTrain12.tmp [,17:21]

slTrain12UTSq <- as.data.frame(
apply(slTrain12.sc[,-1], 2,
function(x){ x^2 })
); names(slTrain12UTSq) <- paste(
names(slTrain12UTSq), "sq",
sep = "."
)


slTrain12UTCb <- as.data.frame(
apply(slTrain12.sc[,-1], 2,
function(x){ x^3 })
); names(slTrain12UTCb) <- paste(
names(slTrain12UTCb), "cb",
sep = "."
)
```

```
sl12.df <- as.data.frame(cbind(slTrain12.sc, slTrain12UTSq, slTrain12UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

sl12LMF <- lm(massUT ~
stemLenMeanUT  + stemLenMeanUT.sq + stemLenMeanUT.cb +
stemLenSumUT + stemLenSumUT.sq + stemLenSumUT.cb +
inflorLenMeanUT  + inflorLenMeanUT.sq + inflorLenMeanUT.cb +
inflorLenSumUT  + inflorLenSumUT.sq + inflorLenSumUT.cb,
data = sl12.df
)

sl12Subset <- expression(
(stemLenMeanUT | !stemLenMeanUT.sq) & (stemLenMeanUT & stemLenMeanUT.sq | !
    stemLenMeanUT.cb) &
(stemLenSumUT | !stemLenSumUT.sq) & (stemLenSumUT & stemLenSumUT.sq | !
    stemLenSumUT.cb) &
(inflorLenMeanUT | !inflorLenMeanUT.sq) & (inflorLenMeanUT & inflorLenMeanUT.
    sq | !inflorLenMeanUT.cb) &
(inflorLenSumUT | !inflorLenSumUT.sq) & (inflorLenSumUT & inflorLenSumUT.sq |
    !inflorLenSumUT.cb)
)

#AIC Dredge
# sl12.AIC.dredge <-dredge(
# sl12LMF,
# subset = sl12Subset,
# rank = AIC
# )

# sl12.AIC.top <- subset(
# sl12.AIC.dredge,
# delta < 2
# )

#BIC Dredge
# sl12.BIC.dredge <-dredge(
# sl12LMF,
# subset = sl12Subset,
# rank = BIC
# )
```

```
# sl12.BIC.top <- subset(
# sl12.BIC.dredge,
# delta < 2
# )

# SL Training PLSR Section -

slUT12.PLSR <- plsr(sl12.df$massUT ~ sl12.df$inflorLenMeanUT + sl12.df$
    inflorLenMeanUT.sq +
sl12.df$stemLenSumUT + sl12.df$stemLenSumUT.sq,
ncomp = 4, data = sl12.df, validation = "CV",
method = "oscorespls")

summary(slUT12.PLSR)

# SL Data Import --

slStemLenMean <- subset(stemLenMeanData12.df, spAbbr == 'SL')

slStemLenSum <- subset(stemLenSumData12.df, spAbbr == 'SL')

slInflorLenMean <- subset(inflorLenMeanData12.df, spAbbr == 'SL')

slInflorLenSum <- subset(inflorLenSumData12.df, spAbbr == 'SL')

slAll12.tmp <-
merge(
merge(
merge(slInflorLenSum, slStemLenMean, by = c("plot", "position", "spAbbr", "
    funGroup","core")),
slStemLenSum, by = c("plot", "position", "spAbbr", "funGroup", "core")),
slInflorLenMean, by = c("plot", "position", "spAbbr", "funGroup", "core"))

# SL Data Transformation --

slUTAll12.df <- slAll12.tmp

slUTAll12.df$stemLenMeanT <- Sqrt.Tns(slUTAll12.df$stemLenMean)
slUTAll12.df$stemLenSumT <- Log.Tns(slUTAll12.df$stemLenSum)
slUTAll12.df$inflorLenMeanT <- Sqrt.Tns(slUTAll12.df$inflorLenMean)
slUTAll12.df$inflorLenSumT <- Sqrt.Tns(slUTAll12.df$inflorLenSum)

slUTAll12.df$stemLenMeanUT <- PredU(slUTAll12.df$stemLenMeanT, slTrain12.tmp$
    stemLenMeanT)
```

```
slUTAll12.df$inflorLenMeanUT <- PredU(slUTAll12.df$inflorLenMeanT, slTrain12.
    tmp$inflorLenMeanT)

slUTAll12.df$inflorLenSumUT <- PredU(slUTAll12.df$inflorLenSumT, slTrain12.tmp
    $inflorLenSumT)
slUTAll12.df$inflorLenSumUT.sq <- (slUTAll12.df$inflorLenSumUT)^2

slUTAll12.df$stemLenSumUT <- PredU(slUTAll12.df$stemLenSumT, slTrain12.tmp$
    stemLenSumT)
slUTAll12.df$stemLenSumUT.sq <- (slUTAll12.df$stemLenSumUT)^2


# SL Data Prediction -

slUTAll12.df$mx <- as.matrix(slUTAll12.df [c(16:19)])

slUTAll12.df$predictMass <- predict(slUT12.PLSR, ncomp = 4, newdata =
    slUTAll12.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- slTrain12.tmp$massUT
slUTAll12.df$massFinalT <- unscale.fn(slUTAll12.df$predictMass)

slUTAll12.df$massFinal <- backLog.Tns(slUTAll12.df$massFinalT)

slUTMass12.df <- data.frame(slUTAll12.df$plot, slUTAll12.df$position,
    slUTAll12.df$spAbbr,
slUTAll12.df$funGroup, slUTAll12.df$massFinal)
names(slUTMass12.df) <- c("plot", "position", "spAbbr", "funGroup", "massFinal
    ")

sl12Test.df <- merge(slTrain12.tmp, slUTMass12.df, by = c("plot", "position"))
sl12Test.df$subtract <- sl12Test.df$mass - sl12Test.df$massFinal
sl12Test.df


##########################
## PREDICTION Accuracy ##
##########################

slCV <- 0.2706
RMSEsl <- data.frame(slCV)

RMSEsl$slBICUn  <- unscale(slCV, slTrain12.tmp$massUT)
```

```
RMSEsl$slBIC <- backLog.Tns(RMSEsl$slBICUn)

sl12TEST.df <- data.frame(predict(slUT12.PLSR, ncomp = 4, newdata = sl12.df))
names(sl12TEST.df) <- c("massUT")

unscale.object <- slTrain12.tmp$massUT

sl12TEST.df$massFinalT <- unscale(sl12TEST.df$massUT, slTrain12.tmp$massUT)

sl12TEST.df$massFinal <- backLog.Tns(sl12TEST.df$massFinalT)

sl12TEST.df$massREAL <- slTrain12.tmp$mass

sl12TEST.df$subtract <- sl12TEST.df$massFinal - sl12TEST.df$massREAL
mean(sl12TEST.df$subtract)
sd(sl12TEST.df$subtract)

sl12LM <- lm(sl12TEST.df$massFinal ~ sl12TEST.df$massREAL -1 , offset = 1.00
    * sl12TEST.df$massREAL)
summary (sl12LM)

mean(sl12TEST.df$massREAL)
sd(sl12TEST.df$massREAL)



# LZCA Training Data Import --

lzcaStemLenMean <- subset(stemLenMeanTrain12.df, spAbbr == 'LZCA')

lzcaStemLenSum <- subset(stemLenSumTrain12.df, spAbbr == 'LZCA')

lzcaInflorLenMean <- subset(inflorLenMeanTrain12.df, spAbbr == 'LZCA')

lzcaInflorLenSum <- subset(inflorLenSumTrain12.df, spAbbr == 'LZCA')

lzcaMassMean <- subset(massMeanTrain12.df, spAbbr == 'LZCA')

lzcaTrain12.tmp <- merge(
merge(
merge(
merge(lzcaMassMean, lzcaStemLenMean, by = c("plot", "position", "spAbbr", "
    funGroup","plsr", "test")),
lzcaStemLenSum, by = c("plot", "position", "spAbbr", "funGroup", "plsr", "test
    ")),
```

```
lzcaInflorLenMean, by = c("plot", "position", "spAbbr", "funGroup", "plsr", "
    test")),
lzcaInflorLenSum, by = c("plot", "position", "spAbbr", "funGroup", "plsr", "
    test"))

# LZCA Training Data Transformation --

lzcaTrain12.tmp$massT <- Log.Tns(lzcaTrain12.tmp$mass)
lzcaTrain12.tmp$stemLenMeanT <- Sqrt.Tns(lzcaTrain12.tmp$stemLenMean)
lzcaTrain12.tmp$stemLenSumT <- Log.Tns(lzcaTrain12.tmp$stemLenSum)
lzcaTrain12.tmp$inflorLenMeanT <- Sqrt.Tns(lzcaTrain12.tmp$inflorLenMean)
lzcaTrain12.tmp$inflorLenSumT <- Sqrt.Tns(lzcaTrain12.tmp$inflorLenSum)

lzcaTrain12.tmp$massUT <- scale(lzcaTrain12.tmp$massT)
lzcaTrain12.tmp$stemLenMeanUT <- scale(lzcaTrain12.tmp$stemLenMeanT)
lzcaTrain12.tmp$stemLenSumUT <- scale(lzcaTrain12.tmp$stemLenSumT)
lzcaTrain12.tmp$inflorLenMeanUT <- scale(lzcaTrain12.tmp$inflorLenMeanT)
lzcaTrain12.tmp$inflorLenSumUT <- scale(lzcaTrain12.tmp$inflorLenSumT)

lzcaTrain12.sc <- lzcaTrain12.tmp [,17:21]

lzcaTrain12UTSq <- as.data.frame(
apply(lzcaTrain12.sc[,-1], 2,
function(x){ x^2 })
); names(lzcaTrain12UTSq) <- paste(
names(lzcaTrain12UTSq), "sq",
sep = "."
)


lzcaTrain12UTCb <- as.data.frame(
apply(lzcaTrain12.sc[,-1], 2,
function(x){ x^3 })
); names(lzcaTrain12UTCb) <- paste(
names(lzcaTrain12UTCb), "cb",
sep = "."
)

lzca12.df <- as.data.frame(cbind(lzcaTrain12.sc, lzcaTrain12UTSq,
    lzcaTrain12UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!
```

```
lzca12LMF <- lm(massUT ~
stemLenMeanUT  + stemLenMeanUT.sq + stemLenMeanUT.cb +
stemLenSumUT + stemLenSumUT.sq + stemLenSumUT.cb +
inflorLenMeanUT  + inflorLenMeanUT.sq + inflorLenMeanUT.cb +
inflorLenSumUT  + inflorLenSumUT.sq + inflorLenSumUT.cb,
data = lzca12.df
)

lzca12Subset <- expression(
(stemLenMeanUT | !stemLenMeanUT.sq) & (stemLenMeanUT & stemLenMeanUT.sq | !
    stemLenMeanUT.cb) &
(stemLenSumUT | !stemLenSumUT.sq) & (stemLenSumUT & stemLenSumUT.sq | !
    stemLenSumUT.cb) &
(inflorLenMeanUT | !inflorLenMeanUT.sq) & (inflorLenMeanUT & inflorLenMeanUT.
    sq | !inflorLenMeanUT.cb) &
(inflorLenSumUT | !inflorLenSumUT.sq) & (inflorLenSumUT & inflorLenSumUT.sq |
    !inflorLenSumUT.cb)
)

#AIC Dredge
# lzca12.AIC.dredge <-dredge(
# lzca12LMF,
# subset = lzca12Subset,
# rank = AIC
# )

# lzca12.AIC.top <- subset(
# lzca12.AIC.dredge,
# delta < 2
# )

#BIC Dredge
# lzca12.BIC.dredge <-dredge(
# lzca12LMF,
# subset = lzca12Subset,
# rank = BIC
# )

# lzca12.BIC.top <- subset(
# lzca12.BIC.dredge,
# delta < 2
# )

# LZCA Training PLSR Section -
```

```
lzcaUT12.PLSR <- plsr(lzca12.df$massUT ~ lzca12.df$stemLenSumUT + lzca12.df$
    stemLenSumUT.sq,
ncomp = 2, data = lzca12.df, validation = "CV", method = "oscorespls")

summary(lzcaUT12.PLSR)
# LZCA Data Import --

lzcaStemLenMean <- subset(stemLenMeanData12.df, spAbbr == 'LZCA')

lzcaStemLenSum <- subset(stemLenSumData12.df, spAbbr == 'LZCA')

lzcaInflorLenMean <- subset(inflorLenMeanData12.df, spAbbr == 'LZCA')

lzcaInflorLenSum <- subset(inflorLenSumData12.df, spAbbr == 'LZCA')

lzcaAll12.tmp <-
merge(
merge(
merge(lzcaStemLenSum, lzcaStemLenMean, by = c("plot", "position", "spAbbr", "
    funGroup","core")),
lzcaInflorLenMean, by = c("plot", "position", "spAbbr", "funGroup", "core")),
lzcaInflorLenSum, by = c("plot", "position", "spAbbr", "funGroup", "core"))

