A Case Study Examining the Relationship Between Human Appropriation of Net Primary Productivity and Landscape Patterns in the US Great Lakes Basin

by

Erin M Barton, Conservation Ecology, U-M School For Environment and Sustainability

A thesis submitted as partial fulfillment of the requirements for the degree of Master of Science (School for Environment and Sustainability) in the University of Michigan

August 2018

Thesis Committee:

Dr. William S. Currie, Chair

Dr. Douglas R. Pearsall

Abstract

Human appropriation of net primary productivity (HANPP) has been proposed as a measure of human pressures on biodiversity; it represents the proportion of energy flow that was historically available to wildlife food webs but has been appropriated for human use, primarily through the harvesting of primary production. This study examined the spatial relationship between HANPP of managed terrestrial landscapes and two abiotic proxy metrics for biodiversity–landscape diversity and local connectedness. Our objectives were 1) to quantify patterns of HANPP in forestlands and croplands, comparing the extraction of NPP in a recent decade against the potential natural vegetation that largely existed on the US side of the Great Lakes prior to European settlement; and 2) to assess spatial patterns of HANPP in comparison to landscape diversity and local connectedness at the county scale across the region. Our analysis considered above and below-ground compartments of NPP and focused on the percent of potential NPP being appropriated (%$HANPP_0$). The mean area-weighted %$HANPP_0$ across our study region was 45%, with the lowest %$HANPP_0$ occurring in counties with >50% forest cover. We observed a significant ($p<0.001$) but weak, negative relationship between %$HANPP_0$ and county means of landscape diversity ($r=-0.53$, $r^2=0.28$) and a significant ($p<0.001$), moderate, negative relationship between %$HANPP_0$ and local connectedness ($r=-0.61$, $r^2=0.36$). Our findings are comparable to global estimate of HANPP on croplands and forestlands, and support previous research indicating HANPP negatively impacts biodiversity. We concluded the calculation of HANPP could be used as an additional tool for conservation professionals during regional-scale landuse planning or conservation decision-making, particularly in mixed-use landscapes that exhibit potential to support biodiversity based on abiotic proxy measures and have high amounts of primary production harvest.

Acknowledgements

This thesis is the creation of many individuals who all lent their time and expertise to its completion. I would like to thank my advisors, Dr. Bill Currie and Dr. Doug Pearsall, for all their help through this process. They both helped encourage me throughout the course of my work and gave guidance that lead to the development of many interesting ideas and improved my final thesis immeasurably. Dr. Pearsall in particular lent his knowledge as a conservation professional working outside of academia. This point of view was essential to developing this project.

I would also like to thank Preeti Rao, whose work with Dr. Currie gave me a template to work from, and whose patient teaching helped me figure out how to use ArcGIS for this project. She also lent aid in using Google Earth Engine and creating shapefiles that moved between different technologies. Shannon Brines also contributed to my ability to use ArcGIS to complete this project, and I would like to acknowledge the help he gave me both in his class and during office hours.

Finally, I would also like to thank my father for lending his expertise in R. No matter how many times I use it, having a second set of eyes to look at lines of code is always helpful. My family and roommates also helped me immeasurably, patiently listening to me talk about nothing but my thesis for the last twelve months, and funding from the Wyss Scholars Program for Conservation of the American West contributed to my ability to complete my degree program.

ABSTRACT I

ACKNOWLEDGEMENTS II

CHAPTER 1 1

Introduction 1

Ecosystem Energy and HANPP in Relation to Biodiversity 2

Abiotic Metrics as Proxies for Biodiversity 3

Study Objectives 4

Methods 5

Study Region–The U.S. Great Lakes Socioecological Gradient 5

Definition of HANPP 7

Spatial Unit of Analysis 8

Data Aggregation and Synthesis 9

HANPP Calculations 11

Data Analysis 12

Results 13

Spatial distribution of %$HANPP_0$ 13

Relationship Between %$HANPP_0$ and Biodiversity Metrics 14

Identifying Counties with Greatest Potential for Biodiversity Conservation 15

Discussion 18

Quantification of HANPP on Managed Lands in the US Great Lakes Region 18

Relationship between HANPP and biodiversity metrics 20

HANPP as a tool for conservation decision-making 22

Conclusion 25

CHAPTER 2: ADDITIONAL RESULTS AND ANALYSIS 27

The Relationship Between %HANPP0 and Landuse/Landcover 27

Crop Analysis 27

Broad Analysis of the Relation Between %$HANPP_0$ and Forestlands 29

Analysis of HANPP Due to Landuse Change 32

Methodological Challenges 33

Review of Different Crop to NPP Conversion Variables 33

Regional HANPP Calculation 39

WORKS CITED 41

APPENDIX A: CROP DATA DEVELOPMENT 49

APPENDIX B: FOREST DATA DEVELOPMENT 57

APPENDIX C: SPATIAL DATA DEVELOPMENT 60

Potential NPP (NPP0) 60

Actual NPP (NPPact) 60

Detailed GIS Methods 60

APPENDIX D: COUNTIES OF CONSERVATION INTEREST 66

Chapter 1

Introduction

Humans have become the dominant influence on Earth's systems, modifying land cover and habitat, altering the global climate, and driving global biodiversity loss (DeFries et al. 2004; Pimm & Raven, 2000; Vitousek et al. 1997). To accommodate the resource needs of the growing human population while also accommodating the resource needs of other species, conservation decision-makers need a deeper understanding of the effects of human activities on ecological conditions for other species. Many studies address the effects of land use on habitat quality, but fewer address the question of how human activities impact ecosystem energy dynamics.

Human appropriation of net primary productivity (HANPP) is a measure of human pressures on biodiversity (Haberl et al. 2014, 2012, 2009, 2004) because it represents the proportion of energy flow that was historically available to wildlife food webs but has been appropriated for human use, primarily through the harvesting of primary production. Close to 29% of global aboveground potential NPP (NPP_0, here defined as the NPP of the potential natural vegetation of a landscape) was human appropriated at the turn of the twentieth century (Haberl et al., 2007). As human populations and needs continue to grow, there is considerable potential for human alteration of ecosystem energy dynamics to impact species. This is particularly true in cropland-dominated landscapes, which are responsible for ca. 50% of global HANPP and appropriate up to 85% of NPP_0 (Haberl et al. 2014). Calculating spatial patterns of HANPP across differing socioecological landscape, such as those that support various cropping and forestry systems, could improve our understanding of interactions in human-environment systems at the landscape scale.

In the present study, we strive to understand the relationship of HANPP to selected conservation metrics and landuses and to analyze patterns of HANPP across a region of conservation interest: a portion of the Great Lakes region of the Upper Midwest, USA (see below). Identifying “high conservation value regions” is of key interest to conservation professionals. A recent study by Anderson et al. (2018) in collaboration with The Nature Conservancy (TNC) identified the spatial distribution of climate resilient sites in the Great Lakes and Tallgrass Prairie regions at a 30m scale resolution. The study defined site resilience as “the capacity of a site to maintain biological diversity, productivity and ecological function as the climate changes” (Anderson et al. 2018). Sites that score higher on the site resilience index are more likely to retain biodiversity going forward. Site resilience is a relatively new and important parameter to consider in assessing conservation value. The site resilience index used by Anderson et al. (2018) integrated two variables: landscape diversity and local connectedness (each described below). Each of these variables was used by Anderson et al. (2018) as abiotic proxies for biodiversity; we obtained their spatial results and used these data in our analysis (hereafter we refer to these two proxies as “biodiversity metrics”).

Ecosystem Energy and HANPP in Relation to Biodiversity

NPP and HANPP are typically quantified in terms of biomass dry weight, but conceptually they represent flows of energy (Currie 2012, Haberl et al. 2014). The flow of energy in ecosystem food webs has been identified as a causal factor controlling species richness (the “species-energy hypothesis”; Hawkins, Porter, & Diniz-Filho, 2003; Mittelbach et al. 2001; Wright, 1983). Spatial variability in the total amount of energy that remains available to ecological food webs after human extraction of NPP may help to explain spatial patterns of biodiversity. Previous studies have found an overall negative relationship between HANPP and biodiversity (Haberl et al. 2012, 2009, 2004; Vačkář et al. 2016).

Abiotic Metrics as Proxies for Biodiversity

Biodiversity metrics are useful because spatial patterns of biodiversity and species richness are often unavailable at regional or larger scales. Few studies have directly examined the relationship between HANPP and species richness for this reason (Haberl et al. 2014). Our study sought to develop and assess HANPP as an additional biodiversity metric for conservation professionals by comparing the spatial distribution of established biodiversity metrics to the spatial distribution of HANPP across our study region (see below).

Anderson and Ferree (2010) provided evidence that regional biodiversity correlates strongly with geophysical settings, including the number of geological classes, latitude, elevation range, dominant vegetation, and the amount of calcareous bedrock. Multiple studies have noted that different forest types and vegetation occur on different soil and topographic types (Abrams, 1992; Host et al. 1987). Landscape diversity was defined in the study by Anderson et al (2018) as an estimate of "the number of microclimates available within a given area. It is measured by counting the variety of landforms, and the density and connectivity of wetlands." A number of studies have used landscape diversity or related measures as an indicator of regional capacity to support biodiversity (Anderson et al. 2016, 2018; Anderson & Ferree, 2010; Lapin and Barnes, 1995; Lawler et al. 2015; Stein, Gerstner, & Kreft, 2014).

Landscape permeability, a variable that draws on fragmentation and connectivity, is likewise associated with biodiversity. It has been used as an indicator of how well habitats can sustain species over the long term (Anderson et al. 2016, 2018). The more permeable a landscape, the more species can move through it and adapt to changed circumstances and maintain gene flows among sub-populations (Fahrig, 2003; Lindenmayer & Fischer, 2006). Unlike landscape diversity, which may be largely independent of human activity when defined based on physiographic variables, permeability is a landscape variable driven by socioecological processes and human land-use (Lawler et al. 2015). Roads, deforestation, urban and suburban build-up–all can create barriers to the movement of species and essential ecological flows. The degree to which these landscape features retard the movements and migrations of wildlife is captured in the concept of local connectedness, defined by Anderson et al (2018) as, "the number of barriers and the degree of fragmentation within the same area." Different landcovers were assigned different resistance scores, with "Developed, High Intensity," having the highest score (20) and natural lands (e.g. forests, wetlands, and natural grasslands) having the lowest (1). Cropland was assigned a score (7) just below that of "Developed, Low Intensity" (Table 3.3 in Anderson et al., 2018).

Study Objectives

Here we examine the spatial relationship between established biodiversity metrics and the HANPP of the dominant terrestrial landuse across a range of intensities, i.e. forestlands and croplands, across our study region. Our purpose is to improve the understanding of these landscape to regional-scale metrics for use in decision-making for biodiversity conservation. Our first objective was to quantify patterns of HANPP in forestland and cropland, comparing the extraction of NPP in a recent decade against the potential natural vegetation that largely existed in the region prior to European settlement. Our second objective was to assess spatial patterns of HANPP in comparison to landscape diversity and local connectedness (as provided by Anderson et al. 2018) at the county scale across the region. Together these two objectives both expand the body of research on distributions of HANPP across different regional landscapes and begin to develop HANPP as a working metric that moves beyond academic discussions.

Methods

Study Region–The U.S. Great Lakes Socioecological Gradient

Our study focuses on part of the Great Lakes region of the Upper Midwest, USA. The Great Lakes contain nearly 21 percent of global and 84 percent of the US surface fresh water (US EPA, 2015). The US portion of the Great Lakes basin contains approximately 10 percent of the US population and is responsible for seven percent of US crop production (US EPA, 2015). The area that we consider (Fig. 1) includes a majority of the US side of the hydrologic basin of the Great Lakes, including all of Michigan and portions of Wisconsin, Ohio, and Indiana. Two ecoregional provinces dominate this area: the Laurentian Mixed Forest province in the north and the Eastern Broadleaf Forest (Continental) province in the south, with a few counties falling within the Prairie Parkland (Temperate) province and the Eastern Broadleaf Forest (Oceanic) province (Bailey, 1994).

This region's heterogeneous landuses and landcover (LU/LC) make it an ideal location to study socioecological system dynamics at the landscape scale. Crop production is one of the most important economic drivers in the Great Lakes today, bringing in more than $15 billion annually in cash receipts to the lake-border states of Michigan, Wisconsin, and Minnesota (Sousounis & Bisanz, 2000). In our study region the southernmost areas make up the northern edge of the US cornbelt and field crops like corn and soybeans dominate, while in the mid-latitude and northern parts crops trend more toward vegetables, fruits (Sousounis & Bisanz, 2000), and hay (Han et al. 2012).

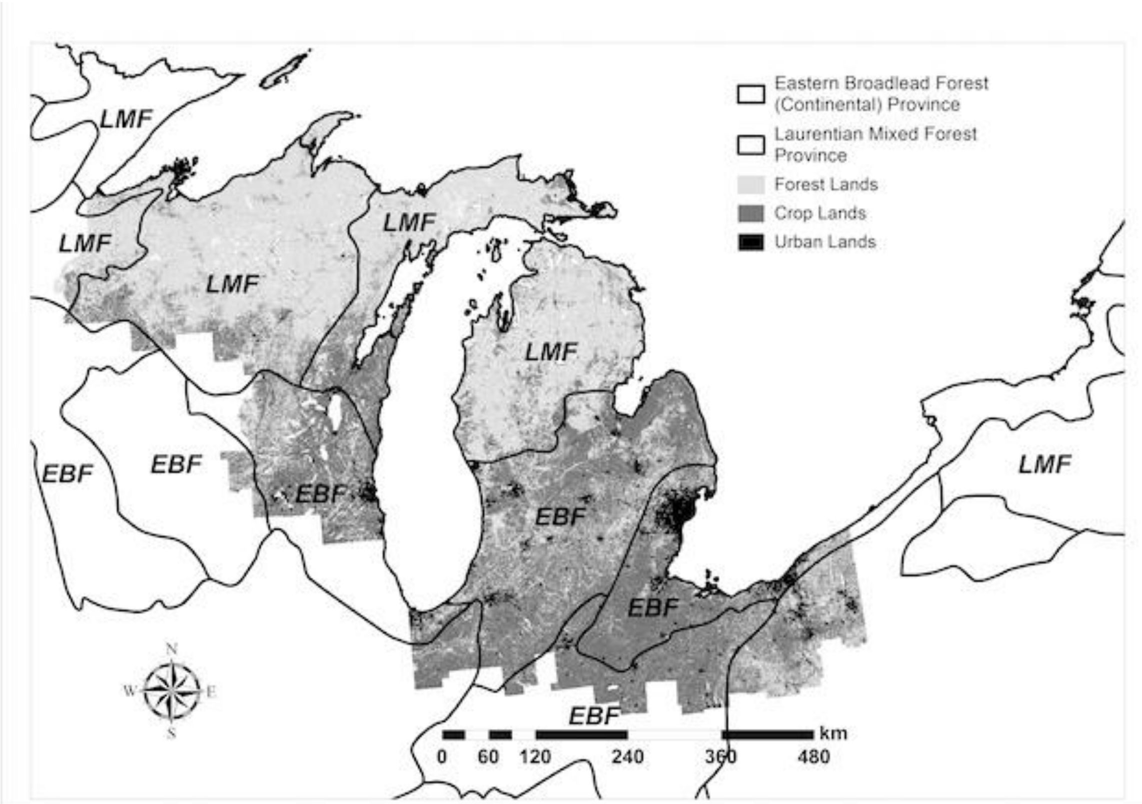


Figure 1: Landcover map of the study region showing forest, crop, and urban lands, along with the region's dominant ecoregional provinces–the Eastern Broadleaf Forest (Continental; EBFC) and the Laurentian Mixed Forest (LMF). Approximately 38% of the study region is cropland and 46% is forestland. Landcover data was retrieved from https://www.mrlc.gov/nlcd11_data.php and developed by Homer et al. (2015). Ecoregional data was retrieved from https://www.fs.fed.us/rm/ecoregions/products/map-ecoregions-united-states/ and developed by Bailey (1994).

The northern sectors of the study region are heavily forested with mixed coniferous and hardwood forests, shifting to boreal ecotones in Michigan's Upper Peninsula (Bogue, 2000). In these areas there is farming, particularly hay (Han et al. 2012), but cropland is eclipsed by forestland. Across Michigan, the forest products industry is worth $20 billion, is responsible for 26,000 jobs, and removes approximately 20 percent of its raw materials from state forestlands (The Michigan Department of Natural Resources, 2018).

Regional demographic and economic change (Brown, 2003; Robinson, 2012; Theobald, 2005) and growing interest in biofuels in the region may drive future landuse decisions and increase landuse intensity on both forestlands and croplands (Gustafson &

Loehle, 2008; Slater, Keegstra, & Donohue, 2010; Kells & Swinton, 2014). Additionally, climate change could drive shifts in landuse and forest make-up (Breffle et al. 2013; Handler et al. 2014). Together these trends could impact how much biomass is extracted from the region and where that extraction takes place, i.e. regional landuse patterns.

Changes in the nature and location of forest and croplands, and the intensity of biomass extraction on them, could affect the region's ability to provide supporting ecological services. The region has undergone a transition over the last 200 years from largely unmanaged forests and small amounts of cropland (i.e. during management by Native American tribal groups) to extremely high amounts of timber extraction in the north during the turn of the 19th century and growing domination of large-scale cropland in the south (Bogue, 2000; Handler et al. 2014). The modern landscape comprises heterogeneous LU/LC types and varied ownership patterns–managed, fragmented forests with altered species compositions coupled with high amounts of cropland throughout much of the mid and southern regions of the Great Lakes (Handler et al. 2014; Whitney, 1987). The choices inherent in this history, such as how much timber to harvest, where to plant crops, or what types of crops to plant, have shaped the present Great Lakes socioecological system (Steen-Adams et al. 2015). Creating future system trajectories that support biodiversity requires conservation professionals to balance human needs with the needs of other species; in this pursuit, multiple landuse planning tools that complement each other and illuminate different aspects of human-environment interaction are a necessity.

Definition of HANPP

For this paper, we adopt a widely-used set of terms related to HANPP (Haberl, 1997; Haberl et al. 2001, 2007; Haberl, Erb, & Krausmann, 2014). HANPP is defined as "the combined effect of harvest and productivity changes induced by land use on the availability of NPP in ecosystems" (Haberl et al. 2007). In other words, this is a somewhat complicated metric to define operationally because it arises from two factors: changes in NPP from human landuse compared to the potential natural vegetation ($HANPP_{luc}$), together with extraction of NPP by human harvest (NPP_h) (eqn. 1).

$$HANPP = HANPP_{luc} + NPP_h \tag{1}$$

$$HANPP = NPP_0 - (NPP_{act} - NPP_h) \tag{2}$$

$$HANPP_{luc} = NPP_0 - NPP_{act} \tag{3}$$

Combining equations (1) and (2) shows that under this set of definitions, $HANPP_{luc}$ can be calculated from potential natural NPP (NPP_0) and actual NPP (NPP_{act}) for the unit of the landscape (eqn. 3). The definitions also address the fact that timber harvests do not remove the entirety of the forest with every harvest by calculating NPP_h of forestlands as a ratio of total forest inventory (Haberl et al. 2001, 2004). We consider both the above and below-ground compartments of NPP and focuses on the percent of NPP_0 being appropriated (%$HANPP_0$). We do not include removals of NPP (NPP_h) due to human-caused fires or livestock.

Spatial Unit of Analysis

We rescaled all spatial data to a 500m pixel resolution and reprojected the data into NAD83 Conus Albers. This projection minimizes spatial distortion within our study region (see Appendix C for more information on spatial data transformations). We use counties as our spatial unit of analysis (n=188 counties) because forest and crop harvest data from the US Forest Service Forest Inventory and Analysis (FIA; Burrill, 2018) and the US Department of Agriculture (USDA; "USDA/NASS QuickStats Ad-hoc Query Tool," 2007, 2012) are aggregated to the county scale. Attempting to use the data at a finer scale introduces high levels of uncertainty (S. Pugh, personal communication, 2017). Additionally, using a scale based on counties as socio-political boundaries will interface better with policy-based planning and studies of demographic change. To compliment the county analysis, we also stratified our data by ecoregions as defined by the US Forest Service (Bailey, 1994). This allowed us to separate out data between counties with different dominant LU/LC patterns within regions where the climate (perhaps most importantly growing season length) and soil type are relatively similar.

Data Aggregation and Synthesis

All spatial analyses were performed using ArcGIS version 10.5.1 (*ArcGIS ArcMap*, 2017), and all data manipulations and statistical analyses were performed using R and Excel. NPP units were transformed into kg C m^{-2} y^{-1} for calculations and final results. In ArcGIS, the zonal statistics function was used to aggregate all values to a county level mean, at which point values were joined to county shapefiles (Fan, 2018).

For NPP_0 (eqn. 2), we used results from Haberl et al (2007) which were calculated at 5 arc min resolution (about 10 km pixel resolution). The researchers derived NPP_0 using the Lund-Potsdam-Jenna Dynamic Global Vegetation Model (LPJ DGVM; Gerten et al. 2004; Sitch et al. 2003) results for a 5-year average over 1998 to 2002 (Haberl et al. 2007). We reprojected and rescaled the data and used zonal statistics to produce a table with the mean NPP_0 in g C m^{-2} y^{-1} of each county in our study region, which we then transformed into kg C m^{-2} y^{-1}.