# LZCA Data Transformation --

lzcaUTAll12.df <- lzcaAll12.tmp

lzcaUTAll12.df$stemLenMeanT <- Sqrt.Tns(lzcaUTAll12.df$stemLenMean)
lzcaUTAll12.df$stemLenSumT <- Log.Tns(lzcaUTAll12.df$stemLenSum)
lzcaUTAll12.df$inflorLenMeanT <- Sqrt.Tns(lzcaUTAll12.df$inflorLenMean)
lzcaUTAll12.df$inflorLenSumT <- Sqrt.Tns(lzcaUTAll12.df$inflorLenSum)

lzcaUTAll12.df$stemLenMeanUT <- PredU(lzcaUTAll12.df$stemLenMeanT, lzcaTrain12
    .tmp$stemLenMeanT)
lzcaUTAll12.df$inflorLenMeanUT <- PredU(lzcaUTAll12.df$inflorLenMeanT,
    lzcaTrain12.tmp$inflorLenMeanT)
lzcaUTAll12.df$inflorLenSumUT <- PredU(lzcaUTAll12.df$inflorLenSumT,
    lzcaTrain12.tmp$inflorLenSumT)

lzcaUTAll12.df$stemLenSumUT <- PredU(lzcaUTAll12.df$stemLenSumT, lzcaTrain12.
    tmp$stemLenSumT)
lzcaUTAll12.df$stemLenSumUT.sq <- (lzcaUTAll12.df$stemLenSumUT)^2
# LZCA Data Prediction -

lzcaUTAll12.df$mx <- as.matrix(lzcaUTAll12.df [c(17:18)])
```

```
lzcaUTAll12.df$predictMass <- predict(lzcaUT12.PLSR, ncomp = 2, newdata =
    lzcaUTAll12.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- lzcaTrain12.tmp$massUT
lzcaUTAll12.df$massFinalT <- unscale.fn(lzcaUTAll12.df$predictMass)

lzcaUTAll12.df$massFinal <- backLog.Tns(lzcaUTAll12.df$massFinalT)

lzcaUTMass12.df <- data.frame(lzcaUTAll12.df$plot, lzcaUTAll12.df$position,
    lzcaUTAll12.df$spAbbr,
lzcaUTAll12.df$funGroup, lzcaUTAll12.df$massFinal)
names(lzcaUTMass12.df) <- c("plot", "position", "spAbbr", "funGroup", "
    massFinal")

lzca12Test.df <- merge(lzcaTrain12.tmp, lzcaUTMass12.df, by = c("plot", "
    position"))
lzca12Test.df$subtract <- lzca12Test.df$mass - lzca12Test.df$massFinal
lzca12Test.df



##########################
## PREDICTION Accuracy ##
##########################

lzcaCV <- 0.3166
RMSElzca <- data.frame(lzcaCV)

RMSElzca$lzcaBICUn  <- unscale(lzcaCV, lzcaTrain12.tmp$massUT)
RMSElzca$lzcaBIC <- backLog.Tns(RMSElzca$lzcaBICUn)

lzca12TEST.df <- data.frame(predict(lzcaUT12.PLSR, ncomp = 2, newdata = lzca12
    .df))
names(lzca12TEST.df) <- c("massUT")

unscale.object <- lzcaTrain12.tmp$massUT

lzca12TEST.df$massFinalT <- unscale(lzca12TEST.df$massUT, lzcaTrain12.tmp$
    massUT)

lzca12TEST.df$massFinal <- backLog.Tns(lzca12TEST.df$massFinalT)

lzca12TEST.df$massREAL <- lzcaTrain12.tmp$mass
```

```
lzca12TEST.df$subtract <- lzca12TEST.df$massFinal - lzca12TEST.df$massREAL
mean(lzca12TEST.df$subtract)
sd(lzca12TEST.df$subtract)

lzca12LM <- lm(lzca12TEST.df$massFinal ~ lzca12TEST.df$massREAL -1 ,  offset =
    1.00 * lzca12TEST.df$massREAL)
summary (lzca12LM)

mean(lzca12TEST.df$massREAL)
sd(lzca12TEST.df$massREAL)


# DC Training Data Import --

dcStemCount <- subset(mainTrain12.df, spAbbr == 'DC'& dataType == 'dcStemNo',
select = c(plot, position, spAbbr, funGroup, plsr, data))

names(dcStemCount) <- c("plot","position", "spAbbr", "funGroup", "plsr", "
    dcStemNo")

dcBCirc <- subset(mainTrain12.df, spAbbr == 'DC'& dataType == 'bCirc',
select = c(plot, position, spAbbr, funGroup, plsr, data))

names(dcBCirc) <- c("plot","position", "spAbbr", "funGroup", "plsr", "bCirc")

dcCirc30 <- subset(mainTrain12.df, spAbbr == 'DC'& dataType == 'circ30',
select = c(plot, position, spAbbr, funGroup, plsr, data))

names(dcCirc30) <- c("plot","position", "spAbbr", "funGroup", "plsr", "circ30"
    )

dcMassMean <- subset(massMeanTrain12.df, spAbbr == 'DC')

dcTrain12.tmp <-
merge(
merge(
merge(dcMassMean, dcCirc30, by = c("plot", "position", "spAbbr", "funGroup","
    plsr")),
dcStemCount, by = c("plot", "position", "spAbbr", "funGroup", "plsr")),
dcBCirc, by = c("plot", "position", "spAbbr", "funGroup", "plsr"))

# DC Training Transformation --

dcTrain12.tmp$massT <- Sqrt.Tns(dcTrain12.tmp$mass)
dcTrain12.tmp$bCircT <- Raw.Tns(dcTrain12.tmp$bCirc)
```

```
dcTrain12.tmp$circ30T <- Raw.Tns(dcTrain12.tmp$circ30)
dcTrain12.tmp$dcStemNoT <- Raw.Tns(dcTrain12.tmp$dcStemNo)

dcTrain12.tmp$massUT <- scale(dcTrain12.tmp$massT)
dcTrain12.tmp$bCircUT <- scale(dcTrain12.tmp$bCircT)
dcTrain12.tmp$circ30UT <- scale(dcTrain12.tmp$circ30T)
dcTrain12.tmp$dcStemNoUT <- scale(dcTrain12.tmp$dcStemNoT)

dcTrain12.sc <- dcTrain12.tmp [,15:18]

dcTrain12UTSq <- as.data.frame(
apply(dcTrain12.sc[,-1], 2,
function(x){ x^2 })
); names(dcTrain12UTSq) <- paste(
names(dcTrain12UTSq), "sq",
sep = "."
)


dcTrain12UTCb <- as.data.frame(
apply(dcTrain12.sc[,-1], 2,
function(x){ x^3 })
); names(dcTrain12UTCb) <- paste(
names(dcTrain12UTCb), "cb",
sep = "."
)

dc12.df <- as.data.frame(cbind(dcTrain12.sc, dcTrain12UTSq, dcTrain12UTCb))


# OC AIC / BIC Models -
# This section will undergo multi-model variable selection to determine
# the best variables for running the PLSR model. CAUTION: Running the full
# model takes a long time.  Grab coffee!

dc12LMF <- lm(massUT ~
bCircUT  + bCircUT.sq + bCircUT.cb +
circ30UT  + circ30UT.sq + circ30UT.cb +
dcStemNoUT  + dcStemNoUT.sq + dcStemNoUT.cb,
data = dc12.df
)

dc12Subset <- expression(
(bCircUT | !bCircUT.sq) & (bCircUT & bCircUT.sq | !bCircUT.cb) &
(circ30UT | !circ30UT.sq) & (circ30UT & circ30UT.sq | !circ30UT.cb) &
(dcStemNoUT | !dcStemNoUT.sq) & (dcStemNoUT & dcStemNoUT.sq | !dcStemNoUT.cb)
```

```
)

#AIC Dredge
# dc12.AIC.dredge <-dredge(
# dc12LMF,
# subset = dc12Subset,
# rank = AIC
# )

# dc12.AIC.top <- subset(
# dc12.AIC.dredge,
# delta < 2
# )

#BIC Dredge
# dc12.BIC.dredge <-dredge(
# dc12LMF,
# subset = dc12Subset,
# rank = BIC
# )

# dc12.BIC.top <- subset(
# dc12.BIC.dredge,
# delta < 2
# )

# DC Training PLSR Section -

dcUT12.PLSR <- plsr(dc12.df$massUT ~ dc12.df$bCircUT + dc12.df$circ30UT,
ncomp = 2, data = dc12.df, validation = "CV",
method = "oscorespls")

summary(dcUT12.PLSR)

# DC Data Import --

dcStemCount <- subset(data12.df, spAbbr == 'DC'& dataType == 'dcStemNo',
select = c(plot, position, spAbbr, funGroup, core, data))

names(dcStemCount) <- c("plot","position", "spAbbr", "funGroup", "core", "
    dcStemNo")

dcBCirc <- subset(data12.df, spAbbr == 'DC'& dataType == 'bCirc',
select = c(plot, position, spAbbr, funGroup, core, data))

names(dcBCirc) <- c("plot","position", "spAbbr", "funGroup", "core", "bCirc")
```

```
dcCirc30 <- subset(data12.df, spAbbr == 'DC'& dataType == 'circ30',
select = c(plot, position, spAbbr, funGroup, core, data))

names(dcCirc30) <- c("plot","position", "spAbbr", "funGroup", "core", "circ30"
    )


dcAll12.tmp <-
merge(
merge(dcBCirc, dcCirc30, by = c("plot", "position", "spAbbr", "funGroup","core
    ")),
dcStemCount, by = c("plot", "position", "spAbbr", "funGroup", "core"))

# DC Data Transformation --

dcUTAll12.df <- dcAll12.tmp

dcUTAll12.df$bCircT <- Raw.Tns(dcUTAll12.df$bCirc)
dcUTAll12.df$circ30T <- Raw.Tns(dcUTAll12.df$circ30)
dcUTAll12.df$dcStemNoT <- Raw.Tns(dcUTAll12.df$dcStemNo)

dcUTAll12.df$bCircUT <- PredU(dcUTAll12.df$bCircT, dcTrain12.tmp$bCircT)
dcUTAll12.df$circ30UT <- PredU(dcUTAll12.df$circ30T, dcTrain12.tmp$circ30T)
dcUTAll12.df$dcStemNoUT <- PredU(dcUTAll12.df$dcStemNoT, dcTrain12.tmp$
    dcStemNoT)

# DC Data Prediction -

dcUTAll12.df$mx <- as.matrix(dcUTAll12.df [c(12:13)])

dcUTAll12.df$predictMass <- predict(dcUT12.PLSR, ncomp = 1, newdata =
    dcUTAll12.df$mx)

## Insert the object to form the basis for unscaling the predicted result.
## This object should contain the scaling attributes from the original scaling

unscale.object <- dcTrain12.tmp$massUT
dcUTAll12.df$massFinalT <- unscale.fn(dcUTAll12.df$predictMass)

dcUTAll12.df$massFinal <- backSqrt.Tns(dcUTAll12.df$massFinalT)

dcUTMass12.df <- data.frame(dcUTAll12.df$plot, dcUTAll12.df$position,
    dcUTAll12.df$spAbbr,
dcUTAll12.df$funGroup, dcUTAll12.df$massFinal)
```

```
names(dcUTMass12.df) <- c("plot", "position", "spAbbr", "funGroup", "massFinal
    ")

dc12Test.df <- merge(dcTrain12.tmp, dcUTMass12.df, by = c("plot", "position",
    "spAbbr", "funGroup"))
dc12Test.df$subtract <- dc12Test.df$mass - dc12Test.df$massFinal
dc12Test.df




##########################
## PREDICTION Accuracy ##
##########################

dcCV <- 0.5580
RMSEdc <- data.frame(dcCV)

RMSEdc$dcBICUn  <- unscale(dcCV, dcTrain12.tmp$massUT)
RMSEdc$dcBIC <- backSqrt.Tns(RMSEdc$dcBICUn)

dc12TEST.df <- data.frame(predict(dcUT12.PLSR, ncomp = 1, newdata = dc12.df))
names(dc12TEST.df) <- c("massUT")

unscale.object <- dcTrain12.tmp$massUT

dc12TEST.df$massFinalT <- unscale(dc12TEST.df$massUT, dcTrain12.tmp$massUT)

dc12TEST.df$massFinal <- backSqrt.Tns(dc12TEST.df$massFinalT)

dc12TEST.df$massREAL <- dcTrain12.tmp$mass

dc12TEST.df$subtract <- dc12TEST.df$massFinal - dc12TEST.df$massREAL
mean(dc12TEST.df$subtract)
sd(dc12TEST.df$subtract)

dc12LM <- lm(dc12TEST.df$massFinal ~ dc12TEST.df$massREAL -1 ,  offset = 1.00
    * dc12TEST.df$massREAL)
summary (dc12LM)

mean(dc12TEST.df$massREAL)
sd(dc12TEST.df$massREAL)
```

```
# Plot Level Biomass Prediction -

bindUT12.df <- rbind(agUTMass12.df, pvUTMass12.df, dcUTMass12.df, lzcaUTMass12
    .df,
slUTMass12.df, lcUTMass12.df)

plotUT12.df <- aggregate(bindUT12.df$massFinal, by=list(bindUT12.df$plot), FUN
     = sum, na.rm = TRUE)
names(plotUT12.df) <- c("plot", "biomassPred")

plotUT12.final <- unique(merge(plotUT12.df, plot12Info.df[,-2:-4], by = "plot"
    ))

biomass12 <- aov(Log.Tns(plotUT12.final$biomassPred) ~ treatment * amf, data =
     plotUT12.final)

TukeyHSD(biomass12, ordered = TRUE)

# For Soil Food Web Data
plotUT12.web <- data.frame(plotUT12.final$plot, plotUT12.final$biomassPred)
names(plotUT12.web) <- c("plot", "plantBio")

write.csv(plotUT12.web, "C:\\Users\\Ohsowski\\Desktop\\plantBio.csv")
# Species Dataframe --
agMass12.df <- aggregate(agUTMass12.df$massFinal,
by=list(agUTMass12.df$plot, agUTMass12.df$spAbbr),
FUN = sum, na.rm = TRUE)

pvMass12.df <- aggregate(pvUTMass12.df$massFinal,
by=list(pvUTMass12.df$plot, pvUTMass12.df$spAbbr),
FUN = sum, na.rm = TRUE)

lcMass12.df <- aggregate(lcUTMass12.df$massFinal,
by=list(lcUTMass12.df$plot, lcUTMass12.df$spAbbr),
FUN = sum, na.rm = TRUE)

slMass12.df <- aggregate(slUTMass12.df$massFinal,
by=list(slUTMass12.df$plot, slUTMass12.df$spAbbr),
FUN = sum, na.rm = TRUE)

lzcaMass12.df <- aggregate(lzcaUTMass12.df$massFinal,
by=list(lzcaUTMass12.df$plot, lzcaUTMass12.df$spAbbr),
FUN = sum, na.rm = TRUE)

dcMass12.df <- aggregate(dcUTMass12.df$massFinal,
```

```
by=list(dcUTMass12.df$plot, dcUTMass12.df$spAbbr),
FUN = sum, na.rm = TRUE)


speciesUT12.df <- rbind(agMass12.df, pvMass12.df, lcMass12.df, lzcaMass12.df,
slMass12.df, dcMass12.df)
names(speciesUT12.df) <- c("plot", "spAbbr", "mass")

plot12Info.df1 <- plot12Info.df[,-2:-4]
species12.final <- unique(merge(speciesUT12.df, plot12Info.df1, by = c("plot")
    ))

qqnorm(Log.Tns(dcUTMass12.df$massFinal))
qqline(Log.Tns(dcUTMass12.df$massFinal))

write.csv(species12.final, "C:\\Users\\Ohsowski\\Desktop\\species.csv")

#Plot Level Biomass Prediction Grass ONLY-

grassUT12.df <- rbind(agUTMass12.df, pvUTMass12.df)

grassUT12.df <- aggregate(grassUT12.df$massFinal, by=list(grassUT12.df$plot),
FUN = sum, na.rm = TRUE)

names(grassUT12.df) <- c("plot", "biomassPred")

grassUT12.final <- unique(merge(grassUT12.df , plot12Info.df[,-2:-4], by = "
    plot"))

grass12 <- aov(Log.Tns(grassUT12.final$biomassPred) ~ treatment * amf, data =
    grassUT12.final)

summary(grass12)
TukeyHSD(grass12, ordered = TRUE)

# Plot Level Biomass Prediction NF ONLY-

nfUT12.df <- rbind(dcUTMass12.df, lzcaUTMass12.df)

nfUT12.df <- aggregate(nfUT12.df$massFinal, by=list(nfUT12.df$plot),
FUN = sum, na.rm = TRUE)

names(nfUT12.df) <- c("plot", "biomassPred")

nfUT12.final <- unique(merge(nfUT12.df, plot12Info.df[,-2:-4], by = "plot"))
```

```
nf12 <- aov(Log.Tns(nfUT12.final$biomassPred) ~ treatment * amf, data = nfUT12
    .final)

summary(nf12)
TukeyHSD(nf12, ordered = TRUE)


PRED <- read.csv("C:/Users/Ohsowski/Documents/PhD/Dissertation/Data/Plant
    Biomass Prediction 2012/plsrThesis_plant_prediction_14feb12.csv")

PRED <- subset(PRED, species == 'lzca' & year == 2012)
summary(PRED)


p <- ggplot(PRED, aes(massREAL, massUTBIC))

p + geom_point() +
scale_y_continuous(limits = c(0, 20)) +
scale_x_continuous(limits = c(0, 20))


########################
## PLOT BIOMASS DATA ##
########################

# Multi-season Plot Data Merging -
sumBio.df <- rbind(plotUT11.final, plotUT12.final)

sumBio.df$massLog <- log10(sumBio.df$biomassPred)
sumBio.df$massRR <- RecipRoot.Tns(sumBio.df$biomassPred)
sumBio.df$season <- ifelse(sumBio.df$season == "2011","A", "B")

qqnorm(sumBio.df$biomassPred)

sumBio.df$plot <- as.factor(sumBio.df$plot)
sumBio.df$treatment <- as.factor(sumBio.df$treatment)
sumBio.df$amf <- as.factor(sumBio.df$amf)
sumBio.df$season <- as.factor(sumBio.df$season)


vignette("using-lsmeans", package="lsmeans")

mean(sumBio.df$biomassPred)

var(sumBio.df$biomassPred)
```

```
scatterplot.matrix( sumBio.df , ellipse = TRUE)

# Generalized Linear Mixed Effects Model


M.null <- lmer(massRR ~ 1 + (1|plot), data = sumBio.df)

M1 <-  lmer(massRR ~ treatment * season * amf  + (1|plot), data = sumBio.df)
anova(M.null, M1, test = "F")

M2 <- update(M1,.~. -treatment:season:amf)
drop1(M2)

M3 <- update(M2,.~. -treatment:season)
drop1(M3)

M4 <- update(M3,.~. -treatment:amf)
drop1(M4)

M5 <- update(M4,.~. - season:amf)
anova(M4, M5)
drop1(M5)

M6 <- update(M5,.~. - amf)
drop1(M6)
anova(M5, M6)

M7 <- update(M6,.~. - treatment)
anova(M6, M7)
drop1(M7)

anova(M.null, M1, M2, M3, M4, M5, M6, M7,  test = 'F')

plot(fitted(M6), residuals(M6),
xlab = "Fitted Values", ylab = "Residuals")
abline(h=0, lty=2)
lines(smooth.spline(fitted(M6), residuals(M6)))


plot(predict(M6, type="response"),
residuals(M6, type= "deviance"))

coefplot2(M6)
fullPvals <- pvals.fnc(M6)
```

```
fullPvals.df <- data.frame(rownames(fullPvals$fixed), fullPvals$fixed$MCMCmean
    ,
fullPvals$fixed$HPD95lower, fullPvals$fixed$HPD95upper,
fullPvals$fixed$pMCMC)

names(fullPvals.df) <- c("treatment", "MCMCmean", "lower", "upper", "pMCMC")
#Correct

totalMass <- aggregate(sumBio.df$biomassPred,
by = list(sumBio.df$season, sumBio.df$treatment),
FUN = mean, na.rm = TRUE)

names(totalMass) <- c("season", "treatment", "data")

totalMassSD <- aggregate(sumBio.df$biomassPred,
by = list(sumBio.df$season, sumBio.df$treatment),
FUN = sd, na.rm = TRUE)

names(totalMassSD) <- c("season", "treatment", "sd")

totalMass.gr <- merge(totalMass, totalMassSD, by = c("season", "treatment"))




graph <- ggplot(data = totalMass.gr, aes(x = treatment, y = data))

limits <- aes(ymax = totalMass.gr$data + totalMass.gr$sd,
ymin = totalMass.gr$data - totalMass.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
theme_bw() +

geom_bar(position = "dodge", stat = "identity") +
geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat = "
    identity") +
facet_grid(. ~ season) +

labs(x = "Carbon Amendment",
y = "Estimated Biomass\n dry mass (g)") +
theme(panel.grid.major = element_line(colour = 'grey85')) +
theme(panel.grid.minor = element_line(colour = 'grey85')) +
theme(legend.position = c(0, -0.40)) +
```

```
theme(axis.text.x = element_text(size = 11, angle = 60)) +
theme(axis.text.y = element_text(size = 13)) +
theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
theme(axis.title.y = element_text(size = 12, angle = 90)) +
theme(legend.title = element_blank()) +
theme(strip.background = element_rect(fill = 'grey85')) +
theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output



# Species Models
species.df <- rbind(species11.final, species12.final)
species.df <-  subset(species.df, mass > 0)

species.df$season <- ifelse(species.df$season == "2011", "A", "B")
species.df$plot <- as.factor(species.df$plot)
species.df$treatment <- as.factor(species.df$treatment)
species.df$amf <- as.factor(species.df$amf)
species.df$season <- as.factor(species.df$season)
species.df$spAbbr <- as.factor(species.df$spAbbr)



# AG Model
agLM.df <-  subset(species.df, spAbbr == "AG")
agLM.df$massLog <- log10(agLM.df$mass)

qqnorm(agLM.df$massLog)
qqline(agLM.df$massLog)

AG.null <- lmer(massLog ~ 1 + (1|plot), data = agLM.df)

AG1 <-  lmer(massLog ~ treatment * season * amf  + (1|plot), data = agLM.df)
drop1(AG1)
anova(AG.null, AG1)

AG2 <- update(AG1,.~. -treatment:season:amf)
drop1(AG2)

AG3 <- update(AG2,.~. -season:amf)
drop1(AG3)

AG4 <- update(AG3,.~. - treatment:amf)
drop1(AG4)
```

```
AG5 <- update(AG4,.~. - treatment:season)
drop1(AG5)

AG6 <- update(AG5,.~. - amf)
drop1(AG6)
anova(AG5, AG6)

AG7 <- update(AG6,.~. - treatment)
drop1(AG7)

anova(AG.null, AG1, AG2, AG3, AG4, AG5, AG6, AG7,  test = 'F')
anova(AG5, AG6)

plot(fitted(AG5), residuals(AG5),
xlab = "Fitted Values", ylab = "Residuals")
abline(h=0, lty=2)
lines(smooth.spline(fitted(AG5), residuals(AG5)))


plot(predict(AG5, type="response"),
residuals(AG5, type= "deviance"))

coefplot2(AG5)
#agPvals <- pvals.fnc(AG5)

agPvals.df <- data.frame(rownames(agPvals$fixed), agPvals$fixed$MCMCmean,
agPvals$fixed$HPD95lower, agPvals$fixed$HPD95upper,
agPvals$fixed$pMCMC)

names(agPvals.df) <- c("treatment", "MCMCmean", "lower", "upper", "pMCMC")
#Correct

agMass <- aggregate(agLM.df$mass,
by = list(agLM.df$season, agLM.df$treatment, agLM.df$amf),
FUN = mean, na.rm = TRUE)

names(agMass) <- c("season", "treatment", "amf", "data")

agMassSD <- aggregate(agLM.df$mass,
by = list(agLM.df$season, agLM.df$treatment, agLM.df$amf),
FUN = sd, na.rm = TRUE)

names(agMassSD) <- c("season", "treatment", "amf", "sd")

agMass.gr <- merge(agMass, agMassSD, by = c("season", "treatment", "amf"))
```

```
graph <- ggplot(data = agMass.gr, aes(x = treatment, y = data, fill = amf))

limits <- aes(ymax = agMass.gr$data + agMass.gr$sd,
ymin = agMass.gr$data - agMass.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
theme_bw() +

geom_bar(position = "dodge", stat = "identity") +
geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat = "
    identity") +
facet_grid(. ~ season) +

labs(x = "Carbon Amendment",
y = "Estimated Biomass\n dry mass (g)") +
scale_y_continuous(limits = c(0, 350)) +
theme(panel.grid.major = element_line(colour = 'grey85')) +
theme(panel.grid.minor = element_line(colour = 'grey85')) +
theme(legend.position = c(0, -0.40)) +
theme(axis.text.x = element_text(size = 11, angle = 60)) +
theme(axis.text.y = element_text(size = 13)) +
theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
theme(axis.title.y = element_text(size = 12, angle = 90)) +
theme(legend.title = element_blank()) +
theme(strip.background = element_rect(fill = 'grey85')) +
theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output



# PV Model
pvLM.df <-  subset(species.df, spAbbr == "PV")
pvLM.df$massLog <- log10(pvLM.df$mass)

qqnorm(pvLM.df$massLog)
qqline(pvLM.df$massLog)

PV.null <- lmer(massLog ~ 1 + (1|plot), data = pvLM.df)
```

```
PV1 <-  lmer(massLog ~ treatment * season * amf  + (1|plot), data = pvLM.df)
coefplot2(PV1)
drop1(PV1)