Data for NPP_{act} (eqn. 2) was obtained from the MODIS Net Primary Productivity MOD17A3H V6 product (Running et al. 2015; Appendix C) using Google Earth Engine. We averaged the MODIS data from 2005 to 2015, to help account for stochastic uncertainty.

We calculated NPP_h (eqn. 2) of croplands based on production and yield data primarily obtained from the USDA Agricultural Census ("USDA/NASS QuickStats Ad-hoc Query Tool," 2007, 2012). This data was input to the equations suggested by Hicke et al (2004; eqn. 4 & 5) to transform field crop production (eqn. 4) into field crop NPP values in kg C m^{-2} yr^{-1}:

$$P = \sum_i \frac{PC_i \times MRY_i \times (1 - MC_i) \times C}{HI_i \times f_{AG,i}} \quad (4)$$

$$NPP = \frac{P}{\sum_i A_i} \quad (5)$$

where i indicates different crop types, PC indicates the production of a crop in reported units (e.g. bushels), P is production in g C yr^{-1}, and A is crop area. We obtained the other input values for the equation–harvest index (HI), fraction of above ground productivity (f_{AG}), moisture content (MC), and percent carbon (C) per unit dry mass–from data compiled by Lobell et al (2002) and Prince et al (2001; Appendix A, Table A-1). For fruit and vegetable crops, we used the equations and parameters presented in Monfreda et al (2008; eqn. 6 & 7; Appendix A, Table A-2), where NPP_i represents the NPP of each crop i, EY represents estimated yield, DF is the dry fraction (1-moisture content), and RS is the root:shoot ratio.

$$NPP_i = \frac{EY_i \times DF_i \times C}{HI_i \times RS_i} \quad (6)$$

$$NPP = \sum_{i=1}^{n} \left(\frac{NPP_i}{fcrop_i}\right) \quad (7)$$

For the forest data, we downloaded data representing volume of live trees harvested from forestlands in ft^3 $acre^{-1}$ from the FIA EVALIDator program for the years 2005-2015 (Burrill, 2018; Appendix B). The use of ratio data accounted for the fact that not all forest is harvested every year. We transformed all NPP_h values for forests and crops into kg C m^{-2} and calculated an area-weighted aggregate value of combined forest and crop NPP_h values by county.

HANPP Calculations

For forestland and cropland separately within each county, we calculated NPP_h at county-scale resolution across our study region (Eqns. 1-3). Investigators often express HANPP as a percentage of NPP_0, the NPP of potential natural vegetation in a unit of the landscape (we write this percentage as $\%HANPP_0$). We calculated $\%HANPP_0$ for forests and croplands separately within each county as well as area-weighted $\%HANPP_0$, combining forests and croplands within each county (eqn. 8, Table 1):

$$\%HANPP_0 = 100 * \frac{(NPP_{hFor} * Area_{For} + NPP_{hAg} * Area_{Ag})}{NPP_0 * Area_{Total}} \tag{8}$$

NPP_{hFor} is the harvested NPP of forestlands per unit area, $Area_{For}$ is the area of forestlands, NPP_{hAg} is the harvested NPP of croplands per unit area, $Area_{Ag}$ is the area of croplands, and $Area_{Total}$ is the total area of managed forest plus crop lands in each county.

Table 1: Summary statistics of area-weighted $\%HANPP_0$ and $\%HANPP_0$ separated into forest and croplands.

	Range	***Mean***	***Median***	***Mode***	***SD***
%HANPP₀	3.2–151	47	48	76	32
%HANPP₀₀ of forestlands	0.049 – 17	4.7	4.2	12	3.0
%HANPP₀ of croplands	21–195	80	76	80	23

We created spatial representations of the $\%HANPP_0$ distribution across our study region by importing the county-scale results into ArcGIS and joining them to county shapefiles.

Data Analysis

To analyze the relationship between $\%HANPP_0$ and landscape diversity and local connectedness, we downloaded and used the spatial data produced by Anderson et al (2018). We analyzed linear regressions between our HANPP results and the county mean values of the two biodiversity metrics, both across the entire study region and stratified by ecoregion. We identified four outlier counties in each relationship: Lake and Cuyahoga Counties in Ohio, Milwaukee County in Wisconsin, and Wayne County, Michigan. All of these counties contain major urban centers and thus exhibited outlier behavior in the relationships between HANPP and biodiversity metrics. Our analysis focuses on forest and croplands, so we chose to remove these four counties from our analysis.

Additionally, we identified counties that combined high potential biodiversity–those with high levels of mean local connectedness or mean landscape diversity–with low intensity of human use-intensity as measured by $\%HANPP_0$. We did so by identifying the 25th, 50th, and 75th percentiles of both $\%HANPP_0$ and the two biodiversity metrics and defining groups based on these statistics. We performed pairwise comparisons of different combinations of $\%HANPP_0$ and either connectedness or landscape diversity that were >50th percentile or <50th percentile using Wilcoxon rank sum test to examine the significance of the differences among groups (Table 2).

Table 2: Results of the pairwise comparisons using Wilcoxon rank sum test examining the differences among groups of different combinations of %HANPP$_0$ and either connectedness or landscape diversity. Each group is a combination of two variables either above or below the 50th percentile. C1/D1 contain counties with high potential for effective biodiversity conservation, as they are low-extraction, high-diversity (LEHD) and low-extraction, high-connectedness (LEHC). The C2/D2 group has high biodiversity potential due to scores on the indices above the 50th percentile, but also has high %HANPP$_0$ values. These are high-extraction, high-connectedness/diversity (HEHC and HEHD) areas where there is a potential for biodiversity-supporting habitat but also extractive activities going on. These are counties where conservation might be costly, but valuable, depending on the cause of the high %HANPP$_0$ values and the socioeconomic drivers impacting landowner decision-making in the region. The C3/D3 and C4/D4 counties are both lower in priority for conservation for existing local biodiversity, as they score an average below the 50th percentile of each of the biodiversity metrics. The C1/D1 groups have the lowest average population estimate and the lowest average road density, as well as the highest average %forest cover. This fits with the other findings of this analysis, which indicate that highly forested, low-use areas are have the highest biodiversity potential.

Group Category	*%HANPP$_0$ percentile*	*Indicator*	*Indicator Percentile*	*Mean 2010 Population Estimate*	*Significant Difference (p≤.05)*	*Mean Road Density Estimate (m/m^2)*	*Significant Difference (p≤.05)*	*Mean %Forest Cover*	*Significant Difference (p≤.05)*
C1	*≤50th*	*connectedness*	*≥50th*	*41411.25*	*C2, C3, C4*	*0.001761446*	*C2, C3, C4*	*67.63474*	*C2, C3, C4*
C2	*≥50th*	*connectedness*	*≥50th*	*57270.48*	*C1, C3*	*0.002157511*	*C1, C3*	*25.11274*	*C1, C3, C4*
C3	*≤50th*	*connectedness*	*≤50th*	*253036.4*	*C1, C2*	*0.002906438*	*C1, C2*	*40.4669*	*C1, C4*
C4	*≥50th*	*connectedness*	*≤50th*	*97638.01*	*C1*	*0.002617986*	*C1*	*14.1498*	*C1, C3*
D1	*≤50th*	*landscape diversity*	*≥50th*	*41280.03*	*D2, D3, D4*	*0.001698855*	*D2, D3, D4*	*67.77231*	*D2, D3, D4*
D2	*≥50th*	*landscape diversity*	*≥50th*	*54110.58*	*D1*	*0.00216339*	*D3, D4*	*23.52386*	*D3, D4*
D3	*≤50th*	*landscape diversity*	*≤50th*	*212672.4*	*D1*	*0.002833123*	*D1, D2*	*45.34228*	*D4*
D4	*≥50th*	*landscape diversity*	*≤50th*	*102211.2*	*D1*	*0.002651708*	*D2*	*13.9452*	*D1, D2, D3*

Results

Spatial distribution of %HANPP$_0$

In our results, forestlands accounted for an average of 4.7 percent of appropriated NPP$_0$, while croplands accounted for an average of 80 percent of NPP$_0$; the overall mean area-weighted %HANPP$_0$ across our region was 45 percent. The highest %HANPP$_0$ values were in the southern counties of our study region in Ohio–the north end of the U.S. corn belt–as well as the fertile regions of southeastern Wisconsin and counties adjacent to Saginaw Bay in Michigan (Fig. 2). These are all counties with extensive and highly productive croplands. An east-west corridor in southern Michigan had lower %HANPP$_0$, as did the northern portion of the Michigan Lower Peninsula and the entirety of the Michigan

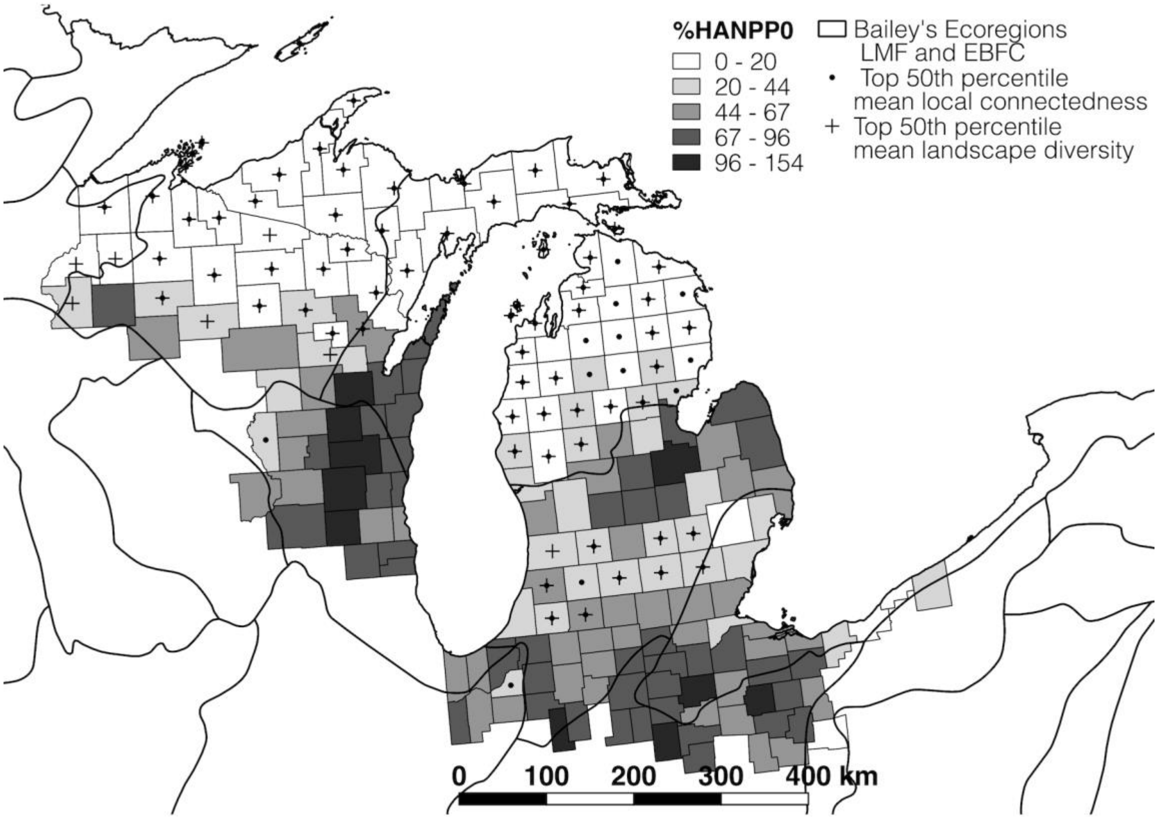


Figure 2: Map showing the spatial distribution of $\%HANPP_0$ in relation to counties with low-extraction, high-connectedness/diversity (LEHC and LEHD). The two dominant ecoregional provinces are shown, with the LMF province covering the northern portions of Michigan and Wisconsin and the EBFC covering the southern portions of these states and northern Ohio and Indiana. Most of the LEHC/LEHD counties are in the LMF province and coincide with area-weighted $\%HANPP_0$ between 3.2 and 44%. The exception is a band in southern Michigan, which coincides with a band of mixed LU/LC, including multiple cities (Detroit, Ann Arbor, Lansing, Grand Rapids, Kalamazoo, Flint, and Jackson, MI; and Ekhart, OH) and their associated suburban and exurban fringes (Fig. 1). These counties have an area-weighted $\%HANPP_0 \leq 44\%$ but are in the EBFC province.

Upper Peninsula (Fig. 2) The east-west corridor in southern Michigan corresponds to a band of urban areas and their associated exurban fringes, while the northern, low $\%HANPP_0$ areas corresponded to regions of dense forest cover (Fig. 1). The lowest associated $\%HANPP_0$ occurs in counties with >50% forest cover and >0 mean connectedness (Chapter 2, Fig. 6).

Relationship Between $\%HANPP_0$ and Biodiversity Metrics

A strong overall pattern in our results was that both landscape diversity and local connectedness exhibited lower values in counties that are experiencing high NPP extraction as measured by $\%HANPP_0$ (Fig. 2; Table 3). This pattern is stronger between local connectedness and %HANPP$_0$ ($r = -0.61$, $r^2 = 0.36$, $p<0.001$), than between %HANPP$_0$ and landscape diversity, particularly in the LMF ecoregional province where 51% of the variation in mean connectedness is explained by %HANPP$_0$ ($r^2 = 0.51$; Table 3). Forestland is more abundant than cropland in the LMF counties, and lower road densities and population (Appendix D) present fewer opportunities for both forest fragmentation and large-scale resource extraction.

The relationship between %HANPP$_0$ and landscape diversity is weak (r^2=0.28) but highly significant ($p<0.001$). R^2 is not improved by stratification by ecoregion, but the LMF ecoregional province again show higher r^2 than the EBFC ecoregional province.

Table 3: Linear regression results examining the relationship between %HANPP$_0$ and the biodiversity metrics. Regressions were done for the whole study region and for the two main ecoregions, Laurentian Mixed Forest (LMF) and Eastern Broadleaf Forest (Continental) (EBFC). All relationships were stronger in the LMF province than in the EBFC province.

REGRESSION		P-VALUE	R	R^2
%HANPP$_0$ vs. MEAN LOCAL CONNECTEDNESS	Region overall	< 2.2e-16	-0.61	0.36
	LMF	*3.86E-13*	*-0.72*	*0.51*
	EBFC	*0.007692*	*-0.26*	*0.06*
%HANPP$_0$ vs. MEAN LANDSCAPE DIVERSITY	Region overall	6.20E-15	-0.53	0.28
	LMF	*1.61E-06*	*-0.52*	*0.26*
	EBFC	*7.81E-02*	*-0.18*	*0.02*

Identifying Counties with Greatest Potential for Biodiversity Conservation

We identified counties that fell within the bottom 50th percentile of %HANPP$_0$ and in the top 50th percentiles of landscape diversity or connectedness (Table 2, Figs. 3 and 4). We refer to these groups as low-extraction, high-diversity (LEHD) and low-extraction, high-connectedness (LEHC) respectively. Most of the LEHD/C counties are located in the northern portion of Michigan’s Lower Peninsula, Michigan’s Upper Peninsula, and northern

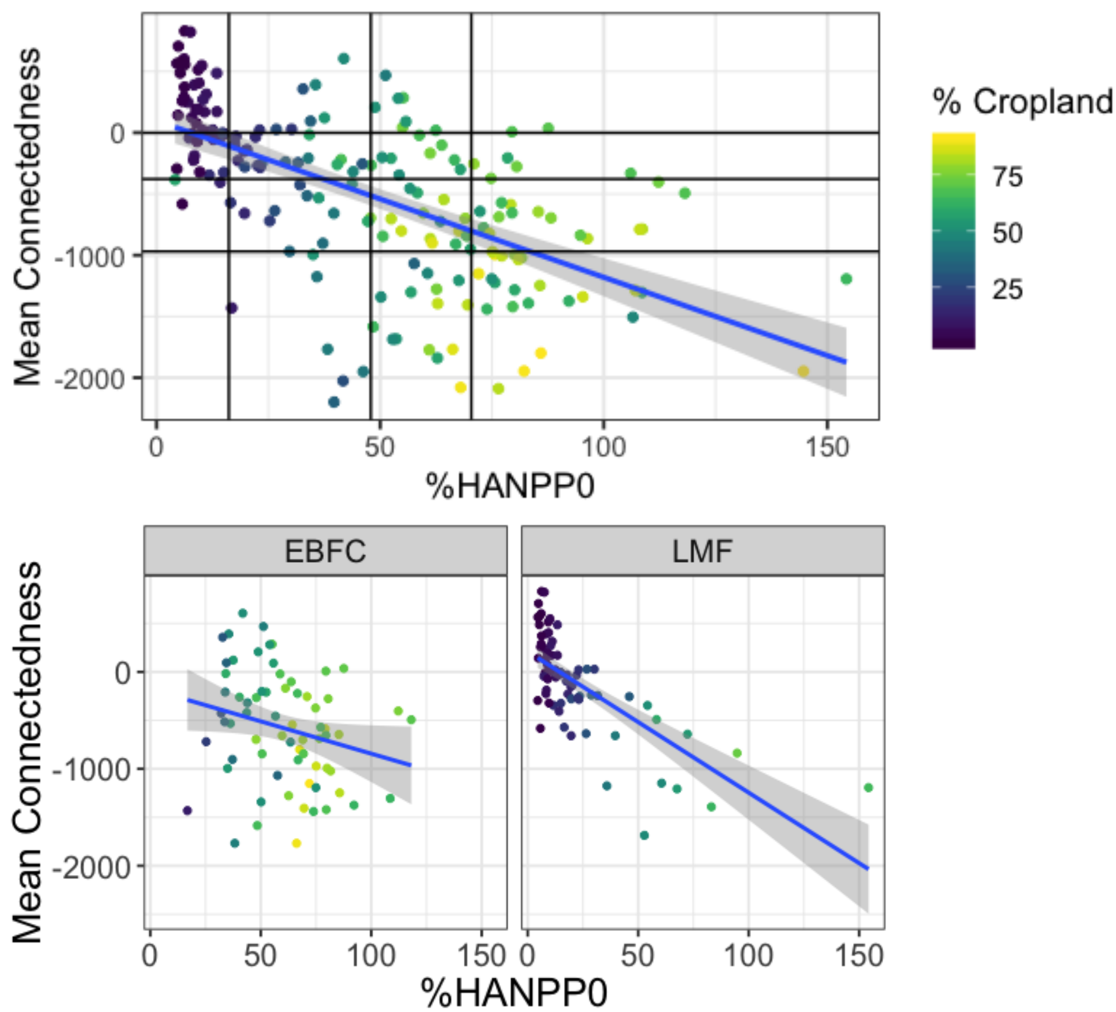


Figure 3: The relationship between mean local connectedness and weighted %HANPP0", with 25th, 50th, and 75th percentile lines shown on the top graph for the whole study region and the counties stratified by ecoregion in the bottom two graphs. Each point represents a singel county in our study regin and the grey area around the line of best fit it the 95% confidence interval. We found that for the whole study region, the relationship between mean connectedness and %$HANPP_0$ is moderate (r^2=0.36) and significant ($p<0.001$). The relationship is much stronger in the LMF province, with 51% of the variation in mean connectedness explained by %$HANPP_0$. In contrast, r^2 is only .06 in the EBFC province. Counties within the bottom 50^{th} percentile of %$HANPP_0$ and in the top 50^{th} percentiles of connectedness are low-extraction, high-connectedness (LEHC) counties. These counties are largely forested (≥50%) and in the LMF province.

Wisconsin. They have >50% forested landcover and ≤10% crop landcover (Appendix D, Tables D-1 and D-2). Hay for forage or pasture is the crop that is planted over the most area across these counties.

Those falling in the opposite arrangement–high-extraction, low-connectedness (HELC) and high-extraction, low-diversity (HELD)–were categorized as high risk and likely high cost for biodiversity conservation due to the combined lack of biodiversity-supporting habitat and high intensity of resource extraction.