PV2 <- update(PV1,.~. -treatment:season:amf)
drop1(PV2)

PV3 <- update(PV2,.~. -treatment:season)
drop1(PV3)

PV4 <- update(PV3,.~. - treatment:amf)
drop1(PV4)

PV5 <- update(PV4,.~. - season:amf)
drop1(PV5)

PV6 <- update(PV5,.~. - treatment)
drop1(PV6)

PV7 <- update(PV6,.~. - amf)
drop1(PV7)

anova(PV.null, PV1, PV2, PV3, PV4, PV5, PV6, PV7,  test = 'F')
anova(PV6, PV7)

plot(fitted(PV6), residuals(PV6),
xlab = "Fitted Values", ylab = "Residuals")
abline(h=0, lty=2)
lines(smooth.spline(fitted(PV6), residuals(PV6)))


plot(predict(PV6, type="response"),
residuals(PV6, type= "deviance"))

coefplot2(PV6)
#pvPvals <- pvals.fnc(PV6)

pvPvals.df <- data.frame(rownames(pvPvals$fixed), pvPvals$fixed$MCMCmean,
pvPvals$fixed$HPD95lower, pvPvals$fixed$HPD95upper,
pvPvals$fixed$pMCMC)

names(pvPvals.df) <- c("treatment", "MCMCmean", "lower", "upper", "pMCMC")
#correct

pvMass <- aggregate(pvLM.df$mass,
```

```
by = list(pvLM.df$season, pvLM.df$amf),
FUN = mean, na.rm = TRUE)

names(pvMass) <- c("season", "amf", "data")

pvMassSD <- aggregate(pvLM.df$mass,
by = list(pvLM.df$season, pvLM.df$amf),
FUN = sd, na.rm = TRUE)

names(pvMassSD) <- c("season", "amf", "sd")

pvMass.gr <- merge(pvMass, pvMassSD, by = c("season", "amf"))




graph <- ggplot(data = pvMass.gr, aes(x = amf, y = data))

limits <- aes(ymax = pvMass.gr$data + pvMass.gr$sd,
ymin = pvMass.gr$data - pvMass.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
theme_bw() +

geom_bar(position = "dodge", stat = "identity") +
geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat = "
    identity") +
facet_grid(. ~ season) +

labs(x = "Carbon Amendment",
y = "Estimated Biomass\n dry mass (g)") +
scale_y_continuous(limits = c(0, 350)) +
theme(panel.grid.major = element_line(colour = 'grey85')) +
theme(panel.grid.minor = element_line(colour = 'grey85')) +
theme(legend.position = c(0, -0.40)) +
theme(axis.text.x = element_text(size = 11, angle = 60)) +
theme(axis.text.y = element_text(size = 13)) +
theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
theme(axis.title.y = element_text(size = 12, angle = 90)) +
theme(legend.title = element_blank()) +
theme(strip.background = element_rect(fill = 'grey85')) +
theme(strip.text.x = element_text(face = 'bold', size = 14))
```

```
graph_output



# DC Model
dcLM.df <-  subset(species.df, spAbbr == "DC")
dcLM.df$massLog <- log10(dcLM.df$mass)

qqnorm(dcLM.df$massLog)
qqline(dcLM.df$massLog)

DC.null <- lmer(massLog ~ 1 + (1|plot), data = dcLM.df)

DC1 <-  lmer(massLog ~ treatment * season * amf  + (1|plot), data = dcLM.df)
coefplot2(DC3)
drop1(DC1)
anova(DC.null, DC1)

DC2 <- update(DC1,.~. -treatment:season:amf)
drop1(DC2)

DC3 <- update(DC2,.~. -treatment:season)
drop1(DC3)

DC4 <- update(DC3,.~. - season:amf)
drop1(DC4)

DC5 <- update(DC4,.~. - treatment:amf)
drop1(DC5)

DC6 <- update(DC5,.~. - amf)
drop1(DC6)

DC7 <- update(DC6,.~. - treatment)
drop1(DC7)

anova(DC.null, DC1, DC2, DC3, DC4, DC5, DC6, DC7,  test = 'F')
anova(DC4, DC5)

plot(fitted(DC4), residuals(DC4),
xlab = "Fitted Values", ylab = "Residuals")
abline(h=0, lty=2)
lines(smooth.spline(fitted(DC4), residuals(DC4)))


plot(predict(DC4, type="response"),
```

```
residuals(DC4, type= "deviance"))

coefplot2(DC4)
#dcPvals <- pvals.fnc(DC4)

dcPvals.df <- data.frame(rownames(dcPvals$fixed), dcPvals$fixed$MCMCmean,
dcPvals$fixed$HPD95lower, dcPvals$fixed$HPD95upper,
dcPvals$fixed$pMCMC)

names(dcPvals.df) <- c("treatment", "MCMCmean", "lower", "upper", "pMCMC")
summary(DC4)



dcMass <- aggregate(dcLM.df$mass,
by = list(dcLM.df$season, dcLM.df$treatment, dcLM.df$amf),
FUN = mean, na.rm = TRUE)

names(dcMass) <- c("season", "treatment", "amf", "data")

dcMassSD <- aggregate(dcLM.df$mass,
by = list(dcLM.df$season, dcLM.df$treatment, dcLM.df$amf),
FUN = sd, na.rm = TRUE)

names(dcMassSD) <- c("season", "treatment", "amf", "sd")

dcMass.gr <- merge(dcMass, dcMassSD, by = c("season", "treatment", "amf"))




graph <- ggplot(data = dcMass.gr, aes(x = treatment, y = data, fill = amf))

limits <- aes(ymax = dcMass.gr$data + dcMass.gr$sd,
ymin = dcMass.gr$data - dcMass.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
theme_bw() +

geom_bar(position = "dodge", stat = "identity") +
geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat = "
    identity") +
facet_grid(. ~ season) +
```

```
labs(x = "Carbon Amendment",
y = "Estimated Biomass\n dry mass (g)") +
scale_y_continuous(limits = c(0, 350)) +
theme(panel.grid.major = element_line(colour = 'grey85')) +
theme(panel.grid.minor = element_line(colour = 'grey85')) +
theme(legend.position = c(0, -0.40)) +
theme(axis.text.x = element_text(size = 11, angle = 60)) +
theme(axis.text.y = element_text(size = 13)) +
theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
theme(axis.title.y = element_text(size = 12, angle = 90)) +
theme(legend.title = element_blank()) +
theme(strip.background = element_rect(fill = 'grey85')) +
theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output




#correct

# LC Model
lcLM.df <-  subset(species.df, spAbbr == "LC")
lcLM.df$massLog <- log10(lcLM.df$mass)

qqnorm(lcLM.df$massLog)
qqline(lcLM.df$massLog)


LC.null <- lmer(massLog ~ 1 + (1|plot), data = lcLM.df)

LC1 <-  lmer(massLog ~ treatment * season * amf  + (1|plot), data = lcLM.df)
coefplot2(LC1)
drop1(LC1)
anova(LC.null, LC1)

LC2 <- update(LC1,.~. -treatment:season:amf)
drop1(LC2)

LC3 <- update(LC2,.~. -treatment:season)
drop1(LC3)

LC4 <- update(LC3,.~. - treatment:amf)
drop1(LC4)
```

```
LC5 <- update(LC4,.~. - season:amf)
drop1(LC5)

LC6 <- update(LC5,.~. - treatment)
drop1(LC6)

LC7 <- update(LC6,.~. - amf)
drop1(LC7)

anova(LC.null, LC1, LC2, LC3, LC4, LC5, LC6, LC7,  test = 'F')
anova(LC1, LC2)
plot(fitted(LC7), residuals(LC7),
xlab = "Fitted Values", ylab = "Residuals")
abline(h=0, lty=2)
lines(smooth.spline(fitted(LC7), residuals(LC7)))

x <- c(AIC(LC.null),AIC(LC1),AIC(LC2), AIC(LC3), AIC(LC4),
AIC(LC5), AIC(LC6), AIC(LC7))
plot(predict(LC7, type="response"),
residuals(LC7, type= "deviance"))

coefplot2(LC1)
#lcPvals <- pvals.fnc(LC1)

lcPvals.df <- data.frame(rownames(lcPvals$fixed), lcPvals$fixed$MCMCmean,
lcPvals$fixed$HPD95lower, lcPvals$fixed$HPD95upper,
lcPvals$fixed$pMCMC)

names(lcPvals.df) <- c("treatment", "MCMCmean", "lower", "upper", "pMCMC")
#correct
lcMass <- aggregate(lcLM.df$mass,
by = list(lcLM.df$season, lcLM.df$treatment, lcLM.df$amf),
FUN = mean, na.rm = TRUE)

names(lcMass) <- c("season", "treatment", "amf", "data")

lcMassSD <- aggregate(lcLM.df$mass,
by = list(lcLM.df$season, lcLM.df$treatment, lcLM.df$amf),
FUN = sd, na.rm = TRUE)

names(lcMassSD) <- c("season", "treatment", "amf", "sd")

lcMass.gr <- merge(lcMass, lcMassSD, by = c("season", "treatment", "amf"))
```

```
graph <- ggplot(data = lcMass.gr, aes(x = treatment, y = data, fill = amf))

limits <- aes(ymax = lcMass.gr$data + lcMass.gr$sd,
ymin = lcMass.gr$data - lcMass.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
theme_bw() +

geom_bar(position = "dodge", stat = "identity") +
geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat = "
    identity") +
facet_grid(. ~ season) +

labs(x = "Carbon Amendment",
y = "Estimated Biomass\n dry mass (g)") +
scale_y_continuous(limits = c(0, 350)) +
theme(panel.grid.major = element_line(colour = 'grey85')) +
theme(panel.grid.minor = element_line(colour = 'grey85')) +
theme(legend.position = c(0, -0.40)) +
theme(axis.text.x = element_text(size = 11, angle = 60)) +
theme(axis.text.y = element_text(size = 13)) +
theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
theme(axis.title.y = element_text(size = 12, angle = 90)) +
theme(legend.title = element_blank()) +
theme(strip.background = element_rect(fill = 'grey85')) +
theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output




# SL Model
slLM.df <-  subset(species.df, spAbbr == "SL")
slLM.df$sqrtLog <- Sqrt.Tns(slLM.df$mass)

qqnorm(slLM.df$sqrtLog)
qqline(slLM.df$sqrtLog)
```

```
SL.null <- lmer(sqrtLog ~ 1 + (1|plot), data = slLM.df)

SL1 <-  lmer(sqrtLog ~ treatment * season * amf  + (1|plot), data = slLM.df)
coefplot2(SL1)
drop1(SL1)
anova(SL.null, SL1)

SL2 <- update(SL1,.~. -treatment:season:amf)
drop1(SL2)

SL3 <- update(SL2,.~. -treatment:season)
drop1(SL3)

SL4 <- update(SL3,.~. - treatment:amf)
drop1(SL4)

SL5 <- update(SL4,.~. - season:amf)
drop1(SL5)

SL6 <- update(SL5,.~. - treatment)
drop1(SL6)

SL7 <- update(SL6,.~. - amf)
drop1(SL7)

anova(SL.null, SL1, SL2, SL3, SL4, SL5, SL6, SL7,  test = 'F')

plot(fitted(SL7), residuals(SL7),
xlab = "Fitted Values", ylab = "Residuals")
abline(h=0, lty=2)
lines(smooth.spline(fitted(SL7), residuals(SL7)))


plot(predict(SL7, type="response"),
residuals(SL7, type= "deviance"))

coefplot2(SL7)
#slPvals <- pvals.fnc(SL7)

slPvals.df <- data.frame(rownames(slPvals$fixed), slPvals$fixed$MCMCmean,
slPvals$fixed$HPD95lower, slPvals$fixed$HPD95upper,
slPvals$fixed$pMCMC)

names(slPvals.df) <- c("treatment", "MCMCmean", "lower", "upper", "pMCMC")
```

```
slMass <- aggregate(slLM.df$mass,
by = list(slLM.df$season),
FUN = mean, na.rm = TRUE)

names(slMass) <- c("season","data")

slMassSD <- aggregate(slLM.df$mass,
by = list(slLM.df$season),
FUN = sd, na.rm = TRUE)

names(slMassSD) <- c("season", "sd")

slMass.gr <- merge(slMass, slMassSD, by = c("season"))




graph <- ggplot(data = slMass.gr, aes(x = season, y = data))

limits <- aes(ymax = slMass.gr$data + slMass.gr$sd,
ymin = slMass.gr$data - slMass.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
theme_bw() +

geom_bar(position = "dodge", stat = "identity") +
geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat = "
    identity") +

labs(x = "Carbon Amendment",
y = "Estimated Biomass\n dry mass (g)") +
scale_y_continuous(limits = c(0, 350)) +
theme(panel.grid.major = element_line(colour = 'grey85')) +
theme(panel.grid.minor = element_line(colour = 'grey85')) +
theme(legend.position = c(0, -0.40)) +
theme(axis.text.x = element_text(size = 11, angle = 60)) +
theme(axis.text.y = element_text(size = 13)) +
theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
theme(axis.title.y = element_text(size = 12, angle = 90)) +
theme(legend.title = element_blank()) +
theme(strip.background = element_rect(fill = 'grey85')) +
```

```
theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output




# LZCA Model
lzcaLM.df <-  subset(species.df, spAbbr == "LZCA")
lzcaLM.df$sqrtLog <- Sqrt.Tns(lzcaLM.df$mass)

qqnorm(lzcaLM.df$sqrtLog)
qqline(lzcaLM.df$sqrtLog)

LZCA.null <- lmer(sqrtLog ~ 1 + (1|plot), data = lzcaLM.df)

LZCA1 <-  lmer(sqrtLog ~ treatment * season * amf  + (1|plot), data = lzcaLM.
    df)
coefplot2(LZCA1)
drop1(LZCA1)

LZCA2 <- update(LZCA1,.~. -treatment:season:amf)
drop1(LZCA2)

LZCA3 <- update(LZCA2,.~. -treatment:season)
drop1(LZCA3)

LZCA4 <- update(LZCA3,.~. - treatment:amf)
drop1(LZCA4)

LZCA5 <- update(LZCA4,.~. - season:amf)
drop1(LZCA5)

LZCA6 <- update(LZCA5,.~. - treatment)
drop1(LZCA6)

LZCA7 <- update(LZCA6,.~. - amf)
drop1(LZCA7)

anova(LZCA.null, LZCA1, LZCA2, LZCA3, LZCA4, LZCA5, LZCA6, LZCA7,  test = 'F',
     REML = FALSE)
anova(LZCA.null, LZCA1)
plot(fitted(LZCA6), residuals(LZCA6),
xlab = "Fitted Values", ylab = "Residuals")
abline(h=0, lty=2)
```

```
lines(smooth.spline(fitted(LZCA1), residuals(LZCA1)))


plot(predict(LZCA5, type="response"),
residuals(LZCA5 type= "deviance"))

coefplot2(LZCA6)
#lzcaPvals <- pvals.fnc(LZCA6)

lzcaPvals.df <- data.frame(rownames(lzcaPvals$fixed), lzcaPvals$fixed$MCMCmean
    ,
lzcaPvals$fixed$HPD95lower, lzcaPvals$fixed$HPD95upper,
lzcaPvals$fixed$pMCMC)

names(lzcaPvals.df) <- c("treatment", "MCMCmean", "lower", "upper", "pMCMC")

x <- c(AIC(LZCA.null),AIC(LZCA1),AIC(LZCA2), AIC(LZCA3), AIC(LZCA4),
AIC(LZCA5), AIC(LZCA6), AIC(LZCA7))
#correct

lzcaMass <- aggregate(lzcaLM.df$mass,
by = list(lzcaLM.df$season, lzcaLM.df$amf),
FUN = mean, na.rm = TRUE)

names(lzcaMass) <- c("season", "amf", "data")

lzcaMassSD <- aggregate(lzcaLM.df$mass,
by = list(lzcaLM.df$season, lzcaLM.df$amf),
FUN = sd, na.rm = TRUE)

names(lzcaMassSD) <- c("season", "amf", "sd")

lzcaMass.gr <- merge(lzcaMass, lzcaMassSD, by = c("season", "amf"))




graph <- ggplot(data = lzcaMass.gr, aes(x = amf, y = data))

limits <- aes(ymax = lzcaMass.gr$data + lzcaMass.gr$sd,
ymin = lzcaMass.gr$data - lzcaMass.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
```

```
theme_bw() +

geom_bar(position = "dodge", stat = "identity") +
geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat = "
    identity") +
facet_grid(. ~ season) +

labs(x = "Carbon Amendment",
y = "Estimated Biomass\n dry mass (g)") +
scale_y_continuous(limits = c(0, 350)) +
theme(panel.grid.major = element_line(colour = 'grey85')) +
theme(panel.grid.minor = element_line(colour = 'grey85')) +
theme(legend.position = c(0, -0.40)) +
theme(axis.text.x = element_text(size = 11, angle = 60)) +
theme(axis.text.y = element_text(size = 13)) +
theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
theme(axis.title.y = element_text(size = 12, angle = 90)) +
theme(legend.title = element_blank()) +
theme(strip.background = element_rect(fill = 'grey85')) +
theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output



# Multiple barplot graph

graph <- ggplot(data = all.df, aes(x = treatment, y = mass, fill = factor(amf)
    ))

limits <- aes(ymax = all.df$mass + all.df$se,
ymin = all.df$mass - all.df$se)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
theme_bw() +

geom_bar(position = "dodge", stat = "identity") +
geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat = "
    identity") +
facet_grid(. ~ season) +

labs(x = "Carbon Amendment",
y = "Estimated Biomass\n dry mass (g)") +
```

```
scale_y_continuous(limits = c(0, 600)) +
theme(panel.grid.major = element_line(colour = 'grey85')) +
theme(panel.grid.minor = element_line(colour = 'grey85')) +
theme(legend.position = c(0, -0.40)) +
theme(axis.text.x = element_text(size = 11, angle = 60)) +
theme(axis.text.y = element_text(size = 13)) +
theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
theme(axis.title.y = element_text(size = 12, angle = 90)) +
theme(legend.title = element_blank()) +
theme(strip.background = element_rect(fill = 'grey85')) +
theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output
```

Appendix D

R Code for Seed Trial Plant Cover

```
# R-Script for Percent Cover Data

library(RODBC)
library(reshape)
library(Hmisc)
library(DMwR)
library(ggplot2)
library(pls)
library(car)
library(lme4)
library(rgl, pos=4)
library(mgcv, pos=4)
library(scatterplot3d)
library(multilevel)
library(lme4)
library(languageR)
library(lsmeans)
library(pbkrtest)
library(lmerTest)
library(HLMdiag)
library(nlme)
library(lattice)
library(influence.ME)
library(LMERConvenienceFunctions)

# Data Transformation Section

Cube.Tns <- function (x) { x ^ 3 }

Square.Tns <- function (x) { x ^ 2 }

Raw.Tns <- function (x) { x }

Sqrt.Tns <- function (x) { sqrt(x) }

Log.Tns <- function (x) { log10(x + 1) }
```

```
RecipRoot.Tns <- function (x) { -1 / sqrt(x) }

Recip.Tns <- function (x) { -1 / (x) }

InvSquare.Tns <- function (x) { -1 / (x ^ 2) }


# Back Transformations
unscale.fn <- function(x){ as.data.frame(unscale(x, unscale.object)) }

backLog.Tns <- function (x){ 10 ^ (x) - 1 }

backSqrt.Tns <- function (x){ (x) ^ 2 }

backRaw.Tns <- function (x){ (x) }

# Prediction Data Centered and Scaled to Unity

## This unity function compares the collected test plants to the plsr data
    analyzed
## x = test variable data; y = main data; required to have same mean /
    variance

PredU <- function (x,y){
center <- x - mean(y)
sdCenter <- sd(y - mean(y))
center / sdCenter
}

# Pearson's Method (Parametric Test)

PearsonsMethod <- function (x) {
cor(x, use = "complete.obs", method = "pearson")
}

# Spearman's Method (Non-Parametric Test)

SpearmansMethod <- function (x,y) {
cor(x, use = "complete.obs", method = "spearman")
}


# Data Import
channel <- odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/
    Data/dissertation_data_exp_2")
main.df <- sqlQuery(channel, "SELECT * FROM DATA_Q WHERE months = 'Fall'")
```

```
close(channel)

soil.tmp <- subset(main.df, dataType == 'soil_px')
soil.tmp$soil_pc <- soil.tmp$data / soil.tmp$gridSize

grass.tmp <- subset(main.df, dataType == 'grass_px')
grass.tmp$grass_pc <- grass.tmp$data / grass.tmp$gridSize

nf.tmp <- subset(main.df, dataType == 'nf_px')
nf.tmp$nf_pc <- nf.tmp$data / nf.tmp$gridSize

forb.tmp <- subset(main.df, dataType == 'forb_px')
forb.tmp$forb_pc <- forb.tmp$data / forb.tmp$gridSize

litter.tmp <- subset(main.df, dataType == 'litter_px')
litter.tmp$litter_pc <- litter.tmp$data / litter.tmp$gridSize

weed.tmp <- subset(main.df, dataType == 'weed_px')
weed.tmp$weed_pc <- weed.tmp$data / weed.tmp$gridSize

tmp <- merge(forb.tmp, nf.tmp, by = c("plot_id", "seasonNum"))
tmp2 <- merge(soil.tmp, grass.tmp, by = c("plot_id", "seasonNum"))
tmp3 <- merge(weed.tmp, litter.tmp, by = c("plot_id", "seasonNum"))

tmp4 <- merge(tmp, tmp2, by = c("plot_id", "seasonNum"))

tmp5 <- merge(tmp4, tmp3, by = c("plot_id", "seasonNum"))

final.df <- data.frame(tmp5$plot_id, tmp5$hgt_cm.y, tmp5$seasonNum, tmp5$
    compost_rate.x.x, tmp5$biochar_rate.x.x, tmp5$myco.x.x, tmp5$soil_pc, tmp5
    $grass_pc, tmp5$forb_pc, tmp5$nf_pc, tmp5$litter_pc, tmp5$weed_pc)

names(final.df) <- c("plot_id", "hgt", "season", "compost_rate", "biochar_rate
    ", "amf", "soil_pc", "grass_pc", "forb_pc", "nf_pc", "litter_pc","weed_pc"
    )

final.df$hgt <- final.df$hgt / 100

#final.df$season <- as.factor(final.df$season)

finalSUB.df <- subset(final.df, season == '3')
summary(finalSUB.df)

final.df$native_pc <- final.df$grass_pc + final.df$nf_pc + final.df$forb_pc
final.df$soil_litter_pc <- final.df$soil_pc + final.df$litter_pc
```

```
final.df$green_pc <- final.df$grass_pc + final.df$nf_pc + final.df$forb_pc +
    final.df$weed_pc
final.df$native_litter_pc <- final.df$grass_pc + final.df$nf_pc + final.df$
    forb_pc + final.df$litter_pc

final.df <- final.df[c(-193:-195, -169:-171),]

final.df$variable <- Log.Tns(final.df$native_pc)
qqnorm(final.df$variable)
qqline(final.df$variable)

# qqnorm(final.df$variable)
# qqline(final.df$variable)


#final.df$compost_rate <- as.factor(final.df$compost_rate)

#final.df$biochar_rate <- as.factor(final.df$biochar_rate)

#final.df <- final.df[c(-196, -197, -198),]

PC.null <- lmer(variable ~ 1 + (1|plot_id), data = final.df, REML = FALSE)

PC1 <- lmer(variable ~ compost_rate*biochar_rate*season*amf*hgt+
            (1|plot_id), data = final.df, REML = FALSE)
summary(PC1)
anova(PC1)
mcp.fnc(PC1)
anova(PC.null, PC1)
relLik(PC.null, PC1)

PC2 <- update(PC1,.~. -compost_rate : biochar_rate : season : amf: hgt)
summary(PC2)
anova(PC2)
anova(PC1, PC2)
relLik(PC1, PC2)

PC3 <- update(PC2,.~. -compost_rate:season:hgt )
summary(PC3)
anova(PC3)
anova(PC2, PC3)
relLik(PC2, PC3)

PC4 <- update(PC3,.~. -compost_rate:biochar_rate:hgt)
summary(PC4)
anova(PC4)
```

```
anova(PC3, PC4)
relLik(PC3, PC4)

pamer.fnc(PC4)
mcp.fnc(PC4)


############################## FINAL MODEL ############################3
PC5 <- update(PC4,.~. - biochar_rate:season:hgt )
summary(PC5)
anova(PC5)
anova(PC4, PC5)
relLik(PC4, PC5)

PC6 <- update(PC5,.~. -compost_rate:biochar_rate:season)
summary(PC6)
anova(PC6)
anova(PC5, PC6)
relLik(PC5, PC6)

PC7 <- update(PC6,.~. -compost_rate:hgt )
summary(PC7)
anova(PC7)
anova(PC6, PC7)
relLik(PC6, PC7)

PC8 <- update(PC7,.~. -biochar_rate:hgt )
summary(PC8)
anova(PC8)
anova(PC7, PC8)
relLik(PC7, PC8)

PC9 <- update(PC8,.~. -biochar_rate:season )
summary(PC9)
anova(PC9)
anova(PC8, PC9)
relLik(PC8, PC9)


PC10 <- update(PC9,.~. -season:hgt )
summary(PC10)
anova(PC10)
anova(PC9, PC10)
relLik(PC9, PC10)

PC11 <- update(PC10,.~. -compost_rate:biochar_rate )
```

```
summary(PC11)
anova(PC11)
anova(PC10, PC11)
relLik(PC10, PC11)

PC12 <- update(PC11,.~. -compost_rate:season)
summary(PC12)
anova(PC12)
anova(PC11, PC12)
relLik(PC11, PC12)