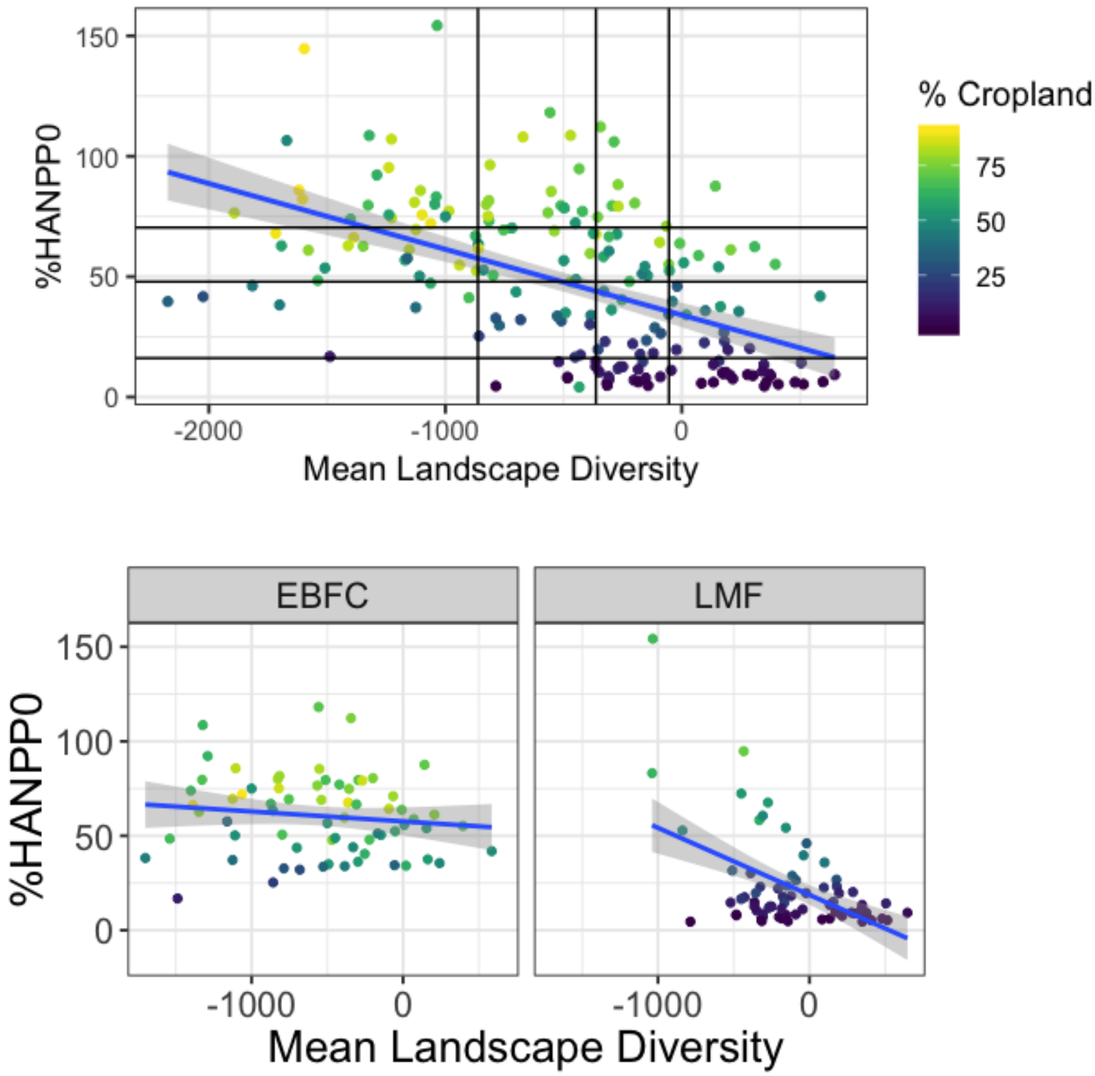


Figure 4: The relationship between mean landscape diversity and $\%HANPP_0$, with the 25^{th}, 50^{th}, and 75^{th} percentiles shown for the whole study region in the top graph and the counties stratified by ecoregion in the bottom two graphs. Each point represents a single county in our study region and the grey border around the line of best fit represents the 95% confidence interval. For the whole region (top graph), the relationship between $\%HANPP_0$ and landscape diversity is weak (r^2=0.28) but significant (p=6.0E-15). R^2 is not improved by stratification by ecoregion (bottom graphs), but the LMF province shows a higher r^2 than the EBFC province. Counties within the bottom 50^{th} percentile of $\%HANPP_0$ and in the top 50^{th} percentiles of connectedness/diversity are low-extraction, high-diversity (LEHD) counties. These counties are largely forested (≥50%) and in the LMF province.

Counties falling in the top 50th percentile of %HANPP$_0$ and in the top 50th percentile of mean local connectedness or mean landscape diversity (HEHC and HEHD respectively) were classified as high risk and high priority for biodiversity conservation. They have high potential to support biodiversity but are also being intensely used in terms of harvest of primary production.

Discussion

Quantification of HANPP on Managed Lands in the US Great Lakes Region

Most HANPP studies have been performed at a global or national scale (Haberl et al. 2007, 2014, 2009, 2004; Krausmann et al. 2013; Plutzar et al. 2016), with fewer examining the regional or local scales (Andersen et al. , 2015; Marull et al. 2016; O'Neill et al. 2007). Yet the landscape and regional scales are important in much conservation decision-making. Our analysis quantified HANPP in a region where it has not previously been examined, adding a new dataset to the body of regional and local HANPP research.

We found that %HANPP$_0$ distribution across our study region aligned well with the global means of %HANPP$_0$ in forest and crop systems, which are approximately 7% and up to 85%, respectively (Haberl et al. 2014). In our region, the mean %HANPP$_0$ of cropland was about 80% and the mean %HANPP$_0$ of forest lands was about 5%. The mean %HANPP$_0$ of forestlands in the Great Lakes region differs more from other regions than it does from the global mean. In Austria (Haberl et al 2001) and Nova Scotia (O'Neill et al. 2004)–two case studies in similarly temperate climates–aboveground %HANPP$_0$ on forestland was found to be about 25% –five times the average in our study region. This difference could be due to decreased activity in the forest products industry in our region over the last several decades (Janowiak et al. 2014; Shivan & Potter-Witter, 2011) combined with more cropland–

approximately 38% of total landcover in our study region as opposed to <10% in Nova Scotia. Although our analysis resulted in mean $\%HANPP_0$ of cropland on par with global means, the county-level data showed a pattern of high variation, ranging from 3.2% to 154% (Table 1). This large range of use-intensity indicates that not all crop-dominated landscape matrices extract high amounts of ecosystem energy. For instance, hay grown for pasture or feed dominates (in terms of area covered) in counties with $\%HANPP_0 \leq 30\%$, but it is also one of three crops, including grain corn and soybeans, that dominate counties with $\%HANPP_0 \geq 100\%$ (data not shown). Thus, the degree of NPP extracted from the landscape may depend on what types of crops are planted, in what combinations, and when and how they are grown and harvested (e.g. type of fertilizer or irrigation used, season of planting, variety of crop planted.

Increasing the percent of forestland in the landscape matrix is a possible strategy for increasing landscape-scale ecosystem energy retention. We found consistently low $\%HANPP_0$ on forestlands ($\leq 17\%$), and in counties with $\geq 30\%$ forestland the area-weighted $\%HANPP_0$ was uniformly low (<45%). This included counties outside of the forestland-dominated LMF province, most notably the east-west band of mixed LU/LC counties that are both LEHC/LEHD counties and include urban areas such as Detroit, Flint, Ann Arbor, Lansing, Kalamazoo, and Grand Rapids. Previous research has shown a correlation between exurban expansion and an increase in tree cover and gross primary productivity; exurban landscapes also display carbon storage levels higher than those in croplands (An et al. 2011; Brown et al. 2008; Currie et al. 2016). Together with our findings this research suggests retention of forestland or afforestation can increase the potential for a mixed LU/LC landscape and matrix to support biodiversity at a county scale.

Despite the notable amount of NPP left in managed forestlands around the Great Lakes compared with that of croplands, high county %HANPP$_0$ values lead croplands to more strongly influence regional mean %HANPP$_0$. Thus, increasing forestland within a mixed LU/LC matrix may not decrease regional %HANPP$_0$, although it may increase landscape patterns that benefit biodiversity. Regional mean %HANPP$_0$ may not be impacted unless croplands undergo conversion to other LU/LC types (such as large-scale crop-to-forest conversions), or crop matrices and planting/harvest techniques are purposefully chosen to increase the amount of NPP left in the ecosystem. In mixed LU/LC areas of biodiversity concern, intensive row crops (e.g. corn, soybeans, sorghum) may be replaced or intermixed with lower-intensity perennial crops, such as hay and alfalfa systems with low harvest rates (Asbjornsen et al. 2014). Graham et al. (2017), in a spatial modeling analysis of an agricultural landscape in Illinois, found that replacement of annual row crops with perennial crops was likely to benefit the biodiversity of pollinators.

Relationship between HANPP and biodiversity metrics

The fact that %HANPP$_0$ exhibited a stronger relationship with connectedness than with landscape diversity indicates that differing physiographic conditions do not affect biomass removal rates as much as biomass removal rates affect the spatial patterns of habitat connectivity and fragmentation and thus the permeability of the landscape for wildlife movement. This may be due to the relatively low degree of diversity of geophysical conditions in our study region (e.g., as compared to mountainous landscapes). The relatively low landscape diversity limits the extent to which the variable can influence how much biomass humans extract from the ecosystem. The relationship between %HANPP$_0$ and landscape diversity was much stronger in the Laurentian Mixed Forest (LMF) ecoregional province ($r^2 = 0.26$), a region that is heavily forested and has more elevation change, remaining wetlands, and diverse geology than the Eastern Broadleaf Forest (Continental; EBFC) ecoregional province ($r^2 = 0.02$; Table 2). The difference between ecoregional provinces implies that the diversity of geophysical settings may act as more of a driver of %$HANPP_0$ in regions with greater variation in landscape diversity, a supposition supported by previous research that has found topographical elements like slope, altitude, and roughness (the flatness/hilliness of a landscape) to be the most predictive of %$HANPP_0$ (Wrbka et al. 2004). However, the overall weak relationship indicates landscape diversity and %$HANPP_0$ do not communicate the same information about socioecological interactions across a landscape. Although the relationship between %$HANPP_0$ and mean local connectedness is stronger, only 36% of the variance in mean local connectedness among all counties could be explained by %$HANPP_0$. Again, the relationship was stronger in the LMF province ($r^2 = 0.51$) and almost non-existent in the EBFC province ($r^2 = 0.06$).

Wrbka et al (2004) similarly found that landform patterns–aspect, roughness, and elevation, variables related to topography–have a moderate to weak relationship with spatial patterns of HANPP, and that the relationship varies notably among geo-ecological units. The research group hypothesized the weak relationship was because their study area consisted of "cultural landscapes," in which the disturbance regime and major energy and material fluxes are controlled by humans. How this control plays out, e.g. what management strategies are used on the land, is constrained not just by the geophysical makeup of the landscape but by interacting social and economic forces. These may be more or less important than ecological constraints in determining management practices at different times and in different spaces.

Our analysis consistently showed a much stronger relationship between $\%HANPP_0$ and the biodiversity metrics in the LMF ecoregional province than in the EBFC, which contained more counties dominated by cropland and urban/exurban land, the most intensively used LU/LC types worldwide (Haberl et al. 2014). The LMF province, on the other hand, contains counties with high percent forest cover that is managed more irregularly (i.e. forest harvests occur only once every few decades or longer, large tracts of forest have protected status that limits resource extraction, and many private forest landowners choose not to harvest their forests at all; Janowiak et al. 2014; Shivan & Potter-Witter, 2011). One explanation is that the socioeconomic forces Wrbka et al (2004) predicted as a third explanatory variable may be more relevant in regions dominated by more intensive extraction of NPP, in which socioeconomic profits and losses are higher and with more immediate effects.

HANPP as a tool for conservation decision-making

Conservation professionals have a wealth of tools and variables at their disposal to aid them in evaluating where to focus conservation efforts. To date, HANPP has largely been studied as an academic metric with few examples of application to conservation planning. Our analysis indicates that there is significant variability in the spatial distribution of $\%HANPP_0$ that is not fully explained by the distribution of mean landscape diversity and that there is a similar (although lesser) variability in mean local connectedness that is not explained by $\%HANPP_0$. This supports the idea that $\%HANPP_0$ may contain additional information about landscape-scale socioecological interactions for conservation professionals when used in conjunction with other metrics of human impacts on biodiversity.

One way %$HANPP_0$ operates as a biodiversity metric is as an ecosystem stress indicator. Extensive research has been done in the Great Lakes region on developing ecosystem stress indicators; percent crop LU/LC has been identified as a major terrestrial stressor on aquatic ecosystems (Johnson et al. 2015). Given that high values of crop LU/LC result in high %$HANPP_0$ and evidence that tree biomass removal (e.g. clear cut harvesting) can negatively impact downstream water quality (Ensign & Mallin, 2001; Wang, Burns, Yanai, Briggs, & Germain, 2006), significant biomass removal from terrestrial landscapes may be a stressor on downstream aquatic ecosystems. In terrestrial ecosystems, one question for conservation professionals is how much energy extraction can occur on a landscape before it crosses a threshold of rapidly declining ecosystem services. Haberl et al (2004) found that there were negative impacts for species richness at %$HANPP_0 \geq 50\%$, which in turn impacts biodiversity conservation. We have found that the mean area-weighted %$HANPP_0$ across forest and croplands in our study region was about 46%, suggesting the region may be close to a threshold past which some species that depend on landscape-scale support will decline.

We identified counties above the 50^{th} percentile of %$HANPP_0$ as HEHC/D counties or HELC/D counties. HEHC/D counties are potentially at-risk–they have a high potential for supporting present biodiversity due to their landscape patterns, but are also being heavily used for resource extraction. These are counties where conservation might be costly, but valuable, depending on the cause of the high %$HANPP_0$ values and the socioeconomic drivers impacting landowner decision-making in the region. Both these counties and HELC/D counties are those that may need ecological restoration either to improve local biodiversity support or create corridors connecting habitats of higher conservation value (Jones et al. 2015).

Habitats of higher conservation value are more likely to be found in the counties in our study region that have an average $\%HANPP_0$ below this 50% threshold. These LEHC/D counties are largely focused in the northern, heavily forested counties (Figs.1 and 2) and may be less costly to conserve as they already have landscapes that can support biodiversity and are not the site of intense resource extraction. In addition to having the highest mean percent forest cover, this county group has significantly lower mean road density ($m\ m^{-2}$) and mean population (Table 2; Appendix D, Tables D-1 and D-2) than the other county groups. In a conservation triage situation (Gerber, 2016) where limited aid must be allocated to regions where the aid will do the most good, the habitats in these counties are ones conservation professionals may want to focus on to protect and connect.

The difference between regional mean $\%HANPP_0$ values and county-level mean $\%HANPP_0$ values invites the question of how landscape-scale extraction patterns translate into local biodiversity impacts. There is evidence that different species may be differentially impacted by land sparing–conserving large tracts of unused land and allowing for smaller areas of more intense extractive management–or land sharing–ensuring human-dominated lands are managed for extraction in an ecologically-friendly way (Gonthier et al. 2014; Kremen, 2015). Because of different responses to land management from different species, both strategies can prove useful in different contexts and complimentary in landscape-level conservation planning (Kremen, 2015). What configuration of crop and forest matrices and what threshold extraction level are best for meeting biodiversity objectives may thus depend on the particular species and ecosystem services of greatest conservation concern.

Conclusion

As a snapshot of the mean LU/LC and the accompanying landscape patterns of the US Great Lakes basin in the first 15 years of the 21st century, this analysis provides an initial quantification of the spatial patterns of HANPP in our study region and shows how HANPP can complement established biodiversity metrics. We observed a moderate, negative relationship between %HANPP$_0$ and mean landscape diversity and local connectedness and a strong pattern of high %HANPP$_0$ in cropland-dominated counties and low %HANPP$_0$ in forestland-dominated counties. These relationships support previous research suggesting that HANPP is negatively correlated with landscape characteristics that likely control species richness and support previous research putting forth HANPP as a metric of human impact on biodiversity (Haberl et al. 2012; Haberl et al., 2004). Our findings suggest HANPP has the potential to be useful to conservation professionals during regional-scale landuse planning or conservation decision-making, particularly in landscapes with a combination of high site potential for biodiversity and high resource extraction activity. Further developing HANPP as a metric may illuminate which LU/LC development should be advocated for or against in the pursuit of biodiversity conservation in managed, mixed use landscapes.

Future research could continue to improve HANPP as a metric for understanding how resource extraction impacts conservation goals. To maintain current levels of site resiliency, further understanding is needed of socioecological processes and landowner decision-making, and how they interact with ecosystems to create specific matrices of LU/LC and use-intensity. Additionally, the forests of the Great Lakes region are managed by a combination of private and public interests; taking a closer look at how different crop and forest management styles impact biomass extraction levels at local scales is an important next step in regional HANPP analysis to support biodiversity conservation.

Chapter 2: Additional Results and Analysis

The Relationship Between %HANPP0 and Landuse/Landcover

Crop Analysis

To supplement our understanding of how %$HANPP_0$ differed between forest and croplands, we conducted a brief exploratory assessment of how %$HANPP_0$ related to the types of crops being grown in a county. Out of 21 crops, we identified five as covering the most land in all counties (n=188): cherries, corn grown for grain, hay, soybeans, and wheat. Of these, cherries were the most extensive crop in only one county (Leelanau Co., Michigan); likewise, wheat was most extensive in only one county (Huron Co., Michigan). In all other counties, hay, soybeans, or grain corn were the most extensive crops (Fig. 4). Soybeans and grain corn are both cash crops, typically grown in large monocultures and with high technological inputs like fertilizer and pesticides. Hay is typically grown for pasture or for animal feed.

Most of the hay, and cherries in the single county that grows cherries over a large area, are grown in the Laurentian Mixed Forest (LMF) ecoregional province. All but one county growing large areas of grain corn are located in the Eastern Broadleaf Forest (Continental; EBFC) province, with the single county outside that province located in the Prairie Parkland (Temperate; PPT) province. One county growing predominantly hay is in the Eastern Broadleaf Forest (Oceanic; EBFO) province, and both soybeans and wheat are grown solely in the EBFC province.

The ten crops with the highest NPP_h value throughout all counties were: apples, grain corn, corn for silage, sweet corn, hay, alfalfa, oats, potatoes, sorghum, and sugar beets. Hay, alfalfa, apples, silage corn, and oats all had $n \leq 5$; the majority of counties grew grain corn, potatoes, sorghum, or sugar beets (n=10) as their most energy intensive crop. Apples, hay, sweet corn, and oats are the most energy intensive crops only in counties in the LMF province. The other crops are split between the LMF and EBFC provinces, with one county in the PPT and one county in the EBFO growing grain corn as their most intensive crop (Fig. 5). For all crops where n>1, there is high variation in the %HANPP$_0$ associated with county in which the crop is growing and multiple outlier counties (Fig. 5). Stratification by ecoregion shows that counties in the LMF province grow crops side-by-side with high forest cover; counties in the EBFC province tend to have low percent forest cover.

The Kruskal-Wallis test showed there is a significant difference among crop types, both in terms of crops that cover the most area per county (p-value = 5.79e-16) and those that have the highest NPP$_h$ per county (p-value = 7.53e-09). The post-hoc test we used, the pairwise comparisons using Wilcoxon rank sum test, indicates that in terms of area there is a significant difference in %HANPP$_0$ between counties growing grain corn and counties growing hay (p-value = 3.60E-11), and between counties growing hay and counties growing soybeans (p-value = 1.50E-11). In terms of NPP$_h$ values, there is a significant difference (p-value ≤ 0.05) between 10 pairs (Tables 4 & 5).

Table 4: Results of the pairwise comparisons using Wilcoxon rank sum test, examining the which counties exhibit significant differences in terms of %HANPP$_0$ based on the crop grown over the most land area (m^2) in each of the counties.

	CHERRIES	*CORN, GRAIN*	*HAY*	*SOYBEANS*
CORN, GRAIN	*0.095*	-	-	-
HAY	*0.98*	*3.6E-11*	-	-
SOYBEANS	*0.095*	*0.094*	*1.5E-11*	-
WHEAT	*1.0*	*0.29*	*0.23*	*0.48*

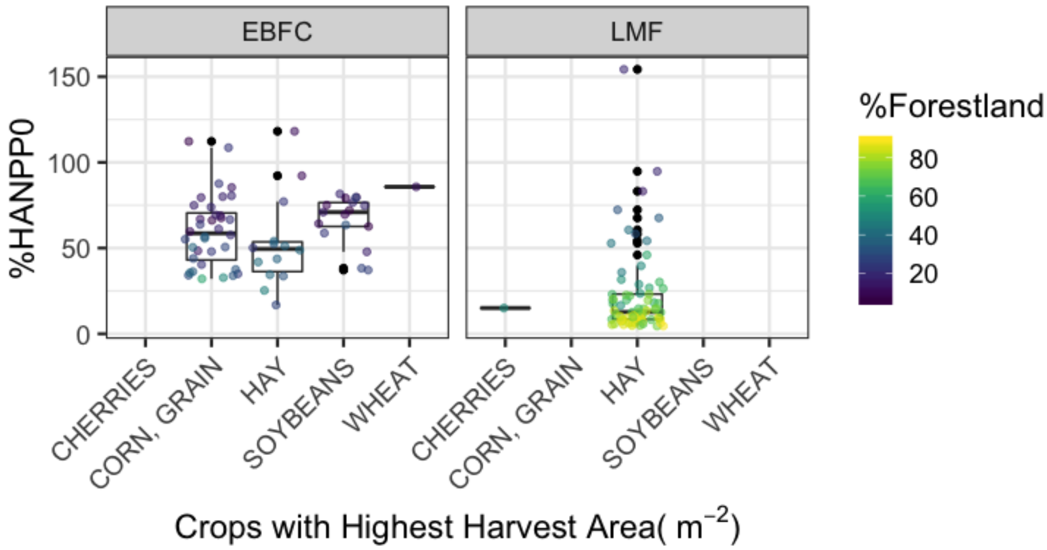
Table 5: Results of the pairwise comparisons using Wilcoxon rank sum test, examining the which counties exhibit significant differences in terms of %HANPP$_0$ based on the crop with the highest NPPh value (kg C m^{-2} yr^{-1}) in each of the counties

	APPLES	CORN, GRAIN	CORN, SILAGE	CORN, SWEET	HAY	HAY, ALFALFA	OATS	POTATOES	SORGHUM
CORN, GRAIN	0.067	-	-	-	-	-	-	-	-
CORN, SILAGE	0.23	0.38	-	-	-	-	-	-	-
CORN, SWEET	0.14	0.025	0.85	-	-	-	-	-	-
HAY	1.o	0.68	0.92	1.0	-	-	-	-	-
HAY, ALFALFA	0.52	0.51	0.46	0.38	0.52	-	-	-	-
OATS	0.38	0.067	0.41	0.41	1.0	0.41	-	-	-
POTATOES	0.067	0.015	0.76	0.41	0.82	0.38	0.14	-	-
SORGHUM	0.015	0.021	0.035	0.00029	0.41	0.68	0.0095	1.2E-07	-
SUGARBEETS	0.091	0.64	0.14	0.021	0.64	0.92	0.057	0.0349	0.38

More in-depth research into the differences in HANPP among crops and crop tending and harvest methods should be explored to provide more precise knowledge on which cropland matrices are more supportive of biodiversity conservation than others in regions where landsharing is necessary. The impact of livestock grazing, which was not analyzed in this project, should also be included in a more in-depth crop energy use analysis. Hay and pasture were the crop uses taking up the most land area in counties with high resiliency potential and low %HANPP$_0$. Thus, understanding how livestock contribute to ecosystem energy dynamics in a mixed forest and crop landscape is an essential next step for understanding the socio-ecological landscape makeup that best promotes biodiversity.