PC13 <- update(PC12,.~. -amf)
summary(PC13)
anova(PC13)
anova(PC12, PC13)
relLik(PC12, PC13)

PC14 <- update(PC13,.~. -hgt)
summary(PC14)
anova(PC14)
anova(PC13, PC14)
relLik(PC13, PC14)

# Wireframe Graph
plot(fitted(M13), residuals(M13), xlab = "Fitted Values", ylab = "Residuals")
  abline(h=0, lty=2) lines(smooth.spline(fitted(M13), residuals(M13)))

wireframe(native_pc ~ compost_rate * biochar_rate, data=final.df, xlab = "
    Compost Rate", ylab = "Biochar Rate)", drape = TRUE, colorkey = TRUE)

p <- wireframe(variable ~ compost_rate * biochar_rate, data=final.df)
npanel <- c(4, 2)
rotx <- c(-50, -80)
rotz <- seq(30, 300, length = npanel[1]+1)
update(p[rep(1, prod(npanel))], layout = npanel,
    panel = function(..., screen) {
        panel.wireframe(..., screen = list(z = rotz[current.column()],
                                           x = rotx[current.row()]))
    })
```

Appendix E

R Code for Soil Food Web Organisms

```
library(RODBC)
library(reshape)
library(Hmisc)
library(pls)
library(DMwR)
library(ggplot2)
library(lme4)
library(MuMIn)

library(multilevel)
library(lme4)
library(languageR)
library(lsmeans)
library(pbkrtest)
library(HH)
library(geoR)
library(car)
library(qcc)
library(pscl)

# Data Transformation Section

Cube.Tns <- function (x) { x ^ 3 }

Square.Tns <- function (x) { x ^ 2 }

Raw.Tns <- function (x) { x }

Sqrt.Tns <- function (x) { sqrt(x) }

Log.Tns <- function (x) { log10(x + 1) }

RecipRoot.Tns <- function (x) { -1 / sqrt(x) }
```

```
Recip.Tns <- function (x) { -1 / (x) }

InvSquare.Tns <- function (x) { -1 / (x ^ 2) }


# Back Transformations

unscale.fn <- function(x){ as.data.frame(unscale(x, unscale.object)) }

backLog.Tns <- function (x){ 10 ^ (x) - 1 }

backSqrt.Tns <- function (x){ (x) ^ 2 }

backRaw.Tns <- function (x){ (x) }

backRecipRoot.Tns <- function (x) {1 / (x) ^2}


# Prediction Data Centered and Scaled to Unity

## This unity function compares the collected test plants to the plsr data
    analyzed
## x = test variable data; y = main data; required to have same mean /
    variance

PredU <- function (x,y){
  center <- x - mean(y)
  sdCenter <- sd(y - mean(y))
  center / sdCenter
}

# Pearson's Method (Parametric Test)

PearsonsMethod <- function (x) {
  cor(x, use = "complete.obs", method = "pearson")
}

# Spearman's Method (Non-Parametric Test)

SpearmansMethod <- function (x,y) {
  cor(x, use = "complete.obs", method = "spearman")
}



# Plot Information Data Frame
```

```
channel <-
  odbcConnectAccess("C:/Users/Ohsowski/Documents/PhD/Dissertation/Data/
    dissertation_data_exp_1_13aug2")

plot12Info.df <- sqlQuery(channel, "SELECT plot, amf, treatment, biochar_rate,
     compost_rate, hgt FROM DATA_WEB")
close(channel)

plot12Info.df <- unique(merge(plotUT12.web, plot12Info.df, by = "plot"))

web.df <- read.csv("C:\\Users\\Ohsowski\\Documents\\PhD\\Dissertation\\Data\\
    Soil Food Web 2012\\soil_food_web_13aug13.csv")


#######################
## Bacterial Biomass ##
#######################

bactBio.df <- subset(web.df, dataType == 'bactBio', select = c(plot, data))

bactBio.df <- unique(merge(bactBio.df, plot12Info.df, by = "plot"))

bactBio.df <- as.data.frame(subset(bactBio.df, amf == 'Y' | amf == 'N'))

bactBio.df <- as.data.frame(subset(bactBio.df, data < 5.8))

# Poisson overdispersion test
qcc.overdispersion.test(bactBio.df$data, type="poisson")

# Log Transformed Data approximates normal distribution
bactBio.df$dataT <- Log.Tns(bactBio.df$data)

qqnorm(bactBio.df$data)
qqline(bactBio.df$data)

# Using generalized linear models to test treatment level data
# Poisson distributions assume mean = variance

mean(bactBio.df$data)
var(bactBio.df$data)

# frequency distribution of the collem data
bactBioBar <- barplot(as.vector(table(bactBio.df$data)), names.arg = seq(1:36)
    )
bactBioBar
```

```
boxplot(dataT ~ treatment + amf, data = bactBio.df)
# Null Model Establishment
bactBio.null <- lm(dataT ~ 1, data = bactBio.df)
summary(bactBio.null)

bactBio.M1 <- lm(dataT ~ treatment * amf * hgt, data = bactBio.df)
summary(bactBio.M1)
anova(bactBio.null, bactBio.M1)

bactBio.M2 <- update(bactBio.M1, . ~ . - treatment:amf:hgt)
summary(bactBio.M2)
anova(bactBio.M1, bactBio.M2)

bactBio.M3 <- update(bactBio.M2, . ~ . - amf:hgt)
summary(bactBio.M3)
anova(bactBio.M2, bactBio.M3)

bactBio.M4 <- update(bactBio.M3, . ~ . - amf:treatment)
summary(bactBio.M4)
anova(bactBio.M3, bactBio.M4)

bactBio.M5 <- update(bactBio.M4, . ~ . - treatment:hgt)
summary(bactBio.M5)
anova(bactBio.M4, bactBio.M5)

bactBio.M6 <- update(bactBio.M5, . ~ . - hgt)
summary(bactBio.M6)
anova(bactBio.M5, bactBio.M6)

bactBio.M7 <- update(bactBio.M6, . ~ . - amf)
summary(bactBio.M7)
anova(bactBio.M6, bactBio.M7)

anova(bactBio.null, bactBio.M1, bactBio.M2 ,bactBio.M3)

x <- c(AIC(bactBio.null),AIC(bactBio.M1),AIC(bactBio.M2), AIC(bactBio.M3)) #
    stores AIC values in a vector
delta <- x - min(x)
delta
# Model #3 is most appropriate

# Check Diagnostic plots
plot(bactBio.M2)

plot(predict(bactBio.M2 , type="response"),
```

```
residuals(bactBio.M2 , type= "deviance"))

plot(hatvalues(bactBio.M2))
plot(rstudent(bactBio.M2))
plot(cooks.distance(bactBio.M2))

influencePlot(bactBio.M2)

# Residuals look good for Bacterial biomass

summary(bactBio.M4)

> summary(bactBio.M4)

# Call:
# glm(formula = dataT ~ treatment, family = gaussian, data = bactBio.df)

# Deviance Residuals:
      # Min          1Q      Median          3Q         Max
# -0.292319  -0.100532   0.002878   0.097328   0.313709

# Coefficients:
                                  # Estimate Std. Error t value Pr(>|t|)
# (Intercept)                        0.33371    0.03114  10.715  < 2e-16 ***
# treatment05T/ha BC                -0.11594    0.04404  -2.632  0.00981 **
# treatment10T/ha BC                -0.10882    0.04404  -2.471  0.01516 *
# treatment20T/ha CP + 00T/ha BC     0.06404    0.04404   1.454  0.14907
# treatment20T/ha CP + 05T/ha BC     0.19008    0.04404   4.316 3.72e-05 ***
# treatment20T/ha CP + 10T/ha BC     0.19489    0.04469   4.361 3.12e-05 ***
# ---
# Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1     1

# (Dispersion parameter for gaussian family taken to be 0.01745883)

    # Null deviance: 3.4504  on 106  degrees of freedom
# Residual deviance: 1.7633  on 101  degrees of freedom
# AIC: -121.65

# Number of Fisher Scoring iterations: 2

# Create final coefficient and multiplicative estimates
bactBio.est <- cbind(Estimate = coef(bactBio.M4), confint(bactBio.M4))
bactBio.final <- exp(bactBio.est)
bactBio.final

# Merged Bact Biomass
```

```
bactBioMn <- aggregate(bactBio.df$data,by = list(bactBio.df$treatment), FUN =
    mean, na.rm = TRUE)

names(bactBioMn) <- c("treatment", "data")

bactBioSD <- aggregate(bactBio.df$data,by = list(bactBio.df$treatment), FUN =
    sd, na.rm = TRUE)

names(bactBioSD) <- c("treatment", "sd")

# No sig diff for amf terms

bactBio.gr <- merge(bactBioMn, bactBioSD, by = c("treatment"))


#Graph
graph <- ggplot(data = bactBio.gr, aes(x = treatment, y = data, fill = factor(
    amf)))

limits <- aes(ymax = bactBio.gr$data + bactBio.gr$sd, ymin = bactBio.gr$data -
     bactBio.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
  theme_bw() +

  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +

  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))
```

```
graph_output


####################
## Fungal Biomass ##
####################

funBio.df <- subset(web.df, dataType == 'funBio', select = c(plot, data))

funBio.df <- unique(merge(funBio.df, plot12Info.df, by = "plot"))

funBio.df <- subset(funBio.df, amf == 'Y' | amf == 'N')

boxcox(funBio.df$data,plotit=T)
boxcox(funBio.M3,lambda = seq(0.5, 1.5, by =0.05),plotit = T)

funBio.df$dataT <- Log.Tns(funBio.df$data)

mean(funBio.df$data)

var(funBio.df$data)

boxplot(dataT ~ treatment + amf, data = funBio.df)

# Null Model Establishment
funBio.null <- lm(dataT ~ 1, data = funBio.df)
summary(funBio.null)

funBio.M1 <- lm(dataT ~ treatment * amf * hgt, data = funBio.df)
summary(funBio.M1)
anova(funBio.null, funBio.M1)

funBio.M2 <- update(funBio.M1, . ~ . - treatment:amf:hgt)
summary(funBio.M2)
anova(funBio.M1, funBio.M2)

funBio.M3 <- update(funBio.M2, . ~ . - amf:hgt)
summary(funBio.M3)
anova(funBio.M2, funBio.M3)

funBio.M4 <- update(funBio.M3, . ~ . - amf:treatment)
summary(funBio.M4)
anova(funBio.M3, funBio.M4)

funBio.M5 <- update(funBio.M4, . ~ . - treatment:hgt)
summary(funBio.M5)
```

```
anova(funBio.M4, funBio.M5)

funBio.M6 <- update(funBio.M5, . ~ . - hgt)
summary(funBio.M6)
anova(funBio.M5, funBio.M6)

funBio.M7 <- update(funBio.M6, . ~ . - amf)
summary(funBio.M7)
anova(funBio.M6, funBio.M7)

anova(funBio.null, funBio.M1, funBio.M2 ,funBio.M3)

x <- c(AIC(funBio.null),AIC(funBio.M1),AIC(funBio.M2), AIC(funBio.M3)) #
    stores AIC values in a vector
delta <- x - min(x)
delta
# Model #3 is most appropriate

# Check Diagnostic plots
plot(funBio.M3)

plot(predict(funBio.M3 , type="response"),
residuals(funBio.M3 , type= "deviance"))

plot(hatvalues(funBio.M3))
plot(rstudent(funBio.M3))
plot(cooks.distance(funBio.M3))

influencePlot(funBio.M3)

summary(funBio.M4)

# Call:
# glm(formula = data ~ treatment, family = gaussian, data = funBio.df)

# Deviance Residuals:
     # Min          1Q      Median          3Q         Max
# -2.00556   -0.61528   -0.09444    0.60556    2.81667

# Coefficients:
                                  # Estimate Std. Error t value Pr(>|t|)
# (Intercept)                         1.2722     0.2244   5.670 1.34e-07 ***
# treatment05T/ha BC                 -0.1278     0.3173  -0.403   0.6880
# treatment10T/ha BC                 -0.3889     0.3173  -1.226   0.2232
# treatment20T/ha CP + 00T/ha BC      0.6222     0.3173   1.961   0.0526 .
# treatment20T/ha CP + 05T/ha BC      1.8111     0.3173   5.708 1.13e-07 ***
```

```
# treatment20T/ha CP + 10T/ha BC   1.8333      0.3173   5.778 8.27e-08 ***
# ---
# Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1      1

# (Dispersion parameter for gaussian family taken to be 0.906171)

# Null deviance: 179.769  on 107  degrees of freedom
# Residual deviance:  92.429  on 102  degrees of freedom
# AIC: 303.68