Broad Analysis of the Relation Between %HANPP$_0$ and Forestlands

To analyze the relationship between HANPP and LU/LC, we performed a set of two regressions examining the relationship between the percent forest cover of a county and its %HANPP$_0$, one as a single regional data set and one stratified by ecoregion (Fig. 6, Table 6). We identified four outlier counties in each relationship. These counties–Lake Co. and Cuyahoga



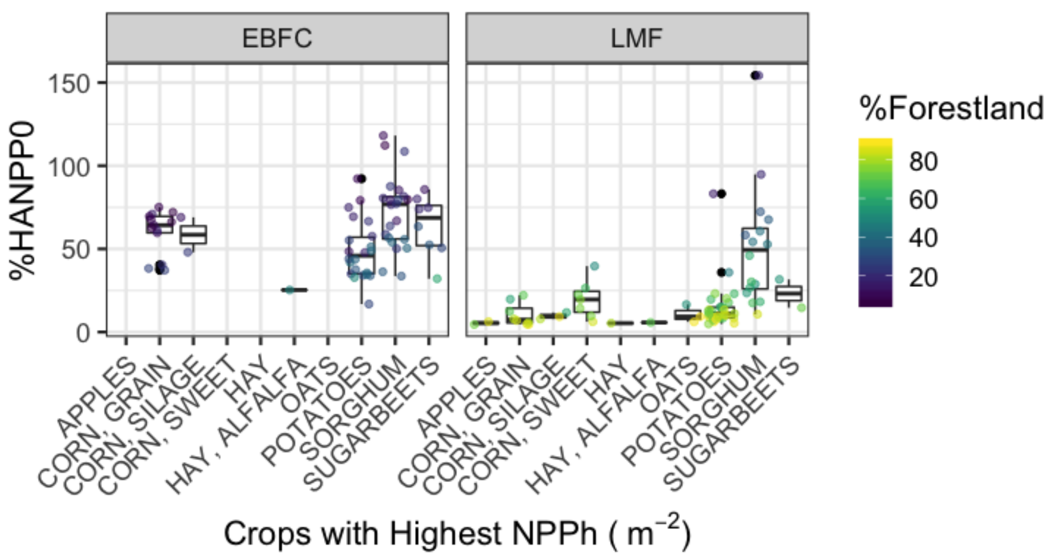


Figure 5: The top row (a) shows %HANPP$_0$ distributed by the crops with the highest harvest areas in all the counties (n=188) and the bottom row (b) shows %HANPP$_0$ distributed by the crops with the highest NPP$_h$ values in all the counties. Large variations in %HANPP$_0$ by crop suggest that methods and timing of crop planting or harvest may play a significant role in how much energy it extracts from the environment. As only the crop with the highest area is shown per column, the agricultural matrix of the county may also contribute to affecting the overall level of NPP harvested from the environment. Note the top graphs show that the crops with the highest NPP$_h$ value per county are different from the crops planted over the most area.

Co., Ohio; Milwaukee Co., Wisconsin; and Wayne Co., Michigan–were contain major urban centers and thus exhibited significantly different LU/LC patterns than the other counties. As

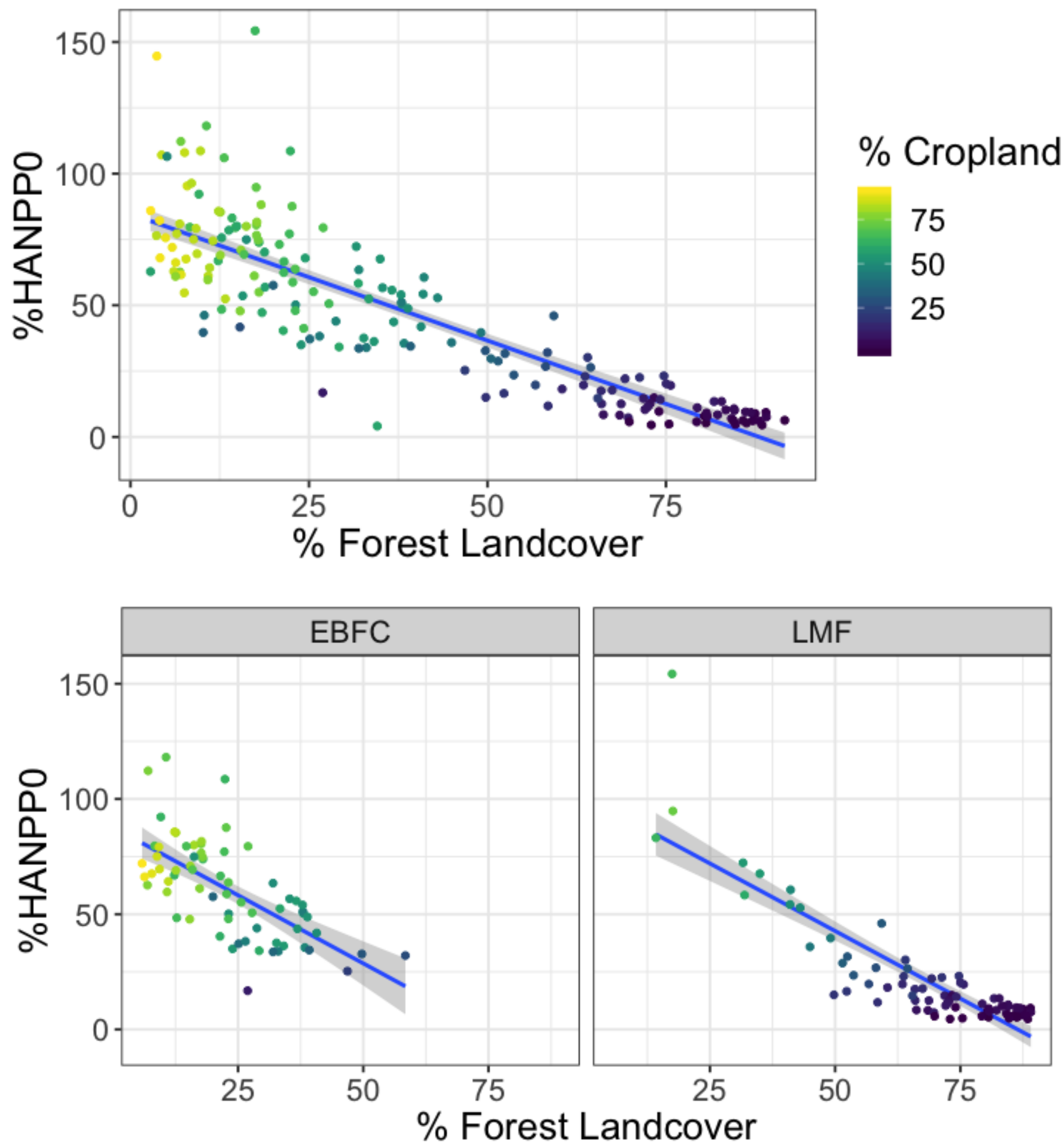


Figure 6: The graphs show the relationship between LU/LC and $\%HANPP_0$, r = 0.85, $r^2 = 0.76$, p<2e-16. The grey area shows the 95% confidence interval of the linear regression. There is considerable more variation about the regression line in areas of low % forest landcover and high % crop landcover than in areas of high % forest landcover. Most of these highly-forested counties are in the LMF province, where the relationship is considerably stronger ($r^2 = 0.85$) than in the EBFC province ($r^2 = 0.42$).

this analysis is focused on managed natural lands, we chose to remove these four counties from our analysis. We also noted the relationships between $\%HANPP_0$ and percent crop landcover was the inverse of its relationship with forest cover; the regression were opposite but nearly identical. Because of this, we chose to just use percent forest cover for the remainder of our analysis, recognizing that lower percent forest cover in a county equated to higher percent crop landcover.

Analysis of HANPP Due to Landuse Change

We chose to focus on $\%HANPP_0$ for our main analysis, but we also briefly analyzed HANPP due to landuse change, or $HANPP_{luc}$. This part of the HANPP equation set is the difference between potential NPP (NPP_0) and actual NPP (NPP_{act}):

$$HANPP_{luc} = NPP_0 - NPP_{act} \tag{3}$$

It measures how much NPP is lost from the ecosystem due to changing LU/LC, like transitions from forestland to cropland. Through our spatial analysis, we found that the corn belt regions of northern Indiana and Ohio showed the most negative values for $HANPP_{luc}$. Negative $HANPP_{luc}$ values occur when $NPP_{act} > NPP_0$, indicating the LU/LC change that occurred in a county resulted in an increase in available ecosystem energy as compared to what was there in the past. Typically, this means there has been some type of technological input into the landscape, such as fertilizer or irrigation, that can artificially raise productivity. In our study site, though, $HANPP_{luc}$ values decrease moving northward through the region (Fig. 7)–the opposite of what would be expected based on other studies (Haberl et al. 2014). One possible reason for this is that in the north, farming generally requires fertilizer and irrigation as the soils are arid and sandy. These inputs would lead to increased productivity on a naturally unproductive landscape. Dividing the distribution of $HANPP_{luc}$ between forestlands and croplands suggests this latter interpretation may be correct, as it is croplands that skew the overall $HANPP_{luc}$ value negative (Fig. 8). On forest lands, a little less than half of counties have low but positive $HANPP_{luc}$ values, suggesting their productivity has slightly decreased over time.

Table 6: Linear regression results examining the relationship between forest cover, %$HANPP_0$ and the two biodiversity metrics, as well as $HANPP_{luc}$ (kg C m^{-2} yr^{-1}) and the two biodiversity metrics.

REGRESSION	P-VALUE	R	R^2
%FOREST ~ MEAN CONNECTEDNESS	< 2.2e-16	0.73	0.53
LMF	*< 2.2e-16*	*0.83*	*0.68*
EBFC	*2.4E-13*	*0.65*	*0.41*
%FOREST ~ MEAN LANDSCAPE DIVERSITY	< 2.2e-16	0.68	0.46
LMF	*1.3E-09*	*0.63*	*0.39*
EBFC	*5.3E-08*	*0.51*	*0.25*
%HANPP0 ~ FORESTED LANDCOVER	<2.0E-16	-0.85	0.76
LMF	*< 2.2e-16*	*-0.88*	*0.78*
EBFC	*2.02E-13*	*-0.65*	*0.42*
$HANPP_{LUC}$ ~ MEAN LOCAL CONNECTEDNESS	1.3E-02	-0.18	0.028
MEAN LANDSCAPE DIVERSITY~ $HANPP_{LUC}$	1.75E-01	-0.10	0.0046

In addition to our spatial analysis, we performed two regression analyses using $HANPP_{luc}$, one comparing it to each of the two biodiversity metrics (Fig. 9). We found that $HANPP_{luc}$ did not strongly correlate with either mean landscape diversity or mean local connectedness, but had a stronger correlation with mean local connectedness ($r = -0.53$, $r^2 = 0.28$, $p<0.001$; Table 6).

Methodological Challenges

Review of Different Crop to NPP Conversion Variables

A note of caution in interpreting the values obtained from the NPP equations is that, beyond the uncertainty inherent in self-reported crops production and yields, harvest indices, a variable essential to translating crop production/yield into NPP, is highly variable among

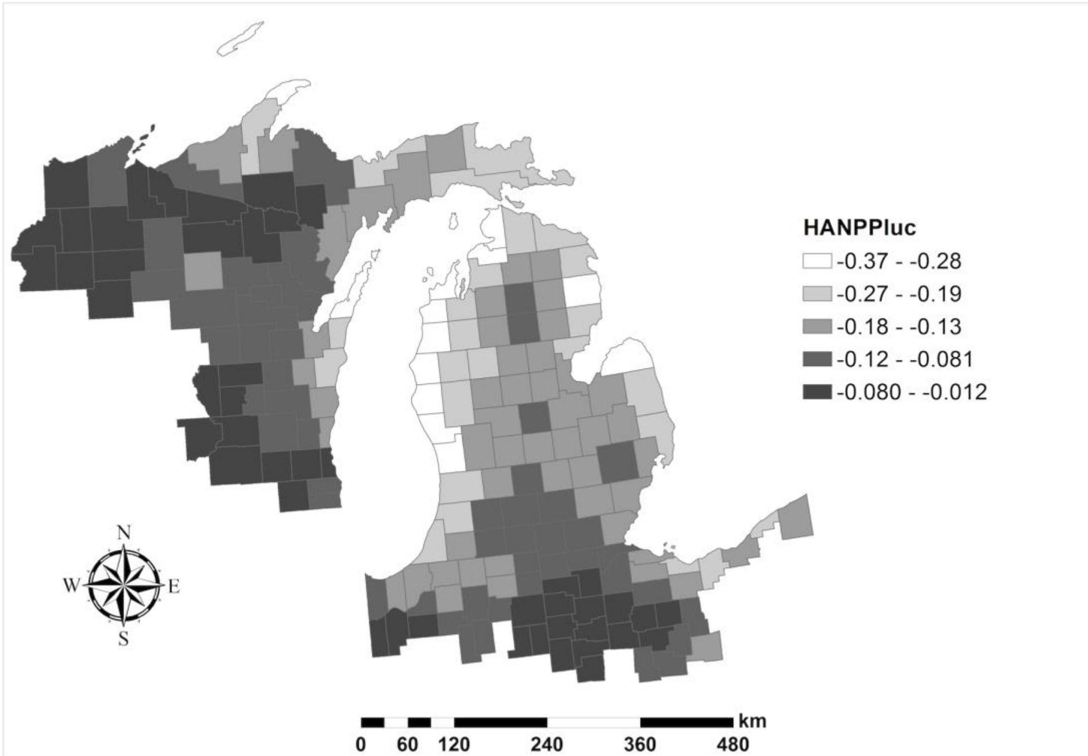


Figure 7: The spatial distribution of HANPP due to landuse change (HANPPluc). Units are in kg C m^{-2} yr^{-1}. As HANPPluc is calculated as the difference between NPP_0 and NPP_{act}, negative values indicate that $NPP_{act} > NPP_0$. This suggests current levels of productivity are greater than those modeled from past ecological data.

studies. Multiple authors have noted the large differences amongst reported harvest indices, and numerous papers reviewed for this study presented quite different values (Smil et al. 1983; Hay, 1995). Since we used the methods presented by Haberl et al (2007), we also chose to use the harvest variables used by those authors, or the references they cited, to maintain consistency between studies. To account for changes in crop production over time, we used the modifiers suggested by Krausman et al (2008, 2013). These harvest variables are also specified to the region, which may help mitigate high levels of uncertainty. For the harvest variables not given by Haberl et al (2007) or Krausman et al (2008, 2013), we used those given in studies cited by these research groups, primarily the model created by

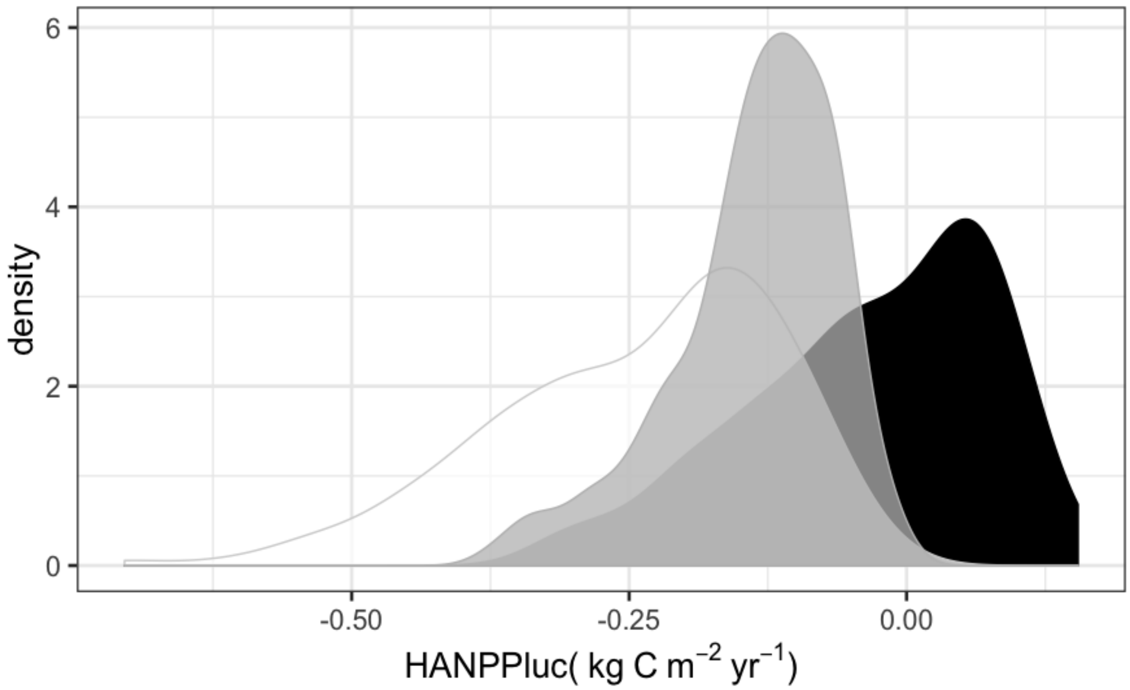


Figure 8: Density plots showing the distribution of $HANPP_{luc}$ on different LU/LC types. The black distribution is $HANPP_{luc}$ on forestlands and the white distribution is $HANPP_{luc}$ on agricultural lands. The grey distribution is the weighted $HANPP_{luc}$ of the combined LU/LC types.

Wirsenius (2000). For fruit and vegetable crops, which were not part of the studies done by the Haberl or Kruasmann research groups, we drew largely on the database created by Monfreda et al (2008), which is the most comprehensive list of HI values, with some values drawn from the work done by Smil (1999).

To estimate the full value of appropriated NPP, we calculated the residues of each crop, and the partition between used residues (e.g. residue that is removed from the land to be used by humans) and the unused residue (e.g. residue that remained unused), noted as the "recovery rate". These estimates are based on grain:straw ratios, also known as residue multipliers and, like the harvest indices, must be observed with caution. These numbers are largely reported for field crops, not fruit and vegetable crops, so residue estimates for these groups are based on more varied sources. As the crop residues not removed from the land–matter thus available to the detritivore branch of the food chain–can be assumed to be the

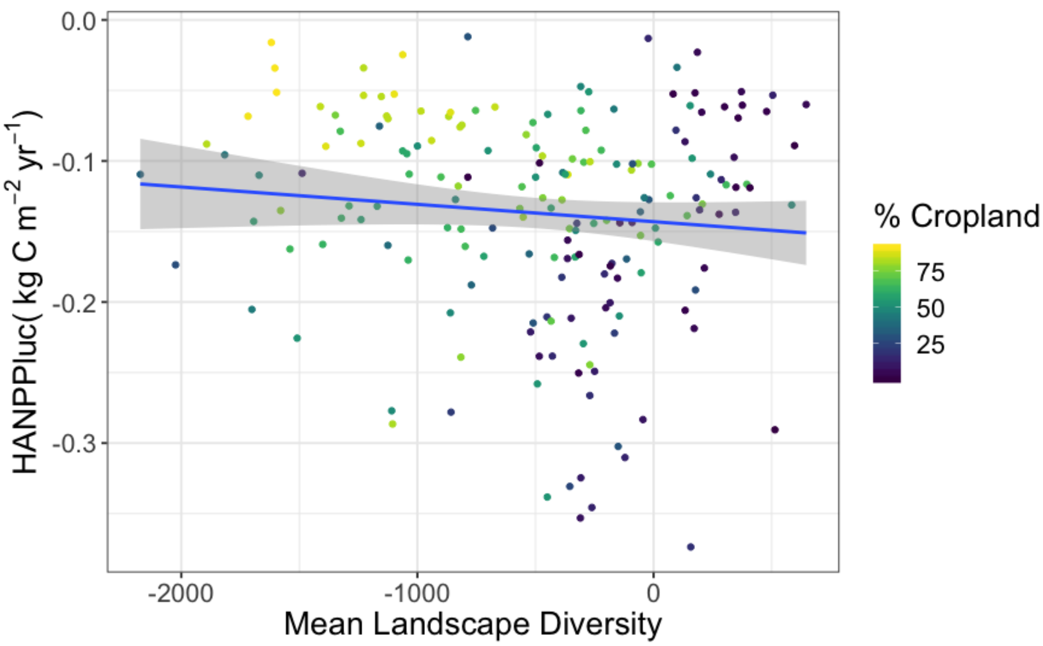


Figure 9(a): The spatial relationship between mean landscape diversity and HANPPluc, with shade indicating to what percent each county it dominated by cropland.

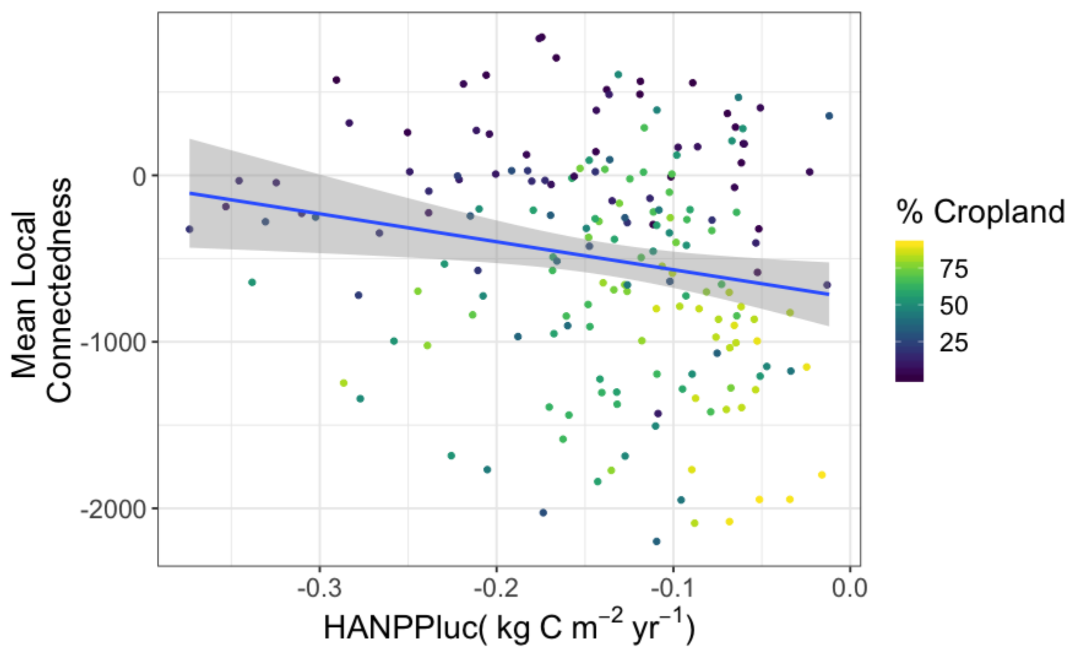


Figure 9(b): The spatial relationship between HANPPluc and mean local connectedness, with shade indicating to what percent each county it dominated by cropland.

Figure 9: Graphs showing the relationship between $HANPP_{luc}$ in kg C m $^{-2}$ y^{-1} and the two biodiversity metrics (each a unitless index). Graph (a) shows how mean landscape diversity impacts $HANPP_{luc}$, and (b) shows how $HANPP_{luc}$ impacts mean local connectedness. The grey area is the 95% confidence interval. Each point represents a county, with darker points indicating low % cropland.

difference between average aboveground NPP and the amount of crop harvested (Krausmann et al. 2008), this value can be assumed to be picked up in the MODIS satellite images and incorporated into the NPP_{act} value. For this reason, we will not use these values as part of the overall analysis.

Different producers deal with residues in different way–farmers may clear residue from the fields, till residue into the soil, or leave residue on top of soils to protect them. This study does not distinguish residue uses to this level of detail, but the uncertainty of residue fates, and how that impacts socioecological metabolisms, should be taken into consideration when examining the results of this study. Haberl et al (2007) also presented a “recovery rate” value based on that presented by Wirsenius (2000). This value was primarily estimated for field crops, and we could find little data for it in other areas of the literature, either for field crops or for the fruit and vegetable crops that make up a significant portion of the crops grown around the Great Lakes. For this reason, we chose not to include these values in the analysis.

Table 7: +Smil (1999) calculated HI based on the equation: (dry matter of harvested crops)/(dry matter of harvested crops + dry matter of crop residues). Smil (1983) collected global averages of crop residue multipliers, derived from FAO data from the 1970s; these are largely field crops. Neither Smil paper differentiated between used and unused crop residues.

++We calculated HI++ using the equation suggested in Smil (1999) and the harvest data we collected from the USDA and different state agricultural departments, while HI+ are the values reported in Smil (1999).

§Haberl et al (2007) and Krausmann et al (2008) use region-specific values drawn largely from Wirsenius (2000); the two rates used, harvest factors and recovery rates, are a multiplier to estimate crop residues and a multiplier to estimate the portion of crop residues used (extracted from the system, e.g. as fodder). There are several crops, largely fruit and vegetable crops, for which there is very little information available. Krausmann et al (2008) assumed for these crops that "crop residues to be the difference between average aboveground NPP per unit of cropland and the amount of primary crop harvested." Wirsenius (2000)–a source for both Krausmann et al (2008) and Haberl et al (2007)–defined the recovery rate as, "'Not recovered' simply means that the generated amount, or a fraction of the generated amount, is not made available for further use within or outside the food system. Thus, by definition, the amount 'not recovered' is lost (from a use point of view, that is). In the FPD model, 'not recovered' is expressed by its reverse quantity, here called 'recovery rate'." Wirsenius (2000) cautioned about the uncertainty of the residue recovery estimates used in the model he created. In cases where data was lacking, a standard value of 90 percent was used for all regions, but this is believed to be an overestimate, particularly with areas that have low yields.

The HI values for fruit and vegetables in the HI column came from work done in Monfreda et al (2008)), while the field crop HI was obtained from Lobell et al (2002). Lobell et al (2002) created these estimates based on a review by Hay (1995) intended for a US-based study. Monfreda et al (2008) also drew on Hay's work, as well as others for cereal crops. The research group's final choice of numbers are meant to "approximates the distribution of global cropland NPP." As this is the most comprehensive list of HI values, we ended up drawing on it for fruit and vegetable crops. we also drew the dry fraction (*DF*)/moisture content (*MC*), fraction of production in aboveground biomass (f_{AG}), and percent carbon content (*C*) from the same sources.

CROP	Residue Multiplier +	HI+	HI++	HI§	HI*	Harvest Residue Factor§	Recovery Rate§	DF*	MC *	fAG*	C*
APPLES	1.67	0.38	0.37		0.3	2.5	0.9	0.16	0.84	0.75	0.45
CHERRIES	1.67	0.38	0.37		0.3	2.5	0.9	0.14	0.86	-0.25	0.45
CHERRIES	1.67	0.38	0.37		0.3	2.5	0.9	0.14	0.86	0.75	0.45
PEACHES	1.67	0.38	0.37		0.3	2.5	0.9	0.14	0.86	-0.25	0.45
PEACHES	1.67	0.38	0.37		0.3	2.5	0.9	0.14	0.86	0.75	0.45
GRAPES	1.67	0.38	0.37		0.3	2.5	0.9	0.19	0.81	0.75	0.45
BLUEBERRIES	1.67	0.38	0.37		0.3	2.5	0.9	0.15	0.85	0.75	0.45
PEAS, GREEN	0.50	0.49			0.45	1	0.9	0.13	0.87	0.85	0.45
BEANS, SNAP	0.50	0.49	0.67		0.45	1	0.9	0.1	0.9	-0.15	0.45
BEANS, SNAP	0.50	0.49	0.67		0.45	1	0.9	0.1	0.9	0.85	0.45
CUCUMBERS		0.38			0.45		0.9	0.04	0.96	-0.15	0.45

Table 7 cont.

CROP	Residue Multiplier +	HI+	HI++	HI§	HI*	Harvest Residue Factor§	Recovery Rate§	DF*	MC *	fAG*	C*
CUCUMBERS		0.38			0.45		0.9	0.04	0.96	0.85	0.45
POTATOES	0.20	0.40	0.83	0.5	0.5	1	0.9	0.28	0.72	0.8	0.45
CORN, SWEET	1.20	0.38	0.45		0.45	1.2	0.7	0.13	0.87	0.85	0.45
CORN, SWEET	1.20	0.38	0.45		0.45	1.2	0.7	0.13	0.87	1.85	0.45
CORN, GRAIN	1.20	0.40	0.45	0.45	0.45	1.2	0.7	0.89	0.11	0.85	0.45
SOYBEANS	1.00	0.52	0.50	0.45	0.4	1.2	0.7	0.9	0.1	0.87	0.45
WHEAT	1.50	0.40	0.40	0.45	0.4	1.2	0.7	0.89	0.11	0.83	0.45
HAY					1	1.3	0.9	0.85	0.15	0.53	0.45
HAY,ALFALFA					1	1.3	0.9	0.85	0.15	0.53	0.45
SORGHUM	1.20	0.40	0.45	0.45	0.4	1.2	0.7	0.9	0.1	0.8	0.45
SORGHUM	1.20	0.40	0.45	0.45	0.4	1.2	0.7	0.9	0.1	0.8	0.45
BARLEY	1.20	0.40	0.45	0.45	0.4	1.2	0.7	0.88	0.12	0.67	0.45
CORN, SILAGE	1.20	0.40	0.45		1	1.3	0.9	0.35	0.65	0.85	0.45
SUNFLOWERS		0.52		0.35	0.35	1.9	0.5	0.9	0.1	0.94	0.45
OATS	1.50	0.40	0.40		0.4	1.2	0.7	0.89	0.11	0.71	0.45
SUGARBEETS	0.10	0.56	0.91	0.65	0.4	0.5	0	0.15	0.85	0.8	0.45

Regional HANPP Calculation

Calculation of HANPP at the regional scale poses multiple challenges. The primary challenge is a lack of comprehensive, consistent harvest data at appropriately fine scales. The finest scale available for forest harvest data from the USFS was the county scale, and even this resolution came with the caveat of high errors due to the method of data collection (Burrill, 2018; S. Pugh, personal communication, 2017). This leads to uncertainty in forest area and average harvest rates, particularly in counties with smaller amounts of forested land. Crop harvest data from the USDA Agricultural Census is similarly limited to the county scale, and is self-reported by farmers, who may own land in multiple counties or may have their data removed from the census due to privacy concerns. Even when harvest data are available, estimating NPP_h requires knowledge of harvest indices (a ratio of the harvested yield of a crop to total aboveground crop biomass), moisture content, and root:shoot ratios for the total variety of crops in the region. These variables have only been calculated for a handful of crops–largely cereal crops–and vary extensively throughout the literature.

Another challenge is the ability of models that estimate NPP_{act} to differentiate among different crop landcover types. The model algorithm behind MODIS NPP datasets has been shown to estimate NPP levels 30 percent lower, on average, than those calculated based on USDA harvest data in cropland-dominated landscapes, is less spatially sensitive in terms of identifying croplands, and is not always able to identify different types of crops (Li et al. 2014). The model also results in a lot of scatter, and so is highly uncertain in its NPP outputs for cropland (Li et al. 2014). The different resolutions and methods of calculation in which each data set used for this analysis originated may also have led to underestimations of $\%HANPP_0$ and $HANPP_{luc}$. The NPP_{act} and NPP_0 models used in this analysis resulted in NPP_{act} values that were consistently higher than the NPP_0 values. These differences lead to high potential error in the $HANPP_{luc}$ variable, in particular, as it is calculated from NPP_{act} and NPP_0 alone. Although the urban development occurring in the region (Brown, 2003; Robinson, 2012) and the high degree of technological inputs on cropland (e.g. fertilizer) in some regions suggest low $HANPP_{luc}$ values are reasonable, the fully negative $HANPP_{luc}$ values we obtained may be a methodological artifact. These uncertainties highlight the need for comprehensive LU/LC data sets and models that include landscape NPP calculations at the local and regional resolutions needed to plan for biodiversity conservation.

Works Cited

Abrams, M. D. (1992). Fire and Development of Oak Forests. *BioScience*, *42*(5), 346–353.

An, L., Brown, D. G., Nassauer, J. I., & Low, B. (2011). Variations in development of exurban residential landscapes: timing, location, and driving forces. *Journal of Land Use Science*, *6*(1), 13–32. https://doi.org/10.1080/1747423X.2010.500686

Andersen, C. B., Donovan, R. K., & Quinn, J. E. (2015). Human Appropriation of Net Primary Production (HANPP) in an Agriculturally-Dominated Watershed, Southeastern USA. *Land*, *4*(2), 513–540. https://doi.org/10.3390/land4020513

Anderson, M. G., Barnett, A., Clark, M., Sheldon, A. O., Prince, J., & Vickery, B. (2016). *Resilient and Connected Landscapes for Terrestrial Conservation* (p. 161). Boston, MA: The Nature Conservancy, Eastern Conservation Science, Eastern Regional Office.

Anderson, M. G., Clark, M. M., Cornett, M. W., Hall, K. R., Olivero Sheldon, A., & Prince, J. (2018). *Resilient Sites for Terrestrial Conservation in the Great Lakes and Tallgrass Prairie*. Boston, MA: The Nature Conservancy, Eastern Conservation Science and North America Region. Retrieved from https://easterndivision.s3.amazonaws.com/Terrestrial/Great_Lakes_Resilience/Great_Lakes_and_Tallgrass_Prairie_Resilience_05_11_18.pdf
Data accessed from http://maps.tnc.org/resilientland/.

Anderson, M. G., & Ferree, C. E. (2010). Conserving the Stage: Climate Change and the Geophysical Underpinnings of Species Diversity. *PLOS ONE*, *5*(7), e11554. https://doi.org/10.1371/journal.pone.0011554

Annual Estimate of the Resident Population: April 1, 2010 to July 1, 2017. (2018, March). US Census Bureua, Population Division.

ArcGIS ArcMap. (2017). (Version 10.5.1). Esri.

Asbjornsen, H., Hernandez-Santana, V., Liebman, M., Bayala, J., Chen, J., Helmers, M., … Schulte, L. A. (2014). Targeting perennial vegetation in agricultural landscapes for enhancing ecosystem services. *Renewable Agriculture and Food Systems; Cambridge*, *29*(2), 101–125. http://dx.doi.org.proxy.lib.umich.edu/10.1017/S1742170512000385

Bailey, R. (1994). Bailey's Ecoregions of the Conterminous United States. vector digital data, United States: US Forest Service. Retrieved from https://www.sciencebase.gov/catalog/item/54244abde4b037b608f9e23d

Birdsey, R. A. (1992). *Carbon Storage and Accumulation in United States Forest Ecosystems* (General Technical Report No. WO-59) (p. 55). US Forest Service. Retrieved from https://www.nrs.fs.fed.us/pubs/gtr/gtr_wo059.pdf

Bogue, M. B. (2000). *Fishing the Great Lakes: An Environmental History, 1783-1933*. Madison, WI: The University of Wisconsin Press.

Breffle, W. S., Muralidharan, D., Donovan, R. P., Liu, F., Mukherjee, A., & Jin, Y. (2013). Socioeconomic evaluation of the impact of natural resource stressors on human-use services in the Great Lakes environment: A Lake Michigan case study. *Resources Policy*, *38*(2), 152–161. https://doi.org/10.1016/j.resourpol.2012.10.004

Brown, D. G. (2003). Land use and forest cover on private parcels in the Upper Midwest USA, 1970 to 1990. *Landscape Ecology*, *18*(8), 777–790. https://doi.org/10.1023/B:LAND.0000014470.16973.cb

Brown, D. G., Robinson, D. T., An, L., Nassauer, J. I., Zellner, M., Rand, W., … Wang, Z. (2008). Exurbia from the bottom-up: Confronting empirical challenges to characterizing a complex system. *Geoforum*, *39*(2), 805–818. https://doi.org/10.1016/j.geoforum.2007.02.010

Burrill, E. A. (2018). The Forest Inventory and Analysis Database: Database Description and User Guide for Phase 2 (version 7.0.1), 942.

Chave Jerome, Coomes David, Jansen Steven, Lewis Simon L., Swenson Nathan G., & Zanne Amy E. (2009). Towards a worldwide wood economics spectrum. *Ecology Letters*, *12*(4), 351–366. https://doi.org/10.1111/j.1461-0248.2009.01285.x

CropScape - NASS CDL Program. (n.d.). Retrieved May 13, 2018, from https://nassgeodata.gmu.edu/CropScape/

Currie, William S. (2012). Energy Flow. In Oxford Bibliographies Online: Ecology. Ed. David Gibson. New York: Oxford University Press. http://oxfordbibliographiesonline.com DOI: 10.1093/OBO/9780199830060-0047.

Currie, W. S., Kiger, S., Nassauer, J. I., Hutchins, M., Marshall, L. L., Brown, D. G., ... Hart, S. K. (2016). Multi-scale heterogeneity in vegetation and soil carbon in exurban residential land of southeastern Michigan, USA. *Ecological Applications*, *26*(5), 1421–1436. https://doi.org/10.1890/15-0817

DeFries, R. S., Foley, J. A., & Asner, G. P. (2004). Land-use choices: balancing human needs and ecosystem function. *Frontiers in Ecology and the Environment*, *2*(5), 249–257. https://doi.org/10.1890/1540-9295(2004)002[0249:LCBHNA]2.0.CO;2

Elvidge, C. D., Nemani, R., & Vogelmann, J. E. (2003). Development Sprawl Impacts on the Terrestrial Carbon Dynamics of the United States. NOAA.

Ensign, S. H., & Mallin, M. A. (2001). Stream water quality changes following timber harvest in a coastal plain swamp forest. *Water Research*, *35*(14), 3381–3390. https://doi.org/10.1016/S0043-1354(01)00060-4

Fahrig, L. (2003). Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *34*, 487–515.

Fan, S. (2018a). Indiana County Boundaries. polygon, Great Lakes Commission des Grands Lacs. Retrieved from https://www.glc.org/greatlakesgis/maplayers

Fan, S. (2018b). Michigan County Boundaries. polygon, Great Lakes Commission des Grands Lacs. Retrieved from https://www.glc.org/greatlakesgis/maplayers

Fan, S. (2018c). Ohio County Boundaries. polygon, Great Lakes Commission des Grands Lacs. Retrieved from https://www.glc.org/greatlakesgis/maplayers

Fan, S. (2018d). Wisconsin County Boundaries. polygon, Great Lakes Commission des Grands Lacs. Retrieved from https://www.glc.org/greatlakesgis/maplayers

Gerber, L. R. (2016). Conservation triage or injurious neglect in endangered species recovery. *Proceedings of the National Academy of Sciences*, *113*(13), 3563–3566. https://doi.org/10.1073/pnas.1525085113

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., & Sitch, S. (2004). Terrestrial vegetation and water balance—hydrological evaluation of a dynamic global vegetation model. *Journal of Hydrology*, *286*(1), 249–270. https://doi.org/10.1016/j.jhydrol.2003.09.029

Gonthier, D. J., Ennis, K. K., Farinas, S., Hsieh, H.-Y., Iverson, A. L., Batary, P., ... Perfecto, I. (2014). Biodiversity conservation in agriculture requires a multi-scale approach. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1791), 20141358–20141358. https://doi.org/10.1098/rspb.2014.1358

Graham, J. B., Nassauer, J. I., Currie, W. S., Ssegane, H., & Negri, M. C. (2017). Assessing wild bees in perennial bioenergy landscapes: effects of bioenergy crop composition, landscape configuration, and bioenergy crop area. *Landscape Ecology; Dordrecht*, *32*(5), 1023–1037. http://dx.doi.org.proxy.lib.umich.edu/10.1007/s10980-017-0506-y

Gustafson, E. J., & Loehle, C. (2008). How will the changing industrial forest landscape affect forest sustainability? Retrieved from https://www.nrs.fs.fed.us/pubs/9745

Haberl, H. (1997). Human Appropriation of Net Primary Production as an Environmental Indicator: Implications for Sustainable Development. *Ambio*, *26*(3), 143–146.

Haberl, H., Erb, K. H., Krausmann, F., Gaube, V., Bondeau, A., Plutzar, C., … Fischer-Kowalski, M. (2007). Quantifying and Mapping the Human Appropriation of Net Primary Production in Earth's Terrestrial Ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(31), 12942–12947.