# Number of Fisher Scoring iterations: 2

# Merged FUNGAL Biomass
# No sig diff for amf terms



funBioMn <- aggregate(funBio.df$data, by = list(funBio.df$treatment), FUN =
    mean, na.rm = TRUE)

names(funBioMn) <- c("treatment", "data")

funBioSD <- aggregate(funBio.df$data, by = list(funBio.df$treatment), FUN = sd
    , na.rm = TRUE)

names(funBioSD) <- c("treatment", "sd")

funBio.gr <- merge(funBioMn, funBioSD, by = c("treatment"))



graph <- ggplot(data = funBio.gr, aes(x = treatment, y = data))

limits <- aes(ymax = funBio.gr$data + funBio.gr$sd, ymin = funBio.gr$data -
    funBio.gr$sd)

dodge <- position_dodge(width = 0.9)


graph_output <- graph +
  theme_bw() +

  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +
```

```
  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output


############################
## Collembella Abundance ##
############################

collem.df <- subset(web.df, dataType == 'collem', select = c(plot, data))

collem.df <- unique(merge(collem.df, plot12Info.df, by = "plot"))

collem.df <- subset(collem.df, amf == 'Y' | amf == 'N')

rownames(collem.df) <-NULL

# Using generalized linear models to test treatment level data
# Poisson distributions assume mean = variance

mean(collem.df$data)
#[1] 10.75

var(collem.df$data)
#[1] 488.7687

# Null Model Establishment
mean(collem.df$data)

var(collem.df$data)



# Null Model Establishment
collem.null <- glm.nb(data ~ 1, data = collem.df)
summary(collem.null)
```

```
collem.M1 <- glm.nb(data ~ treatment * amf * hgt, data = collem.df)
summary(collem.M1)
anova(collem.null, collem.M1)

collem.M2 <- update(collem.M1, . ~ . - treatment:amf:hgt)
summary(collem.M2)
anova(collem.M1, collem.M2)

collem.M3 <- update(collem.M2, . ~ . - treatment:hgt)
summary(collem.M3)
anova(collem.M2, collem.M3)

collem.M4 <- update(collem.M3, . ~ . - amf:treatment)
summary(collem.M4)
anova(collem.M3, collem.M4)

collem.M5 <- update(collem.M4, . ~ . - amf:hgt)
summary(collem.M5)
anova(collem.M4, collem.M5)

collem.M6 <- update(collem.M5, . ~ . - hgt)
summary(collem.M6)
anova(collem.M5, collem.M6)

collem.M7 <- update(collem.M6, . ~ . - amf)
summary(collem.M7)
anova(collem.M6, collem.M7)



x <- c(AIC(collem.null),AIC(collem.M1),AIC(collem.M2), AIC(collem.M3)) #
    stores AIC values in a vector
delta <- x - min(x)
delta
# Model #5 is most appropriate

# Check Diagnostic plots
plot(collem.M3)

plot(predict(collem.M3 , type="response"),
residuals(collem.M3, type= "deviance"))


plot(rstudent(collem.M3))
plot(cooks.distance(collem.M3))
```

```
influencePlot(collem.M3)

# Model #5 is most appropriate

glht_pois <- glht(collem.M3, linfct = mcp(treatment = "Tukey"))
summary(glht_pois)


# Residuals look good for Collembella abundance

collemMn <- aggregate(collem.df$data, by = list(collem.df$treatment, collem.df
    $amf),FUN = mean, na.rm = TRUE)

names(collemMn) <- c("treatment", "amf", "data")

collemSD <- aggregate(collem.df$data, by = list(collem.df$treatment, collem.df
    $amf), FUN = sd, na.rm = TRUE)

names(collemSD) <- c("treatment", "amf", "sd")

collem.gr <- merge(collemMn, collemSD, by = c("treatment", "amf"))



graph <- ggplot(data = collem.gr, aes(x = treatment, y = data, fill = factor(
    amf)))

limits <- aes(ymax = collem.gr$data + collem.gr$sd,
              ymin = collem.gr$data - collem.gr$sd)

dodge <- position_dodge(width = 0.9)

graph_output <- graph +
  theme_bw() +

  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +

  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
```

```
        theme(axis.text.y = element_text(size = 13)) +
        theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
        theme(axis.title.y = element_text(size = 12, angle = 90)) +
        theme(legend.title = element_blank()) +
        theme(strip.background = element_rect(fill = 'grey85')) +
        theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output


##################################
## Bacterial Feeding Nematodes ##
##################################

bactFDNem.df <- subset(web.df, dataType == 'bactFDNem', select = c(plot, data)
    )

bactFDNem.df <- unique(merge(bactFDNem.df, plot12Info.df, by = "plot"))

bactFDNem.df <- subset(bactFDNem.df, amf == 'Y' | amf == 'N')

rownames(bactFDNem.df) <- NULL

# Using generalized linear models to test treatment level data
# Poisson distributions assume mean = variance

mean(bactFDNem.df$data)

var(bactFDNem.df$data)

bactFDNem.null <- glm.nb(data ~ 1, data = bactFDNem.df)
summary(bactFDNem.null)

bactFDNem.M1 <- glm.nb(data ~ treatment * amf * hgt, data = bactFDNem.df)
summary(bactFDNem.M1)
anova(bactFDNem.null, bactFDNem.M1)

bactFDNem.M2 <- update(bactFDNem.M1, . ~ . - treatment:amf:hgt)
summary(bactFDNem.M2)
anova(bactFDNem.M1, bactFDNem.M2)

bactFDNem.M3 <- update(bactFDNem.M2, . ~ . - treatment:hgt)
summary(bactFDNem.M3)

bactFDNem.M4 <- update(bactFDNem.M3, . ~ . - amf:treatment)
anova(bactFDNem.M2, bactFDNem.M3)
```

```
summary(bactFDNem.M4)
anova(bactFDNem.M3, bactFDNem.M4)

bactFDNem.M5 <- update(bactFDNem.M4, . ~ . - amf:hgt)
summary(bactFDNem.M5)
anova(bactFDNem.M4, bactFDNem.M5)

bactFDNem.M6 <- update(bactFDNem.M5, . ~ . - hgt)
summary(bactFDNem.M6)
anova(bactFDNem.M5, bactFDNem.M6)

bactFDNem.M7 <- update(bactFDNem.M6, . ~ . - amf)
summary(bactFDNem.M7)
anova(bactFDNem.M6, bactFDNem.M7)



x <- c(AIC(bactFDNem.null),AIC(bactFDNem.M1),AIC(bactFDNem.M2),
     AIC(bactFDNem.M3)) # stores AIC values in a vector
delta <- x - min(x)
delta
# Model #5 is most appropriate

# Check Diagnostic plots
plot(bactFDNem.M3)

plot(predict(bactFDNem.M3 , type="response"),
residuals(bactFDNem.M3, type= "deviance"))


plot(rstudent(bactFDNem.M3))
plot(cooks.distance(bactFDNem.M3))

influencePlot(bactFDNem.M3)

# Merged Bact Feeding Nematodes Biomass
# No sig diff for amf terms

bactFDNemMn <- aggregate(bactFDNem.df$data, by = list(bactFDNem.df$treatment,
    bactFDNem.df$amf), FUN = mean, na.rm = TRUE)

names(bactFDNemMn) <- c("treatment", "amf", "data")

bactFDNemSD <- aggregate(bactFDNem.df$data, by = list(bactFDNem.df$treatment,
    bactFDNem.df$amf), FUN = sd, na.rm = TRUE)
```

```
names(bactFDNemSD) <- c("treatment", "amf", "sd")

bactFDNem.gr <- merge(bactFDNemMn, bactFDNemSD, by = c("treatment", "amf"))


graph <- ggplot(data = bactFDNem.gr, aes(x = treatment, y = data, fill =
    factor(amf)))

limits <- aes(ymax = bactFDNem.gr$data + bactFDNem.gr$sd,
              ymin = bactFDNem.gr$data - bactFDNem.gr$sd)

dodge <- position_dodge(width = 0.9)

graph_output <- graph +
  theme_bw() +

  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +

  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output


################################
## Fungal Feeding Nematodes ##
################################

funFDNem.df <- subset(web.df, dataType == 'funFDNem', select = c(plot, data))

funFDNem.df <- unique(merge(funFDNem.df, plot12Info.df, by = "plot"))

funFDNem.df <- subset(funFDNem.df, amf == 'Y' | amf == 'N')
```

```
rownames(funFDNem.df) <- NULL


# Using generalized linear models to test treatment level data
# Poisson distributions assume mean = variance

mean(funFDNem.df$data)

var(funFDNem.df$data)

funFDNem.null <- glm.nb(data ~ 1, data = funFDNem.df)
summary(funFDNem.null)

funFDNem.M1 <- glm.nb(data ~ treatment * amf * hgt, data = funFDNem.df)
summary(funFDNem.M1)
anova(funFDNem.null, funFDNem.M1)

funFDNem.M2 <- update(funFDNem.M1, . ~ . - treatment:amf:hgt)
summary(funFDNem.M2)
anova(funFDNem.M1, funFDNem.M2)

funFDNem.M3 <- update(funFDNem.M2, . ~ . - treatment:hgt)
summary(funFDNem.M3)

funFDNem.M4 <- update(funFDNem.M3, . ~ . - amf:treatment)
anova(funFDNem.M2, funFDNem.M3)
summary(funFDNem.M4)
anova(funFDNem.M3, funFDNem.M4)

funFDNem.M5 <- update(funFDNem.M4, . ~ . - amf:hgt)
summary(funFDNem.M5)
anova(funFDNem.M4, funFDNem.M5)

funFDNem.M6 <- update(funFDNem.M5, . ~ . - hgt)
summary(funFDNem.M6)
anova(funFDNem.M5, funFDNem.M6)

funFDNem.M7 <- update(funFDNem.M6, . ~ . - amf)
summary(funFDNem.M7)
anova(funFDNem.M6, funFDNem.M7)


x <- c(AIC(funFDNem.null),AIC(funFDNem.M1),AIC(funFDNem.M2),
     AIC(funFDNem.M3)) # stores AIC values in a vector
delta <- x - min(x)
delta
```

```
# Model #5 is most appropriate

# Check Diagnostic plots
plot(funFDNem.M3)

plot(predict(funFDNem.M3 , type="response"),
residuals(funFDNem.M3, type= "deviance"))


plot(rstudent(funFDNem.M3))
plot(cooks.distance(funFDNem.M3))

influencePlot(funFDNem.M3)


glht_pois <- glht(funFDNem.M5, linfct = mcp(treatment = "Tukey"))
summary(glht_pois)


# Merged Bact Feeding Nematodes Biomass
# No sig diff for amf terms
funFDNemMn <- aggregate(funFDNem.df$data, by = list(funFDNem.df$treatment,
    funFDNem.df$amf),
                                        FUN = mean, na.rm = TRUE)

names(funFDNemMn) <- c("treatment", "amf", "data")

funFDNemSD <- aggregate(funFDNem.df$data, by = list(funFDNem.df$treatment,
    funFDNem.df$amf), FUN = sd, na.rm = TRUE)

names(funFDNemSD) <- c("treatment", "amf", "sd")

funFDNem.gr <- merge(funFDNemMn, funFDNemSD, by = c("treatment", "amf"))


graph <- ggplot(data = funFDNem.gr, aes(x = treatment, y = data, fill = factor
    (amf)))

limits <- aes(ymax = funFDNem.gr$data + funFDNem.gr$sd,
              ymin = funFDNem.gr$data - funFDNem.gr$sd)

dodge <- position_dodge(width = 0.9)

graph_output <- graph +
  theme_bw() +
```

```
  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +

  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output


##########################
## Predatory Nematodes ##
##########################

predFDNem.df <- subset(web.df, dataType == 'predNem', select = c(plot, data))

predFDNem.df <- unique(merge(predFDNem.df, plot12Info.df, by = "plot"))

predFDNem.df <- subset(predFDNem.df, amf == 'Y' | amf == 'N')

rownames(predFDNem.df) <- NULL

predFDNem.df <- predFDNem.df[-43, ]


# Using generalized linear models to test treatment level data
# Poisson distributions assume mean = variance

mean(predFDNem.df$data)

var(predFDNem.df$data)


predFDNem.null <- glm.nb(data ~ 1, data = predFDNem.df)
summary(predFDNem.null)
```

```
predFDNem.M1 <- glm.nb(data ~ treatment * amf * hgt, data = predFDNem.df)
summary(predFDNem.M1)
anova(predFDNem.null, predFDNem.M1)

predFDNem.M2 <- update(predFDNem.M1, . ~ . - treatment:amf:hgt)
summary(predFDNem.M2)
anova(predFDNem.M1, predFDNem.M2)

predFDNem.M3 <- update(predFDNem.M2, . ~ . - treatment:hgt)
summary(predFDNem.M3)

predFDNem.M4 <- update(predFDNem.M3, . ~ . - amf:treatment)
anova(predFDNem.M2, predFDNem.M3)
summary(predFDNem.M4)
anova(predFDNem.M3, predFDNem.M4)

predFDNem.M5 <- update(predFDNem.M4, . ~ . - amf:hgt)
summary(predFDNem.M5)
anova(predFDNem.M4, predFDNem.M5)

predFDNem.M6 <- update(predFDNem.M5, . ~ . - hgt)
summary(predFDNem.M6)
anova(predFDNem.M5, predFDNem.M6)

predFDNem.M7 <- update(predFDNem.M6, . ~ . - amf)
summary(predFDNem.M7)
anova(predFDNem.M6, predFDNem.M7)


predFDNem.null <- glm.nb(data ~ 1, data = predFDNem.df)
summary(predFDNem.null)


predFDNem.M1 <- glm.nb(data ~ treatment * amf, data = predFDNem.df)
summary(predFDNem.M1)
anova(predFDNem.null, predFDNem.M1)

predFDNem.M2 <- update(predFDNem.M1, . ~ . - treatment:amf)
summary(predFDNem.M2)
anova(predFDNem.M1, predFDNem.M2)

predFDNem.M3 <- update(predFDNem.M2, . ~ . - amf)
summary(predFDNem.M3)
anova(predFDNem.M2, predFDNem.M3)
```