Haberl, H., Erb, K.-H., & Krausmann, F. (2014). Human Appropriation of Net Primary Production: Patterns, Trends, and Planetary Boundaries. *Annual Review of Environment and Resources*, *39*(1), 363–391. https://doi.org/10.1146/annurev-environ-121912-094620

Haberl, H., Erb, K.-H., Krausmann, F., Loibl, W., Schulz, N., & Weisz, H. (2001). Changes in ecosystem processes induced by land use: Human appropriation of aboveground NPP and its influence on standing crop in Austria. *Global Biogeochemical Cycles*, *15*(4), 929–942. https://doi.org/10.1029/2000GB001280

Haberl, H., Erb, K.-H., Plutzar, C., Fischer-Kowalski, M., & Krausmann, F. (2012). Human Appropriation of Net Primary Productivity (HANPP) as an Indicator for Pressures on Biodiversity. In *Scientific Committee on Problems of the Environment (SCOPE) Series : Sustainability Indicators : A Scientific Assessment* (pp. 271–283). Washington, US: Island Press. Retrieved from http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10222014

Haberl, H., Gaube, V., Díaz-Delgado, R., Krauze, K., Neuner, A., Peterseil, J., … Vadineanu, A. (2009). Towards an integrated model of socioeconomic biodiversity drivers, pressures and impacts. A feasibility study based on three European long-term socio-ecological research platforms. *Ecological Economics*, *68*(6), 1797–1812. https://doi.org/10.1016/j.ecolecon.2008.11.013

Haberl, H., Schulz, N. B., Plutzar, C., Erb, K. H., Krausmann, F., Loibl, W., … Zulka, P. (2004). Human appropriation of net primary production and species diversity in agricultural landscapes. *Agriculture, Ecosystems & Environment*, *102*(2), 213–218. https://doi.org/10.1016/j.agee.2003.07.004

Han, W., Yang, Z., Di, L., & Mueller, R. (2012). CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Computers and Electronics in Agriculture*, *84*, 111–123. https://doi.org/10.1016/j.compag.2012.03.005

Handler, S., Duveneck, M. J., Iverson, L., Peters, E., Scheller, R. M., Wythers, K. R., … Ziel, R. ; (2014). *Michigan forest ecosystem vulnerability assessment and synthesis: a report from the Northwoods Climate Change Response Framework project*. Retrieved from http://www.treesearch.fs.fed.us.ezproxy1.lib.asu.edu/pubs/45688

Hawkins, B. A., Porter, E. E., & Diniz-Filho, J. A. F. (2003). Productivity and History as Predictors of the Latitudinal Diversity Gradient of Terrestrial Birds. *Ecology*, *84*(6), 1608–1623.

Hay, R. K. M. (1995). Harvest index: a review of its use in plant breeding and crop physiology. *Annals of Applied Biology*, *126*(1), 197–216. https://doi.org/10.1111/j.1744-7348.1995.tb05015.x

Hicke, J. A., Lobell, D. B., & Asner, G. P. (2004). Cropland Area and Net Primary Production Computed from 30 Years of USDA Agricultural Harvest Data. *Earth Interactions*, *8*(10), 1–20. https://doi.org/10.1175/1087-3562(2004)008<0001:CAANPP>2.0.CO;2

Host, G. E., Pregitzer, K. S., Ramm, C. W., Hart, J. B., & Cleland, D. T. (1987). Landform-Mediated Differences in Successional Pathways Among Upland Forest Ecosystems in Northwestern Lower Michigan. *Forest Science*, *33*(2), 445–457.

Janowiak, M. K., Iverson, L. R., Mladenoff, D. J., Peters, E., Wythers, K. R., Xi, W., ... Ziel, R. (2014). Forest ecosystem vulnerability assessment and synthesis for northern Wisconsin and western Upper Michigan: a report from the Northwoods Climate Change Response Framework project. Retrieved from http://www.nrs.fs.fed.us/pubs/46393

Johnson, L. B., Kovalenko, K. E., Host, G. E., Brady, V. J., Bracey, A. M., Brown, T. N., ... Niemi, G. J. (2015). *Great Lakes Environmental Indicators Testing and Refinement: Final Report* (U.S. EPA GLNPO Project Identifier: EPAGLNPO-2010-NS-5-1071-795. Natural Resources Research Institute Technical Report No. NRRI/TR-2015/56).

Jones, A., Schindel, M., & Scott, S. (2015). *Mapping Habitat Connectivity for the Great Sage-Grouse in Oregon's Sage-Grouse Conservation Partnership (SageCon) Assessment Area.* The Nature Conservancy (Portland, OR) in partial fulfillment of the BLM Cooperative Agreement L12AC2061. Retrieved from https://www.researchgate.net/profile/Aaron_Jones17/publication/301341799_Mapping_Habitat_Connectivity_for_Greater_Sage-Grouse_in_Oregon%27s_Sage-Grouse_Conservation_Partnership_SageCon_Assessment_Area/links/57132fc108ae39beb87a54ae.pdf?origin=publication_detail

Kells, B. J., & Swinton, S. M. (2014). Profitability of Cellulosic Biomass Production in the Northern Great Lakes Region. *Agronomy Journal; Madison*, *106*(2), 397–406.

Krausmann, F., Erb, K.-H., Gingrich, S., Haberl, H., Bondeau, A., Gaube, V., ... Searchinger, T. D. (2013). Global human appropriation of net primary production doubled in the 20th century. *Proceedings of the National Academy of Sciences*, *110*(25), 10324–10329. https://doi.org/10.1073/pnas.1211349110

Krausmann, F., Erb, K.-H., Gingrich, S., Lauk, C., & Haberl, H. (2008). Global patterns of socioeconomic biomass flows in the year 2000: A comprehensive assessment of supply, consumption and constraints. *Ecological Economics*, *65*(3), 471–487. https://doi.org/10.1016/j.ecolecon.2007.07.012

Kremen, C. (2015). Reframing the land-sparing/land-sharing debate for biodiversity conservation. *Annals of the New York Academy of Sciences*, *1355*(1), 52–76. https://doi.org/10.1111/nyas.12845

Lawler, J. J., Ackerly, D. D., Albano, C. M., Anderson, M. G., Dobrowski, S. Z., Gill, J. L., ... Weiss, S. B. (2015). The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. *Conservation Biology*, *29*(3), 618–629. https://doi.org/10.1111/cobi.12505

Li, Z., Liu, S., Tan, Z., Bliss, N. B., Young, C. J., West, T. O., & Ogle, S. M. (2014). Comparing cropland net primary production estimates from inventory, a satellite-based model, and a process-based model in the Midwest of the United States. *Ecological Modelling*, *277*, 1–12. https://doi.org/10.1016/j.ecolmodel.2014.01.012

Lindenmayer, D. B., & Fischer, J. (2006). *Habitat Fragmentation and Landscape Change: An Ecological and Conservation Synthesis*. Washington, UNITED STATES: Island Press.

Retrieved from http://ebookcentral.proquest.com/lib/umichigan/detail.action?docID=3317424

Lobell, D. B., Hicke, J. A., Asner, G. P., Field, C. B., Tucker, C. J., & Los, S. O. (2002). Satellite estimates of productivity and light use efficiency in United States agriculture, 1982–98. *Global Change Biology*, *8*(8), 722–735. https://doi.org/10.1046/j.1365-2486.2002.00503.x

Marull, J., Font, C., Tello, E., Fullana, N., Domene, E., Pons, M., & Galán, E. (2016). Towards an energy–landscape integrated analysis? Exploring the links between socio-metabolic disturbance and landscape ecology performance (Mallorca, Spain, 1956–2011). *Landscape Ecology*, *31*(2), 317–336. https://doi.org/10.1007/s10980-015-0245-x

Midwest Vegetable Production Guide for Commercial Growers 2018. (2018). Purdue Agricultural Communications Service. Retrieved from https://ag.purdue.edu/btny/ppdl

Mittelbach, G. G., Steiner, C. F., Scheiner, S. M., Gross, K. L., Reynolds, H. L., Waide, R. B., ... Gough, L. (2001). What Is the Observed Relationship between Species Richness and Productivity? *Ecology*, *82*(9), 2381–2396.

MOD17A3H.006: Terra Net Primary Production Yearly Global 500m. (2015). USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota: NASA EOSDIS Land Processes DAAC.

Monfreda, C., Ramankutty, N., & Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, *22*(1), GB1022. https://doi.org/10.1029/2007GB002947

MYD17A3H.006: Aqua Net Primary Production Yearly Global 500m. (2015). USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota: NASA EOSDIS Land Processes DAAC.

Ohio Department of Agriculture. (2007). *2007 Ohio Department of Agriculture Annual Report and Statistics*. Retrieved from http://www.agri.ohio.gov/divs/Admin/Docs/AnnReports/ODA_Comm_AnnRpt_2007.pdf

Ohio Department of Agriculture. (2012). *Ohio Department of Agriculture 2012 Annual Report and Statistics*. State of Ohio. Retrieved from http://www.agri.ohio.gov/divs/Admin/Docs/AnnReports/ODA_Comm_AnnRpt_2012.pdf

Ohio Department of Agriculture. (2013). *Ohio Department of Agriculture 2013 Annual Report and Statistics*. State of Ohio. Retrieved from http://www.agri.ohio.gov/divs/communications/docs/ODA_Comm_AnnRpt_2013.pdf

O'Neill, D. W., Tyedmers, P. H., & Beazley, K. F. (2007). Human appropriation of net primary production (HANPP) in Nova Scotia, Canada. *Regional Environmental Change*, *7*(1), 1–14. https://doi.org/10.1007/s10113-006-0021-1

Pimm, S. L., & Raven, P. (2000). Biodiversity: Extinction by numbers. *Nature*, *403*(6772), 843–845. https://doi.org/10.1038/35002708

Plutzar, C., Kroisleitner, C., Haberl, H., Fetzel, T., Bulgheroni, C., Beringer, T., ... Erb, K.-H. (2016). Changes in the spatial patterns of human appropriation of net primary production (HANPP) in Europe 1990–2006. *Regional Environmental Change*, *16*(5), 1225–1238. https://doi.org/10.1007/s10113-015-0820-3

Prince, S. D., Haskett, J., Steininger, M., Strand, H., & Wright, R. (2001). Net Primary Production of U.s. Midwest Croplands from Agricultural Harvest Yield Data.

Ecological Applications, *11*(4), 1194–1205. https://doi.org/10.1890/1051-0761(2001)011[1194:NPPOUS]2.0.CO;2

Robinson, D. T. (2012). Land-cover fragmentation and configuration of ownership parcels in an exurban landscape. *Urban Ecosystems*, *15*(1), 53–69. https://doi.org/10.1007/s11252-011-0205-4

Shivan, G. C., & Potter-Witter, K. (2011). An Examination of Michigan's Logging Sector in the Emerging Bioenergy Market. *Forest Products Journal; Madison*, *61*(6), 459–465.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., ... Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*(2), 161–185. https://doi.org/10.1046/j.1365-2486.2003.00569.x

Slater, S., Keegstra, K., & Donohue, T. J. (2010). The US Department of Energy Great Lakes Bioenergy Research Center: Midwestern Biomass as a Resource for Renewable Fuels. *BioEnergy Research*, *3*(1), 3–5. https://doi.org/10.1007/s12155-009-9075-5

Smil, V. (1983). Crop Residues. In *Biomass Energies* (pp. 163–236). Springer, Boston, MA. https://doi.org/10.1007/978-1-4613-3691-4_4

Smil, V. (1999). Crop Residues: Agriculture's Largest HarvestCrop residues incorporate more than half of the world's agricultural phytomass. *BioScience*, *49*(4), 299–308. https://doi.org/10.2307/1313613

Sousounis, P. J., & Bisanz, J. M. (Eds.). (2000). *Preparing for a changing climate: the potential consequences of climate variability and change: Great Lakes*. Ann Arbor, Mich: Great Lakes Regional Assessment, University of Michigan, Atmospheric, Oceanic and Space Sciences Department.

Steen-Adams, M. M. ., 2, Langston, N., Adams, M. D. O. ., & Mladenoff, D. J. . (2015). Historical framework to explain long-term coupled human and natural system feedbacks: application to a multiple-ownership forest landscape in the northern Great Lakes region, USA. *Ecology & Society*, *20*(1), 140–160. https://doi.org/10.5751/ES-06930-200128

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, *17*(7), 866–880. https://doi.org/10.1111/ele.12277

Taylor, C., Teran, J., Vale, K., & Woodstock, H. (2012). *2012 Wisconsin Agricultural Statistics* (Annual Statistics Bulletin). State of Wisconsin: USDA's National Agricultural Statistics Service Wisconsin Field Office & the Wisconsin Department of Agriculture, Trade, and Consumer Protection. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/Annual_Statistical_Bulletin/bulletin2012_web.pdf

Taylor, C., Vale, K., & Woodstock, H. (2013). *2013 Wisconsin Agricultural Statistics* (Annual Statistics Bulletin). State of Wisconsin: USDA's National Agricultural Statistics Service Wisconsin Field Office & the Wisconsin Department of Agriculture, Trade, and Consumer Protection. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/Annual_Statistical_Bulletin/bulletin2013_web.pdf

The Michigan Department of Natural Resources. (2018). Commercial Timber Sales. Retrieved July 27, 2018, from https://www.michigan.gov/dnr/0,4570,7-350-79136_79237_80912---,00.html

Theobald, D. M. (2005). Landscape Patterns of Exurban Growth in the USA from 1980 to 2020. *Ecology and Society, 10*(1), 32.

Turner, D. P., Ritts, W. D., Law, B. E., Cohen, W. B., Yang, Z., Hudiburg, T., … Duane, M. (2007). Scaling net ecosystem production and net biome production over a heterogeneous region in the western United States, 16.

US EPA, R. 05. (2015, September 18). Great Lakes Facts and Figures [Overviews and Factsheets]. Retrieved July 15, 2018, from https://www.epa.gov/greatlakes/great-lakes-facts-and-figures

USDA/NASS QuickStats Ad-hoc Query Tool. (2007, 2012). Retrieved March 20, 2018, from https://quickstats.nass.usda.gov/

USDA's National Agricultural Statistics Service Indiana Field Office. (2007). *Indiana 2007-2008 Agricultural Statistics: Crop Summary* (Annual Statistics Bulletin) (pp. 31–34). Retrieved from https://www.nass.usda.gov/Statistics_by_State/Indiana/Publications/Annual_Statistical_Bulletin/0708/pg31-34.pdf

USDA's National Agricultural Statistics Service Indiana Field Office. (2012). *Indiana 2012-2013 Agricultural Statistics: Crop Summary* (Annual Statistics Bulletin) (p. 33). Retrieved from https://www.nass.usda.gov/Statistics_by_State/Indiana/Publications/Annual_Statistical_Bulletin/1213/pg33.pdf

USDA's National Agricultural Statistics Service Michigan Field Office. (2008). *Michigan Agricultural Statistics: 2007-2008* (Annual Statistics Bulletin). State of Michigan. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Michigan/Publications/Annual_Statistical_Bulletin/stats08/agstat-all-08.pdf

USDA's National Agricultural Statistics Service Michigan Field Office. (2013). *Michigan Agricultural Statistics: 2012-13* (Annual Statistics Bulletin). State of Michigan. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Michigan/Publications/Annual_Statistical_Bulletin/stats13/agstat13.pdf

Vačkář, D., Harmáčková, Z. V., Kaňková, H., & Stupková, K. (2016). Human transformation of ecosystems: Comparing protected and unprotected areas with natural baselines. *Ecological Indicators*, *66*, 321–328. https://doi.org/10.1016/j.ecolind.2016.02.001

Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human Domination of Earth9s Ecosystems. *Science*, *277*(5325), 494–499. https://doi.org/10.1126/science.277.5325.494

Wang, X., Burns, D. A., Yanai, R. D., Briggs, R. D., & Germain, R. H. (2006). Changes in stream chemistry and nutrient export following a partial harvest in the Catskill Mountains, New York, USA. *Forest Ecology and Management*, *223*(1–3), 103–112. https://doi.org/10.1016/j.foreco.2005.10.060

Whitney, G. G. (1987). An Ecological History of the Great Lakes Forest of Michigan. *Journal of Ecology*, *75*(3), 667–684. https://doi.org/10.2307/2260198

Wirsenius, S. (2000). Human use of land and organic materials: Modeling the turnover of biomass in the global food system, 1–255.

Wrbka, T., Erb, K.-H., Schulz, N. B., Peterseil, J., Hahn, C., & Haberl, H. (2004). Linking pattern and process in cultural landscapes. An empirical study based on spatially explicit indicators. *Land Use Policy*, *21*(3), 289–306. https://doi.org/10.1016/j.landusepol.2003.10.012

Wright, D. H. (1983). Species-Energy Theory: An Extension of Species-Area Theory. *Oikos*, *41*(3), 496–506. https://doi.org/10.2307/3544109

Zandstra, B. H., & Price, H. C. (1988, January). Yields of Michigan Vegetable Crops (E1565). Retrieved March 11, 2018, from http://msue.anr.msu.edu/resources/yields_of_michigan_vegetable_crops_e1565

Zanne, A. E., Lopez-Gonzalez, G., Coomes, D. A., Ilic, J., Jansen, S., Lewis, S. L., ... Chave, J. (2009). Data from: Towards a worldwide wood economics spectrum. https://doi.org/10.5061/dryad.234

Appendix A: Crop Data Development

This appendix explains how we developed the crop harvest data set and transformed it into NPP values, including the conversion constants we used as inputs for eqn. 4-7. To create the data set, we used data from the USDA National Agricultural Statistics Service (NASS) Quick Stats tool. We downloaded field crop data for Indiana, Michigan, Wisconsin, and Ohio for the years 2007 and 2012. These were the most recent available years in which the USDA's Agricultural Census had been collected.

```
#CSV files used are downloads from the USDA's NASS Quick Stats tool
#packages used
>install.packages("dplyr")
>install.packages("readr")
>library(readr)
>library(plyr)
>library(dplyr)
#select production values from Michigan county yield data, from USDA Agricultural  Census (2012)
>MI_countyYeild <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/AgHarvests/MI/MI_countyYeild.csv")
>MI_countyProduction <- filter(MI_countyYeild, grepl('PRODUCTION', MI_countyYeild$'Data Item'))
>MI_countyProduction <- MI_countyProduction[c(1:21)]
>MI_countyArea <- filter(MI_countyYeild, grepl('ACRES HARVESTED', MI_countyYeild$'Data Item'))
>MI_countyArea <- filter(MI_countyArea, grepl('TOTAL', MI_countyArea$Domain))
>MI_countyArea <- MI_countyArea[c(1:21)]
#Wisconson
>WI_countyYeild <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/AgHarvests/WI/WI_countyYeild.csv")
>WI_countyProduction <- filter(WI_countyYeild, grepl('PRODUCTION', WI_countyYeild$'Data Item'))
>WI_countyArea <- filter(WI_countyYeild, grepl('AREA', WI_countyYeild$'Data Item'))
>WI_countyArea <- filter(WI_countyYeild, grepl('ACRES HARVESTED', WI_countyYeild$'Data Item'))
>WI_countyArea <- filter(WI_countyArea, grepl('TOTAL', WI_countyArea$Domain))
#Indiana
>IN_countyYeild <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/AgHarvests/IN/IN_countyYeild.csv")
>IN_countyProduction <- filter(IN_countyYeild, grepl('PRODUCTION', IN_countyYeild$'Data Item'))
>IN_countyArea <- filter(IN_countyYeild, grepl('AREA', IN_countyYeild$'Data Item'))
>IN_countyArea <- filter(IN_countyYeild, grepl('ACRES HARVESTED', IN_countyYeild$'Data Item'))
>IN_countyArea <- filter(IN_countyArea, grepl('TOTAL', IN_countyArea$Domain))
#Ohio
>OH_countyYeild <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/AgHarvests/OH/OH_countyYeild.csv")
>OH_countyProduction <- filter(OH_countyYeild, grepl('PRODUCTION', OH_countyYeild$'Data Item'))
>OH_countyArea <- filter(OH_countyYeild, grepl('AREA', OH_countyYeild$'Data Item'))
OH_countyArea <- filter(OH_countyYeild, grepl('ACRES HARVESTED', OH_countyYeild$'Data Item'))
OH_countyArea <- filter(OH_countyArea, grepl('TOTAL', OH_countyArea$Domain))
```

```
#merging the state ag data sets and area data sets
>Production12 <- rbind(IN_countyProduction,WI_countyProduction,OH_countyProduction,
MI_countyProduction)
>Area12 <- rbind(IN_countyArea,MI_countyArea, WI_countyArea, OH_countyArea)
```

From the census downloads, we used R to filter out all values that were not of interest to us and values listed as unreported, leaving production values and acreage values. Removal of crops that had unreported data left us with data on about 78% of field crops the census identified as growing in our study region. Unreported data, listed as "(D)" in the NASS downloads, was such due to privacy concerns for that particular grower.

We further filtered the data so that we only examined the data pertaining to field crops for which we had conversion constants that would allow us to convert crop production values into NPP (Table A-1), and collapsed some nuances in the data. Seasonal hay varieties

```
#cleaning data
>Production12 <- filter(Production12, !grepl("CONTRACT", Production12$"Data Item"))
>Production12 <- filter(Production12, !grepl("SURVEY", Production12$Program))
>Production12 <- filter(Production12, !grepl('(D)', Production12$Value))
>Production12$uniqueID <- paste(Production12$`State ANSI`,Production12$County)
>Production12$Commodity[grepl("CORN, SILAGE", Production12$`Data Item`)]<-"CORN, SILAGE"
>Production12$Commodity[grepl('CORN, GRAIN', Production12$`Data Item`)]<-"CORN, GRAIN"
>Production12$Commodity[grepl('HAY', Production12$`Data Item`)]<-"HAY"
>Production12$Commodity[grepl('HAY, ALFALFA', Production12$`Data Item`)]<-"HAY, ALFALFA"
>Production12$Commodity[grepl('HAYLAGE, ALFALFA', Production12$`Data Item`)]<-"HAY, ALFALFA"
>Production12$crop <- ifelse((grepl("CORN, GRAIN", Production12$'Data Item')), "cornGrain",
            ifelse((grepl("CORN, SILAGE", Production12$'Data Item')), "cornSilage",
               ifelse((grepl(", ALFALFA", Production12$'Data Item')), "hayAlfalfa",
                  ifelse((grepl("HAY", Production12$'Data Item')), "hay",
                     ifelse((grepl("SOYBEANS", Production12$'Data Item')), "soy",
                        ifelse((grepl("WHEAT", Production12$'Data Item')), "wheat",
                           ifelse((grepl("SORGHUM", Production12$'Data Item')), "sorghum",
                              ifelse((grepl("BARLEY", Production12$'Data Item')), "barley",
                                 ifelse((grepl("RICE", Production12$'Data Item')), "rice",
                                    ifelse((grepl("SUNFLOWER", Production12$'Data Item')),
"sunflowers",
                                       ifelse((grepl("OATS", Production12$'Data Item')),
"oats",
                                          ifelse((grepl("SUGARBEETS", Production12$'Data
Item')), "sugarbeets",
                                             NA))))))))))))
Production12 <- Production12[c("Year","State","State ANSI","County","uniqueID", "Data Item", "crop",
"Commodity","Value","valueUnits")]
```

and “HAYLAGE” were all collapsed into “HAY,” for instance, and we did not distinguish between irrigated and non-irrigated crops. Filtering out crops for which we did not have conversion constants left us with a data set containing roughly 91 percent of the reported field crop production and acreage data for the four states within our study region. We combined the filtered field crop data with a data table of NPP conversion constants (Table A-1) and ran eq. 4 with the two data sets. We performed a similar process to filter out the area on which each crop was reported as being planted.

```
>cropVariables <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/CSVfiles/Crops&CropConversions/c
ropVariables.csv")
>Production12 <- merge(Production12, cropVariables, by=c("crop","valueUnits"), all.x=TRUE,
all.y=TRUE)
>Production12 <- unique(Production12)
#make the formula
#remove commas from the numbers in the Value column
>Production12$Value <- as.numeric(gsub(",", "", as.character(Production12$Value)))
>attach(Production12)
>for (i in length(Production12$crop)) {
 Value=Production12$Value
 MRY=Production12$kg.Conversion.Factors
 MC=Production12$MC
 C=Production12$C
 HI=Production12$HI
 FAG=Production12$FAG
 i=((Value*MRY*(1-MC)*C)/(HI*FAG))
 Production12$P.kgC.yr <- i
}
detach(Production12)


#filter field crop Area data
>Area12$Commodity[grepl('CORN, SILAGE', Area12$`Data Item`)]<-"CORN, SILAGE"
>Area12$Commodity[grepl('CORN, GRAIN', Area12$`Data Item`)]<-"CORN, GRAIN"
>Area12$Commodity[grepl('HAY', Area12$`Data Item`)]<-"HAY"
>Area12$Commodity[grepl('HAY, ALFALFA', Area12$`Data Item`)]<-"HAY, ALFALFA"
>Area12$Commodity[grepl('HAYLAGE, ALFALFA', Area12$`Data Item`)]<-"HAY, ALFALFA"
>Area12 <- filter(Area12, !grepl('CONTRACT', Area12$'Data Item'))
>Area12 <- filter(Area12, !grepl('SURVEY', Area12$Program))
```

```
>Area12$crop <- ifelse((grepl("CORN, GRAIN", Area12$'Data Item')), "cornGrain",
                ifelse((grepl("CORN, SILAGE", Area12$'Data Item')), "cornSilage",
                    ifelse((grepl(", ALFALFA", Area12$'Data Item')), "hayAlfalfa",
                        ifelse((grepl("HAY", Area12$'Data Item')), "hay",
                            ifelse((grepl("SOYBEANS", Area12$'Data Item')), "soy",
                                ifelse((grepl("WHEAT", Area12$'Data Item')), "wheat",
                                    ifelse((grepl("SORGHUM", Area12$'Data Item')), "sorghum",
                                        ifelse((grepl("BARLEY", Area12$'Data Item')), "barley",
                                            ifelse((grepl("RICE", Area12$'Data Item')), "rice",
                                                ifelse((grepl("SUNFLOWER", Area12$'Data Item')),
"sunflowers",
                                                    ifelse((grepl("OATS", Area12$'Data Item')), "oats",
                                                        ifelse((grepl("SUGARBEETS", Area12$'Data Item')),
"sugarbeets",
                                                            NA))))))))))))
>Area12reference <- Area12[c("Year","State","State ANSI","County","crop", "Commodity","Data
Item","Value","CV (%)")]
>write.csv(Area12reference, "Area12reference.csv")
>Area12 <- Area12[c("Year","State","State ANSI","County","crop","Commodity","Value")]
>Area12$uniqueID <- paste(Area12$`State ANSI`,Area12$County)
>colnames(Area12)[7] <- "Acreage"
>Area12$Acreage <- as.numeric(gsub(",", "", as.character(Area12$Acreage)))
>Area12sum <- Area12 %>% group_by(Year, State, County, uniqueID, Commodity) %>%
summarise(county.harvest.acreage=sum(Acreage))
>Production12 <-
Production12[c("Year","State","County","uniqueID","Commodity","Value","valueUnits","kg.Conversion.
Factors","MC","HI","FAG","C","P.kgC.yr")]
>FieldAgNPP12 <- merge(Production12, Area12sum, by=c("Year","State","County", "uniqueID",
"Commodity"))
>FieldAgNPP12$Area.m2 <- FieldAgNPP12$county.harvest.acreage*4046.86
>write.csv(FieldAgNPP12, "FieldAgNPP12_reference.csv")

##repeat for 2007 census data, then combine two years into one data table
>FieldAgNPP0712 <- rbind(FieldAgNPP07sum,FieldAgNPP12sum)
```

For fruit and vegetable crops, we repeated this process for the top five fruits and top five vegetables grown within in our study region, in terms of how frequently they were recorded in the census (i.e. how many counties reported the crop being grown). Only acres harvested was available for fruit and vegetable data. For this reason, we used eqn. 6 and obtained yield estimates from alternate sources (Ohio Department of Agriculture, 2007, 2012, 2013; USDA's National Agricultural Statistics Service Indiana Field Office, 2007, 2012; USDA's

National Agricultural Statistics Service Michigan Field Office, 2008, 2013; Taylor et al. 2012, 2013).

```
library(readr)
library(dplyr)
> FruitNutCrops <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/CSVfiles/Crops&CropConversions/Fruit
&Nut_2007&2012.csv")
>AreaFN <- filter(FruitNutCrops, grepl('ACRES', FruitNutCrops$'Data Item'))
>AreaFN_cleaned <- filter(AreaFN, !grepl('NOT', AreaFN$'Data Item'))
>AreaFN_cleaned <- filter(AreaFN_cleaned, !grepl('GROWN', AreaFN_cleaned$'Data Item'))
>AreaFN_cleaned <- filter(AreaFN_cleaned, !grepl('NON-BEARING', AreaFN_cleaned$'Data Item'))
>AreaFN_cleaned <- filter(AreaFN_cleaned, grepl('CENSUS', AreaFN_cleaned$Program))
>AreaFN_cleaned <- filter(AreaFN_cleaned, !grepl('(D)', AreaFN_cleaned$Value))
>AreaFN_cleaned <- filter(AreaFN_cleaned, !grepl('(Z)', AreaFN_cleaned$Value))
>AreaFN_cleaned$valueUnits[grepl('ACRES', AreaFN_cleaned$`Data Item`)]<-"acres"
>AreaFN_cleaned$Commodity[grepl('BERRIES, OTHER', AreaFN_cleaned$Commodity)]<-"BERRIES"
>AreaFN_cleaned$Commodity[grepl('NON-CITRUS TOTALS', AreaFN_cleaned$Commodity)]<-
"NONCITRUSTOTALS"
>AreaFN_cleaned$Commodity[grepl('TREE NUT', AreaFN_cleaned$Commodity)]<-"TREENUT"
>AreaFN_cleaned$Commodity[grepl('NON-CITRUS', AreaFN_cleaned$Commodity)]<-"NONCITRUS"
>AreaFN_cleaned$Commodity[grepl('PLUMS & PRUNES', AreaFN_cleaned$Commodity)]<-
"PLUMS&PRUNES"
>AreaFN_cleaned$Commodity[grepl('APPLES', AreaFN_cleaned$Commodity)]<-"APPLES"
>write.csv(AreaFN_cleaned, "Fruit&NutAcreage_0712.csv")

>X2007_Vegetabls <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/AgHarvests/2007_Vegetabls.csv")
>X2012_Vegetables <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/AgHarvests/2012_Vegetables.csv")
>VegCrops <- rbind(X2007_Vegetabls, X2012_Vegetables)
>AreaV <- filter(VegCrops, grepl('ACRES', VegCrops$'Data Item'))
>AreaV_cleaned <- filter(AreaV, !grepl('NOT', AreaV$'Data Item'))
>AreaV_cleaned <- filter(AreaV_cleaned, grepl('CENSUS', AreaV_cleaned$Program))
>AreaV_cleaned <- filter(AreaV_cleaned, !grepl('(D)', AreaV_cleaned$Value))
>AreaV_cleaned <- filter(AreaV_cleaned, !grepl('(Z)', AreaV_cleaned$Value))
>AreaV_cleaned$valueUnits[grepl('ACRES', AreaV_cleaned$`Data Item`)]<-"acres"
>write.csv(AreaV_cleaned, "VegAcreage_0712.csv")
> AreaFV <- rbind(AreaFN_cleaned, AreaV_cleaned)
> library(tools)
>AreaFV[[6]] <- tolower(AreaFV[[6]])
>AreaFV[[6]] <- toTitleCase(AreaFV[[6]])
#fruit, veg, field crop bind
>names(AreaFV)[names(AreaFV) == 'Value'] <- 'Acreage'
>AreaFV$Acreage <- as.numeric(gsub(",", "", as.character(AreaFV$Acreage)))
>AreaFV$uniqueID <- paste(AreaFV$`State ANSI`,AreaFV$County)
#remove spaces in uniqueID column
>AreaFV$uniqueID <- gsub('\\s+', '',AreaFV$uniqueID)
>AreaFV$uniqueID <- tolower(AreaFV$uniqueID)
```

```
#reduce counties to only those within our study region
>GLBasinCounties <- read_csv("/Volumes/EMBARTO/tables/GLBasinCounties.csv") #on flash drive
>GLBasin_AreaFV <- merge(AreaFV, GLBasinCounties, by = c("uniqueID"), all=FALSE)
#add up the acreage of each crop to see which crops have the most acreage in the four states–select the top five fruit and top five vegetables
>ddply(X, c("x"), subset, rank(Commodity)<=14)
>X <- ddply(GLBasin_AreaFV, "Commodity", summarise, x=sum(Acreage))
>Top10GLFV <- subset(GLBasin_AreaFV, GLBasin_AreaFV$Commodity=="SWEET CORN"|GLBasin_AreaFV$Commodity=="POTATOES"|GLBasin_AreaFV$Commodity=="PEAS"|GLBasin_AreaFV$Commodity=="BLUEBERRIES"|GLBasin_AreaFV$Commodity=="GRAPES"|GLBasin_AreaFV$Commodity=="CUCUMBERS"|GLBasin_AreaFV$Commodity=="PEACHES"|GLBasin_AreaFV$Commodity=="CHERRIES"|GLBasin_AreaFV$Commodity=="BEANS"|GLBasin_AreaFV$Commodity=="APPLES", select=c("Year","State","State ANSI","County", "uniqueID", "Commodity","Data Item","Acreage","valueUnits"))
>Top10GLFV <- filter(Top10GLFV, !grepl("FRESH MARKET", Top10GLFV$'Data Item'))
>Top10GLFV$Commodity <- tolower(Top10GLFV$Commodity)
>Top10GLFV$Commodity <- ifelse((grepl("BLUEBERRIES", Top10GLFV$'Data Item')), "BLUEBERRIES",
        ifelse((grepl("PEAS, GREEN",Top10GLFV$'Data Item')), "PEAS, GREEN",
          ifelse((grepl("BEANS, SNAP", Top10GLFV$'Data Item')), "BEANS, SNAP",
            ifelse((grepl("PEACHES", Top10GLFV$'Data Item')), "PEACHES",
              ifelse((grepl("APPLES", Top10GLFV$'Data Item')), "APPLES",
                ifelse((grepl("CHERRIES", Top10GLFV$'Data Item')), "CHERRIES",
                  ifelse((grepl("GRAPES", Top10GLFV$'Data Item')), "GRAPES",
                    ifelse((grepl("CUCUMBERS", Top10GLFV$'Data Item')),
"CUCUMBERS",
                      ifelse((grepl("POTATOES", Top10GLFV$'Data Item')),
"POTATOES",
                        ifelse((grepl("CORN", Top10GLFV$'Data Item')),
"CORN, SWEET",
                          NA))))))))))

write.csv(Top10GLFV, "Top10GLFV_reference.csv")
>write.csv(Top10GLFV, "Top10GLFV_reference.csv")
>attach(Top10GLFV)
#convert acres into meters^2
>Top10GLFV$Area.m2 <- Top10GLFV$Acreage*4046.86
>totAcreCropVar <- Top10GLFV %>% group_by(Year, State, County, uniqueID, Commodity) %>%
>summarise(county.harvest.m2=sum(Area.m2))
>detach(Top10GLFV)
>attach(totAcreCropVar)
#differentiate perennial vs annual crops; perennial root biomass stays in the system, while annual root biomass does not. Thus, only annual crops need to be divided by the FAG (fraction above-ground productivity) variable
#sources for perennial crops
#http://www.michigan.gov/documents/mdard/MI_Ag_Facts__Figures_474011_7.pdf
#fruit and nut trees "have a perennial life cycle"
https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_2_County_Level/Michigan/miappxb.pdf
```

```
>totAcreCropVar$Perennial.Annual <- ifelse((grepl("BLUEBERRIES", totAcreCropVar$Commodity)),
        "perennial",
        ifelse((grepl("APPLES", totAcreCropVar$Commodity)), "perennial",
                    ifelse((grepl("CHERRIES", totAcreCropVar$Commodity)), "perennial",
                        ifelse((grepl("PEACHES", totAcreCropVar$Commodity)), "perennial",
                            ifelse((grepl("GRAPES", totAcreCropVar$Commodity)), "perennial",
                                "annual")))))
>fruitvegVariables <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/CSVfiles/Crops&CropConversions/fr
uitvegVariables2.csv")
fruitvegVariables <- merge(fruitvegVariables, FVcropVariables, by=c("Commodity", "YieldUnits"),
all=TRUE)
>FVNPPsum <- merge(totAcreCropVar, fruitvegVariables, by=c("State","Year","Commodity"), all=TRUE)
#use equation from Monfreda et al (2008) NPP=(EY*DF*C)/(HI*RS); where RS=Fag (ratio of below to
above ground productivity), DF=dry faction or 1-Moisture Content, C=.45 gC/g dry matter, EY=metric
tons of economic yield per unit area
>FVNPPsum$EY.kg.m2 <- FVNPPsum$EY.kg.acre/4046.86
>FVNPPsum$MRY.kg <- FVNPPsum$EY.kg.m2*FVNPPsum$county.harvest.m2
>for (i in length(FVNPPsum$uniqueID)) {
 MRY=FVNPPsum$MRY.kg
 DF=FVNPPsum$DF
 C=FVNPPsum$C
 HI=FVNPPsum$HI
 FAG=FVNPPsum$fAG
 m2=FVNPPsum$county.harvest.m2
 if(isTRUE(FVNPPsum$Perennial.Annual=="annual")){
  i=((MRY * DF * C) / (HI*FAG))
 } else {
  i=((MRY * DF * C) / (HI))
 }
 FVNPPsum$P.kgC.yr <- i
 FVNPPsum$NPP.kgC.yr.m2 <- FVNPPsum$P.kgC.yr/m2
}
```

We combined all crops into a single data table to improve analysis. For some minor crops, such as peas, where we were unable to find production or yield data in our study counties, we either used average yield data collected by Midwest University Extension Services (Zandstra & Price, 1988) or used yields from the same year from neighboring states. We summarized all crop NPP calculations at the county level.

Table A-1: Field crop conversion variables from Lobell et al (2002) and Prince et al (2001). These values convert 93% of the field crop data from the Agricultural census. MC= moisture content (1-Dry Fraction (DF)), HI=harvest index, FAG=fraction above ground biomass (related to root:shoot ratio) and C is percent carbon.

Crop	Production Units	Conversion Factors–Units to Kilograms	MC	HI	FAG	C
cornGrain	bu	25.4	0.11	0.45	0.85	0.45
soy	bu	27.22	0.1	0.4	0.87	0.45
wheat	bu	27.22	0.11	0.4	0.83	0.45
hay	tons	1000	0.15	1	0.53	0.45
hayAlfalfa	tons	1000	0.15	1	0.53	0.45
sorghum	bu	27.22	0.1	0.4	0.8	0.45
sorghum	tons	1000	0.1	0.4	0.8	0.45
barley	bu	27.22	0.12	0.4	0.67	0.45
cornSilage	tons	1000	0.65	1	0.85	0.45
sunflowers	lb	0.4536	0.1	0.35	0.94	0.45
oats	bu	27.22	0.11	0.4	0.71	0.45
sugarbeets	tons	1000	0.85	0.4	0.8	0.45

Table A-2-1: Fruit and vegetable conversion variables from Monfreda et al. (2008). These values convert the top 10 fruit and vegetable crops by area in the study region.

Crop	Yield Units	Conversion Factors–Units to Kilograms	HI	DF	MC	fAG
apples	lbs/acre	0.4536	0.3	0.16	0.84	0.75
cherries	tons/acre	907.1847	0.3	0.14	0.86	-0.25
cherries	lbs/acre	0.4536	0.3	0.14	0.86	0.75
peaches	tons/acre	907.1847	0.3	0.14	0.86	-0.25
peaches	lbs/acre	0.4536	0.3	0.14	0.86	0.75
grapes	tons/acre	907.1847	0.3	0.19	0.81	0.75
blueberries	lbs/acre	0.4536	0.3	0.15	0.85	0.75
peas, green	tons/acre	907.1847	0.45	0.13	0.87	0.85
beans, snap	cwt/acre	50.8023	0.45	0.1	0.9	-0.15
beans, snap	tons/acre	907.1847	0.45	0.1	0.9	0.85
cucumbers	cwt/acre	50.8023	0.45	0.04	0.96	-0.15
cucumbers	tons/acre	907.1847	0.45	0.04	0.96	0.85
potatoes	cwt/acre	50.8023	0.5	0.28	0.72	0.8
corn, sweet	tons/acre	907.1847	0.45	0.13	0.87	0.85
corn, sweet	cwt/acre	50.8023	0.45	0.13	0.87	1.85

Appendix B: Forest Data Development

To develop the forest harvest data input into the NPP_h variable, we downloaded forest harvest data from the US Forest Service Forest Inventory and Analysis EVALIDator tool and used a combination of R and Excel to transform the downloaded values into into NPP values in kg C m^{-2} yr^{-1}.

We used a ratio of forest harvest per acre. The EVALIDator tool defines the numerator we used as the "average annual harvest removals of live trees (trees ≥ 5in DBH), in cubic feet, on forest land," and the denominator is defined as the "area of forestland, in acres." We retrieved this ratio estimate for Indiana, Michigan, Ohio, and Wisconsin in the years 2005-2015.

To find the density (kg C ft^{-3}) of each species commonly found in the Forest Types by which the EVALIDator data was organized (Burrill, 2018), we used data from the Global Wood Density database. The database recorded wood density in g cm^{-3} (oven dry mass/fresh volume). To convert dry mass to C mass we used the proportions 0.521 (softwoods) and 0.491 (hardwoods; Birdsey, 1992).