```
x <- c(AIC(predFDNem.null),AIC(predFDNem.M1),AIC(predFDNem.M2),
      AIC(predFDNem.M3)) # stores AIC values in a vector
delta <- x - min(x)
delta
# Model #5 is most appropriate

# Check Diagnostic plots
plot(predFDNem.M3)

plot(predict(predFDNem.M3 , type="response"),
residuals(predFDNem.M3, type= "deviance"))


plot(rstudent(predFDNem.M3))
plot(cooks.distance(predFDNem.M3))

influencePlot(predFDNem.M3)
glht_pois <- glht(predFDNem.M4, linfct = mcp(treatment = "Tukey"))
summary(glht_pois)

# Merged Bact Feeding Nematodes Biomass
# No sig diff for amf terms

predNemMn <- aggregate(predFDNem.df$data, by = list(predFDNem.df$treatment),
    FUN = mean, na.rm = TRUE)

names(predNemMn) <- c("treatment", "data")

predNemSD <- aggregate(predFDNem.df$data, by = list(predFDNem.df$treatment),
    FUN = sd, na.rm = TRUE)

names(predNemSD) <- c("treatment", "sd")

predNem.gr <- merge(predNemMn, predNemSD, by = c("treatment"))


graph <- ggplot(data = predNem.gr, aes(x = treatment, y = data))

limits <- aes(ymax = predNem.gr$data + predNem.gr$sd,
              ymin = predNem.gr$data - predNem.gr$sd)

dodge <- position_dodge(width = 0.9)

graph_output <- graph +
  theme_bw() +
```

```
  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +

  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output


############################
## Microbe Feeding Mites ##
############################

predMite.df <- subset(web.df, dataType == 'predMite', select = c(plot, data))

predMite.df <- unique(merge(predMite.df, plot12Info.df, by = "plot"))

predMite.df <- subset(predMite.df, amf == 'Y' | amf == 'N')

rownames(predMite.df) <- NULL
predMite.df <- predMite.df[-64, ]

# Using generalized linear models to test treatment level data
# Poisson distributions assume mean = variance

mean(predMite.df$data)

var(predMite.df$data)

predMite.null <- glm.nb(data ~ 1, data = predMite.df)
summary(predMite.null)

predMite.M1 <- glm.nb(data ~ treatment * amf * hgt, data = predMite.df)
summary(predMite.M1)
anova(predMite.null, predMite.M1)
```

```
predMite.M2 <- update(predMite.M1, . ~ . - treatment:amf:hgt)
summary(predMite.M2)
anova(predMite.M1, predMite.M2)

predMite.M3 <- update(predMite.M2, . ~ . - treatment:hgt)
summary(predMite.M3)

predMite.M4 <- update(predMite.M3, . ~ . - amf:treatment)
anova(predMite.M2, predMite.M3)
summary(predMite.M4)
anova(predMite.M3, predMite.M4)

predMite.M5 <- update(predMite.M4, . ~ . - amf:hgt)
summary(predMite.M5)
anova(predMite.M4, predMite.M5)

predMite.M6 <- update(predMite.M5, . ~ . - hgt)
summary(predMite.M6)
anova(predMite.M5, predMite.M6)

predMite.M7 <- update(predMite.M6, . ~ . - amf)
summary(predMite.M7)
anova(predMite.M6, predMite.M7)


x <- c(AIC(predMite.null),AIC(predMite.M1),AIC(predMite.M2),
     AIC(predMite.M3)) # stores AIC values in a vector
delta <- x - min(x)
delta
# Model #5 is most appropriate

# Check Diagnostic plots
plot(predMite.M3)

plot(predict(predMite.M3 , type="response"),
residuals(predMite.M3, type= "deviance"))


plot(rstudent(predMite.M3))
plot(cooks.distance(predMite.M3))

influencePlot(predMite.M3)

# Merged Bact Feeding Nematodes Biomass
# No sig diff for amf terms
```

```
predMiteMn <- aggregate(predMite.df$data, by = list(predMite.df$treatment),
                                    FUN = mean, na.rm = TRUE)

names(predMiteMn) <- c("treatment", "data")

predMiteSD <- aggregate(predMite.df$data, by = list(predMite.df$treatment),
                                    FUN = sd, na.rm = TRUE)

names(predMiteSD) <- c("treatment", "sd")

predMite.gr <- merge(predMiteMn, predMiteSD, by = c("treatment"))


graph <- ggplot(data = predMite.gr, aes(x = treatment, y = data))

limits <- aes(ymax = predMite.gr$data + predMite.gr$sd,
              ymin = predMite.gr$data - predMite.gr$sd)

dodge <- position_dodge(width = 0.9)

graph_output <- graph +
  theme_bw() +

  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +

  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output


####################
## Predatory Mites ##
```

```
####################

predMite.df <- subset(web.df, dataType == 'predMite', select = c(plot, data))

predMite.df <- unique(merge(predMite.df, plot12Info.df, by = "plot"))

predMite.df <- subset(predMite.df, amf == 'Y' | amf == 'N')

# Using generalized linear models to test treatment level data
# Poisson distributions assume mean = variance

mean(predMite.df$data)

var(predMite.df$data)


predMite.null <- glm.nb(data ~ 1, data = predMite.df)
summary(predMite.null)

predMite.M1 <- glm.nb(data ~ treatment * amf * hgt, data = predMite.df)
summary(predMite.M1)
anova(predMite.null, predMite.M1)

predMite.M2 <- update(predMite.M1, . ~ . - treatment:amf:hgt)
summary(predMite.M2)
anova(predMite.M1, predMite.M2)

predMite.M3 <- update(predMite.M2, . ~ . - treatment:hgt)
summary(predMite.M3)

predMite.M4 <- update(predMite.M3, . ~ . - amf:hgt)
anova(predMite.M2, predMite.M3)
summary(predMite.M4)
anova(predMite.M3, predMite.M4)

predMite.M5 <- update(predMite.M4, . ~ . - amf:hgt)
summary(predMite.M5)
anova(predMite.M4, predMite.M5)

predMite.M6 <- update(predMite.M5, . ~ . - hgt)
summary(predMite.M6)
anova(predMite.M5, predMite.M6)

predMite.M7 <- update(predMite.M6, . ~ . - amf)
summary(predMite.M7)
anova(predMite.M6, predMite.M7)
```

```
x <- c(AIC(predMite.null),AIC(predMite.M1),AIC(predMite.M2),
     AIC(predMite.M3)) # stores AIC values in a vector
delta <- x - min(x)
delta
# Model #5 is most appropriate

# Check Diagnostic plots
plot(predMite.M3)

plot(predict(predMite.M5 , type="response"),
residuals(predMite.M5, type= "deviance"))


plot(rstudent(predMite.M5))
plot(cooks.distance(predMite.M5))

influencePlot(predMite.M5)

# Merged Bact Feeding Nematodes Biomass
# No sig diff for amf terms

predMiteMn <- aggregate(predMite.df$data, by = list(predMite.df$treatment),
    FUN = mean, na.rm = TRUE)

names(predMiteMn) <- c("treatment", "data")

predMiteSD <- aggregate(predMite.df$data, by = list(predMite.df$treatment),
    FUN = sd, na.rm = TRUE)

names(predMiteSD) <- c("treatment", "sd")

predMite.gr <- merge(predMiteMn, predMiteSD, by = c("treatment"))



graph <- ggplot(data = predMite.gr, aes(x = treatment, y = data))

limits <- aes(ymax = predMite.gr$data + predMite.gr$sd,
              ymin = predMite.gr$data - predMite.gr$sd)

dodge <- position_dodge(width = 0.9)

graph_output <- graph +
  theme_bw() +
```

```
  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +

  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output



########################################
## FUNGAL / Bacterial Ratio Biomass ##
########################################

FBR.df <- subset(web.df, dataType == 'FBRatio', select = c(plot, data))

FBR.df <- unique(merge(FBR.df, plot12Info.df, by = "plot"))

FBR.df <- subset(FBR.df, amf == 'Y' | amf == 'N')

FBR.df <- FBR.df[c(-89, -102), ]


rownames(FBR.df) <- NULL

# Using generalized linear models to test treatment level data
# Poisson distributions assume mean = variance

mean(FBR.df$data)

var(FBR.df$data)


# frequency distribution of the collem data
```

```
fBRBar <- barplot(as.vector(table(FBR.df$data)), names.arg = seq(1:90))
fBRBar

# Poisson Distribution doesn't fit
points(fBRBar,dpois(seq(1:90),2.140768)* sum(table(FBR.df$data)),
      cex=2,type="b",col="sienna",lwd=2,pch=19)

#negative binomial distribution fits the data
points(fBRBar,dnbinom(seq(1,90), mu = mean(FBR.df$data),
      size=(mean(FBR.df$data)^2)/(var(FBR.df$data) - mean(FBR.df$data)))*sum(
    table(FBR.df$data)),
      type="b",cex=2,col="salmon",lwd=2,pch=19)


# Null Model Establishment
fBR.null <- glm(data ~ 1,data = FBR.df,family = poisson)
summary(fBR.null)


fBR.M1 <- glm(data ~ treatment * amf, data = FBR.df, family = poisson)
summary(fBR.M1)

# Model Comparison
anova(fBR.null, fBR.M1,test="Chisq")


# Create negative binomial distribution

fBR.M1 <- glm.nb(data  ~ treatment * amf * hgt, data = FBR.df)
summary(fBR.M1)
anova(fBR.null, fBR.M1)

fBR.M2 <- update(fBR.M1, . ~ . - treatment:amf:hgt)
summary(fBR.M2)
anova(fBR.M1, fBR.M2)

fBR.M3 <- update(fBR.M2, . ~ . - amf:hgt)
summary(fBR.M3)
anova(fBR.M2, fBR.M3)

fBR.M4 <- update(fBR.M3, . ~ . - amf:treatment)
summary(fBR.M4)
anova(fBR.M3, fBR.M4)

fBR.M5 <- update(fBR.M4, . ~ . - treatment:hgt)
summary(fBR.M5)
```

```
anova(fBR.M4, fBR.M5)

fBR.M6 <- update(fBR.M5, . ~ . - hgt)
summary(fBR.M6)
anova(fBR.M5, fBR.M6)

fBR.M7 <- update(fBR.M6, . ~ . - amf)
summary(fBR.M7)
anova(fBR.M6, fBR.M7)


x <- c(AIC(fBR.null),AIC(fBR.M1),AIC(fBR.M2),
     AIC(fBR.M3)) # stores AIC values in a vector
delta <- x - min(x)
delta

plot(fBR.M3)

summary(fBR.M3)

FBRMn <- aggregate(FBR.df$data, by = list(FBR.df$treatment), FUN = mean, na.rm
     = TRUE)

names(FBRMn) <- c("treatment", "data")

FBRSD <- aggregate(FBR.df$data, by = list(FBR.df$treatment), FUN = sd, na.rm =
     TRUE)

names(FBRSD) <- c("treatment", "sd")

FBR.gr <- merge(FBRMn, FBRSD, by = c("treatment"))


graph <- ggplot(data = FBR.gr, aes(x = treatment, y = data))

limits <- aes(ymax = FBR.gr$data + FBR.gr$sd,
              ymin = FBR.gr$data - FBR.gr$sd)

dodge <- position_dodge(width = 0.9)

graph_output <- graph +
  theme_bw() +

  geom_bar(position = "dodge", stat = "identity") +
  geom_errorbar(limits, width = 0.5, color = "black", position = dodge, stat =
     "identity") +
```

```
  labs(x = "Carbon Amendment",
       y = "Factor") +
         theme(panel.grid.major = element_line(colour = 'grey85')) +
         theme(panel.grid.minor = element_line(colour = 'grey85')) +
         theme(legend.position = c(0, -0.40)) +
         theme(axis.text.x = element_text(size = 11, angle = 60)) +
         theme(axis.text.y = element_text(size = 13)) +
         theme(axis.title.x = element_text(size = 12, vjust = 0.1)) +
         theme(axis.title.y = element_text(size = 12, angle = 90)) +
         theme(legend.title = element_blank()) +
         theme(strip.background = element_rect(fill = 'grey85')) +
         theme(strip.text.x = element_text(face = 'bold', size = 14))

graph_output
```