Table B-1: Partial table showing an example of the combined data from FIA's Database User Guide (Burrill, 2018) and Zanne et al.'s (2009) Global Wood Density Database, along with conversion from g cm^{-3} to kg C ft^{-3}

Binomial	**Common Name**	**Wood density (g cm^{-3})**	**Region**	**Wood Density kg ft^{-3}**	**Kg C ft^{-3}**	**USFS Forest Type**
Pseudotsuga menziesii	Douglas-fir	0.453	NorthAmerica	13	6.7	DF
Pinus ponderosa	ponderosapine	0.38	NorthAmerica	11	5.6	na
Pinus jeffreyi	Jeffreypine	0.37	NorthAmerica	10	5.5	na
Abies lasiocarpa	subalpinefir	0.31	NorthAmerica	8.8	4.6	FirSp
Abies balsamea	balsamfir	0.33	NorthAmerica	9.3	4.9	SpFir
Abies concolor	whitefir	0.37	NorthAmerica	10	5.5	FirSp
Abies magnifica	Californiaredfir	0.36	NorthAmerica	10	5.3	FirSp
Abies amabilis	Pacificsilverfir	0.40	NorthAmerica	11	5.9	FirSp
Abies procera	noblefir	0.37	NorthAmerica	10	5.5	FirSp

We used R to combine the wood density data with the FIA data

```
#GWD from Zanne, A.E., Lopez-Gonzalez, G.*, Coomes, D.A., Ilic, J., Jansen, S., Lewis, S.L., Miller, R.B.,
Swenson, N.G., Wiemann, M.C., and Chave, J. 2009. Global wood density database. Dryad. Identifier:
http://hdl.handle.net/10255/dryad.235.
>GlobalWoodDensityDatabase <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/ForestHarvest/GlobalWoodDensityD
atabase.csv")
#FIA group info from https://www.fia.fs.fed.us/library/database-
documentation/current/ver70/FIADB%20User%20Guide%20P2_7-0_ntc.final.pdf
>FIA_TreeGrpSpp <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/ForestHarvest/FIA_TreeGrpSpp.csv",
col_types = cols(East = col_character(), MAJGRP = col_character(), West = col_character()))
>FIA_Density <- merge(FIA_TreeGrpSpp, GlobalWoodDensityDatabase, by=c("Binomial"))
>FIA_SppGrps_East <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/ForestHarvest/FIA_SppGrps-
East.csv", col_types = cols(East = col_character()))
>FIA_SppGrps_West <-
read_csv("~/Desktop/MAThesis_MasterFolder/Thesis_Data_Sets/ForestHarvest/FIA_SppGrps-
West.csv", col_types = cols(West = col_character()))
>FIA_Density <- merge(FIA_Density, FIA_SppGrps_East, by=c("East"))
>FIA_Density <- merge(FIA_Density, FIA_SppGrps_West, by=c("West"))
>FIA_Density <- FIA_Density[c(-12,-16)]
>names(data)[3]<-"new_name"
>names(FIA_Density)[15] <- "SppGrpNameEast"
>names(FIA_Density)[17] <- "SppGrpNameWest"
>names(FIA_Density)[5] <- "CommonName"
>names(FIA_Density)[16] <- "HWorSW_East"
>names(FIA_Density)[18] <- "HWorSW_West"
>names(FIA_Density)[1] <- "SppGrpCode_W"
>names(FIA_Density)[2] <- "SppGrpCode_E"
>FIA_Density$`Wood density (g/cm^3), oven dry mass/fresh volume` <-
as.numeric(FIA_Density$`Wood density (g/cm^3), oven dry mass/fresh volume`)
#convert g/cm3 to kg/ft3 to match the FIA output
>FIA_Density$WoodDensity.kg.ft3_ovendrymass.freshvol <- FIA_Density$`Wood density (g/cm^3),
oven dry mass/fresh volume`*28.3168466
#multiply density by % C to get carbon weight, numbers based on average softwood (SW) and average
hardwood (HW) percent carbon from http://www.nrs.fs.fed.us/pubs/gtr/gtr_wo059.pdf
>FIA_Density$kgC.ft3 <-
ifelse(FIA_Density$HWorSW_East=="SW",FIA_Density$WoodDensity.kg.ft3_ovendrymass.freshvol*0.
521, FIA_Density$WoodDensity.kg.ft3_ovendrymass.freshvol*0.491)
```

Once the density was calculated in R, we further manipulated the data in Excel using methods developed by Dr. William Currie, Stephanie Hart, and Preeti Rao (W. Currie and P. Rao, personal communication, 2017). We found that the average density of all the forest types that appeared in our study region (6.5 kg C ft$^{-3)}$ was similar to the average density value

(6.4 kg C ft^{-3}; Turner et al., 2007) used in the body of research this study builds upon (W. Currie, S. Hart, and P. Rao, personal communication, 2017). Because of this, we ultimately decided to simply use 6.4 kg C ft^{-3} as our density value to maintain consistency throughout the body of research.

Appendix C: Spatial data development

Potential NPP (NPP0)

For the NPP_0 portion of the equation, we drew on the work done by the Haberl et al (2007). The research group calculated NPP_0 in g C m^{-2} $yr^{-1}$for the entire globe, at 5 arc min (≈10km pixel resolution). The data downloaded as an ASCII grid with no projection. We reprojected it to match the other data in NAD83 Conus Albers, and used zonal statistics to get a table with the summed NPP_0 of each county. We divided the sums by 1000 to get the amount in kg C m^{-2} yr^{-1} and then joined the output sum to the Great Lakes county shapefile to get a total NPP_0 for each county in kg C m^{-2} yr^{-1}.

Actual NPP (NPPact)

To calculate NPP_{act}, we used MODIS NPP data extracted and subsetted using Google Earth Engine (P. Rao, personal communication, 2018). We obtained the MODIS NPP data between 2005 and 2015. For each year, the morning and afternoon data was averaged to account for potential cloud cover at different times of day. We then took the average NPP_{act} value over the entire decade. This average was used to help account for errors in the data, and in the harvest and potential NPP data. The MODIS data is stored at a 500m pixel resolution, in units of kg C m^{-2} yr^{-1}. We resampled the data to 1 km^2 using the resample function in ArcGIS so that it would match the rest of the data. We then used zonal statistics to get the total NPP_{act} per county in kg C m^{-2} yr^{-1}.

Detailed GIS Methods

For manipulating spatial data, we used a combination of ArcGIS and Google Earth Engine (GEE). In ArcGIS, we used a number of transformations. The original data we used can be seen in Table 3. The primary GIS transformations we used were on the NPP_{act} and NPP_0 data, but we also used the program to turn non-spatial NPP_h data into spatial data. The workflow used can be seen in Figure C-1. We used GEE to obtain and parse MODIS data (Running et al. 2015), averaging it first between morning and afternoon satellite passes to reduce error due to cloud cover, and then taking the decadal mean (2005-2015) to use as the NPP_{act} variable. The chosen period matches up with the period over which the USFS forest harvest data was sampled and overlaps the two years from which the crop harvest data was sampled. We downloaded the averaged data from GEE in 500m pixel resolution and NAD83 Conus Albers datum/projection and then uploaded it into ArcGIS for further manipulation.

We used a mask made from the 2011 NLCD to make sure we only had NPP_{act} data for the areas we examined in our study: managed lands, e.g. pasture/hay, crop, and the different landcover types that include tree cover, e.g. deciduous forests, evergreen forests, mixed forests, shrub/scrub (which includes young trees, such as post-harvest regrowth, and stunted trees) and woody wetlands. Exclude landcovers include urban areas, herbaceous wetlands, grasslands, barren ground, open water, and perennial ice. We did the same masking with the re-projected NPP_0 data. We input these masked data layers into the "Zonal Statistics as Table" function to get the mean NPP per county in kg C m^{-2} yr^{-1} and then transferred the data to R. We used R to combine NPP_{act}, NPP_0 and NPP_h using the equation put forward by the Haberl research group: $HANPP=NPP_0-(NPP_{act}–NPP_h)$.

Table C-1: Details of the original and final spatial data for the actual and potential NPP data layers.

Data Layer	NPPact	NPPpot
Original Coordinate System	Sinusoidal	unprojected lat/long
Original Projection	N/A	undefined
Original Datum	N/A	undefined
Original Extent	global	global
Original Resolution	500m	5 arcminutes
Original Units	kgC/m2	gC/m2
Final coordinate system	NAD83 Conus Albers	NAD83 Conus Albers
Final projection	Albers	Albers
Final datum	D_North_American_1983	D_North_American_1983
Final units	kgC/m2	kgC/m2
Feature Class	raster	raster
Source	Google Earth Engine/NASA	Haberl et al (2007)

Table C-2: Original spatial data sources and formats

Project Use	Data Name	Data Type	Resolution	Units	Temporal granularity	Spatial Extent	Coordinate System, Datum, Projection	Citation
NPPact	MYD17A3H: MODIS/ Aqua Net Primary Production Yearly L4 Global 500 m SIN Grid V006	MODIS NPP, morning pass	500mx 500m	kg C m^{-2} yr^{-1}	annual	global	Sinusoidal	Running, S., Mu, Q., Zhao, M. (2015). MYD17A3H MODIS/Aqua Net Primary Production Yearly L4 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MYD17A3H.006
NPPact	MOD17A3H: MODIS/ Terra Net Primary Production Yearly L4 Global 500 m SIN Grid V006	MODIS NPP, afternoon pass	500mx 500m	kg C m^{-2} yr^{-1}	annual	global	Sinusoidal	Running, S., Mu, Q., Zhao, M. (2015). MOD17A3H MODIS/Terra Net Primary Production Yearly L4 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MOD17A3H.006

Table C-2 cont.

Project Use	Data Name	Data Type	Resolution	Units	Temporal granularity	Spatial Extent	Coordinate System, Datum, Projection	Citation
NPPpot	NPP0: net primary production of the potential vegetation	NPP ASCII grid	5' (≈10kmx 10km)	g C m^{-2} yr^{-1}	annual	global	unknown	Helmut Haberl, Karl-Heinz Erb, Fridolin Krausmann, Veronika Gaube, Alberte Bondeau, Christof Plutzar, Somone Gingrich, Wolfgang Lucht and Marina Fischer-Kowalski. 2007. Quantifying and mapping the global human appropriation of net primary production in Earth's terrestrial ecosystem. Proceedings of the National Academy of Sciences of the USA. 104: 12942-12947.
NPPh	Michigan fruit/ vegetable agricultural statistics 2012	non-spatial	NA	NA	annual	state	NA	USDA's National Agricultural Statistics Service Michigan Field Office. (2013). Michigan Agricultural Statistics: 2012-13 (Annual Statistics Bulletin). State of Michigan. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Michigan/Publications/Annual_Statistical_Bulletin/stats13/agstat13.pdf
NPPh	Michigan fruit/vegetable agricultural statistics 2007	non-spatial	NA	NA	annual	state	NA	USDA's National Agricultural Statistics Service Michigan Field Office. (2008). Michigan Agricultural Statistics: 2007-2008 (Annual Statistics Bulletin). State of Michigan. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Michigan/Publications/Annual_Statistical_Bulletin/stats08/agstat-all-08.pdf
NPPh	Ohio fruit/vegetable agricultural statistics 2007	non-spatial	NA	NA	annual	state	NA	Boggs, R. J., O'Brien, D., Hargett, G., Brown, C., & Showalter, S. (2008). 2007 Ohio Agricultural Statistics (Annual Statistics Bulletin) (p. 106). USDA's National Agricultural Statistics Service Ohio Field Office and the Ohio Department of Agriculture.

Table C-2 cont.

Project Use	Data Name	Data Type	Resolution	Units	Temporal granularity	Spatial Extent	Coordinate System, Datum, Projection	Citation
NPPh	Ohio fruit/vegetable agricultural statistics 2007	non-spatial	NA	NA	annual	state	NA	Ohio Department of Agriculture. (2007). 2007 Ohio Department of Agriculture Annual Report and Statistics. Retrieved from http://www.agri.ohio.gov/divs/Admin/Docs/AnnReports/ODA_Comm_AnnRpt_2007.pdf
NPPh	Ohio fruit/vegetable agricultural statistics 2013	non-spatial	NA	NA	annual	state	NA	Ohio Department of Agriculture. (2013). Ohio Department of Agriculture 2013 Annual Report and Statistics. State of Ohio. Retrieved from http://www.agri.ohio.gov/divs/communications/docs/ODA_Comm_AnnRpt_2013.pdf
NPPh	Ohio fruit/vegetable agricultural statistics 2012	non-spatial	NA	NA	annual	state	NA	Ohio Department of Agriculture. (2012). Ohio Department of Agriculture 2012 Annual Report and Statistics. State of Ohio. Retrieved from http://www.agri.ohio.gov/divs/Admin/Docs/AnnReports/ODA_Comm_AnnRpt_2012.pdf
NPPh	Indiana fruit/vegetable statistics 2007	non-spatial	NA	NA	annual	state	NA	USDA's National Agricultural Statistics Service Indiana Field Office. (2007). Indiana 2007-2008 Agricultural Statistics: Crop Summary (Annual Statistics Bulletin) (pp. 31–34). Retrieved from https://www.nass.usda.gov/Statistics_by_State/Indiana/Publications/Annual_Statistical_Bulletin/0708/pg31-34.pdf
NPPh	Indiana fruit/vegetable statistics 2012	non-spatial	NA	NA	annual	state	NA	USDA's National Agricultural Statistics Service Indiana Field Office. (2012). Indiana 2012-2013 Agricultural Statistics: Crop Summary (Annual Statistics Bulletin) (p. 33). Retrieved from https://www.nass.usda.gov/Statistics_by_State/Indiana/Publications/Annual_Statistical_Bulletin/1213/pg33.pdf

Table C-2 cont.

Project Use	Data Name	Data Type	Resolution	Units	Temporal granularity	Spatial Extent	Coordinate System, Datum, Projection	Citation
NPPh	Wisonsin fruit/vegetable statistics 2007	non-spatial	NA	NA	annual	state	NA	Taylor, C., Vale, K., & Woodstock, H. (2013). 2013 Wisconsin Agricultural Statistics (Annual Statistics Bulletin). State of Wisconsin: USDA's National Agricultural Statistics Service Wisconsin Field Office & the Wisconsin Department of Agriculture, Trade, and Consumer Protection. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/Annual_Statistical_Bulletin/bulletin2013_web.pdf
NPPh	Wisonsin fruit/vegetable statistics 2012	non-spatial	NA	NA	annual	state	NA	Taylor, C., Teran, J., Vale, K., & Woodstock, H. (2012). 2012 Wisconsin Agricultural Statistics (Annual Statistics Bulletin). State of Wisconsin: USDA's National Agricultural Statistics Service Wisconsin Field Office & the Wisconsin Department of Agriculture, Trade, and Consumer Protection. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/Annual_Statistical_Bulletin/bulletin2012_web.pdf
NPPh	NASS quickstats	Non-spatial	NA	NA	annual	Month	NA	USDA/NASS QuickStats Ad-hoc Query Tool. (n.d.). Retrieved March 20, 2018, from https://quickstats.nass.usda.gov/

Appendix D: Counties of Conservation Interest

The following tables show a selection of the socioecological characteristics of counties of conservation interest:

Table D-1: Landscape characteristics of counties in the top 50th percentile of mean landscape diversity and the bottom 50th percentile of %HANPP0.

uniqueID	*2010 Population Estimate*	*% Crop Cover*	*% Forest Cover*	*Roads (m m^{-2}))*	*HANPP (kg C m^{-2})*
26alger	9579	2.224044899	87.16623678	0.001419538	0.068798305
26baraga	8855	2.384312277	89.00044255	0.001094513	-0.013258647
26benzie	17508	7.476621193	66.23341377	0.002377849	-0.227360752
26chippewa	38614	9.351378271	72.0589167	0.001217388	0.027462054
26delta	37049	6.404876776	82.23114002	0.001504923	0.136927516
26dickinson	26155	3.490785244	87.07086753	0.00148994	0.273767546
26gogebic	16399	1.434667474	88.45990859	0.001134048	0.208705934
26houghton	36724	4.779193	84.5991257	0.001562639	0.037383309
26iron	11809	2.311708194	87.48619803	0.001251963	0.355074076
26keweenaw	2169	0.032785158	80.61870457	0.000668811	-0.364094058
26lake	11511	5.442682817	84.17680122	0.001743847	0.10473479
26luce	6599	1.083138408	85.78664485	0.001134001	-0.062971157
26mackinac	11107	2.651917171	79.29143649	0.001188226	0.003732538
26manistee	24590	10.70453876	69.71807204	0.002005764	-0.186816108
26marquette	67083	0.94887772	86.2705184	0.001474621	0.135047444
26montmorency	9782	4.667653105	80.7744234	0.001789231	0.163079552
26ontonagon	6776	4.408479175	89.06289971	0.000808267	0.224583958
26oscoda	8603	2.955789395	84.7382594	0.002125425	0.160721695
26schoolcraft	8470	1.485621537	75.42690198	0.000984298	0.150775535
55ashland	16143	5.155742455	86.13839531	0.001218888	0.459966304
55bayfield	15006	6.208098537	84.53008472	0.001617959	0.345707178
55douglas	44134	4.975037347	83.37501172	0.001350423	0.423791413
55florence	4398	4.967081245	86.88511646	0.001297949	0.123572226
55forest	9296	2.859472891	88.01642041	0.001298049	0.269558804
55iron	5924	1.623343871	84.42590605	0.001093594	0.496548735
55menominee	4268	0.476141983	91.63087942	0.001460515	-0.197771056
55oneida	35936	2.79597911	74.02315643	0.001596073	0.390613551
55sawyer	16566	4.338850558	80.67259868	0.001206843	0.332332539
55vilas	21441	1.337178817	69.85573875	0.001825138	-0.041990001

Table D-2: Landscape characteristics of counties in the top 50th percentile of mean connectedness and the bottom 50th percentile of $\%HANPP_0$.

uniqueID	2010 Population Estimate	% Agricultural Landcover	% Forest Cover	Roads (m/m^2)	HANPP (kgC/m^2)
26alger	9579	2.224044899	87.16623678	0.001419538	2450490097
26baraga	8855	2.384312277	89.00044255	0.001094513	2401111656
26benzie	17508	7.476621193	66.23341377	0.002377849	899470472.8
26cheboygan	26067	6.35743655	68.51500191	0.001740167	2064511363
26chippewa	38614	9.351378271	72.0589167	0.001217388	4314871972
26crawford	14054	0.822367191	80.3863929	0.002524708	1459202183
26delta	37049	6.404876776	82.23114002	0.001504923	3075781266
26dickinson	26155	3.490785244	87.07086753	0.00148994	2012441187
26gogebic	16399	1.434667474	88.45990859	0.001134048	2962358928
26houghton	36724	4.779193	84.5991257	0.001562639	2699200472
26iron	11809	2.311708194	87.48619803	0.001251963	3136208981
26kalkaska	17141	4.955259575	73.33107685	0.00200349	1478227304
26keweenaw	2169	0.032785158	80.61870457	0.000668811	1525080323
26lake	11511	5.442682817	84.17680122	0.001743847	1488236642
26luce	6599	1.083138408	85.78664485	0.001134001	2400431913
26mackinac	11107	2.651917171	79.29143649	0.001188226	2818715487
26manistee	24590	10.70453876	69.71807204	0.002005764	1443313004
26marquette	67083	0.94887772	86.2705184	0.001474621	4847832235
26montmorency	9782	4.667653105	80.7744234	0.001789231	1456834912
26ontonagon	6776	4.408479175	89.06289971	0.000808267	3442230165
26oscoda	8603	2.955789395	84.7382594	0.002125425	1480146051
26otsego	24438	7.323082158	72.79033544	0.00201794	1362131379
26roscommon	8470	0.865577647	72.94156249	0.0022866	1501887213
26schoolcraft	16143	1.485621537	75.42690198	0.000984298	3163659036
55ashland	15006	5.155742455	86.13839531	0.001218888	2666929180
55bayfield	44134	6.208098537	84.53008472	0.001617959	3902161001
55douglas	4398	4.975037347	83.37501172	0.001350423	3477360831
55florence	9296	4.967081245	86.88511646	0.001297949	1288483052
55forest	5924	2.859472891	88.01642041	0.001298049	2710289727
55iron	4268	1.623343871	84.42590605	0.001093594	2079041946
55menominee	16566	0.476141983	91.63087942	0.001460515	945096244.3
55oneida	35936	2.79597911	74.02315643	0.001596073	3201025347
55sawyer	16566	4.338850558	80.67259868	0.001206843	3497470078

These counties all fall within the top 50^{th} percentile of either mean landscape diversity or mean local connectedness, and the bottom 50^{th} percentile of $\%HANPP_0$. Thus, we have identified them as sites of high conservation value, given that they have both landscapes supportive of long-term biodiversity protection and, at present, low resource extraction rates. Road data was obtained from (Elvidge et al. 2003) and population data was obtained from the US Census Bureua ("Annual Estimate of the Resident Population: April 1, 2010 to July 1, 2017," 2018).

SEAS SCHOOL FOR ENVIRONMENT AND SUSTAINABILITY
UNIVERSITY OF MICHIGAN

Deep Blue – Library License for Submitted Opus

I hereby grant to the Regents of the University of Michigan the non-exclusive right to retain, reproduce and distribute the deposited work (the Work) in whole or in part, in and from its electronic format. This agreement does not represent a transfer of copyright to the University of Michigan.

The University of Michigan may make and keep more than one copy of the Work for purposes of security, backup, preservation and access, and may migrate the Work to any medium or format for the purpose of preservation and access in the future. The University of Michigan will not make any alteration, other than as allowed by this agreement, to the Work.

I represent and warrant to the University of Michigan that the Work is my original work. I also represent that the Work does not, to the best of my knowledge, infringe or violate any rights of others.

I further represent and warrant that I have obtained all necessary rights to permit the University of Michigan to reproduce and distribute the Work and that any third-party owned content is clearly identified and acknowledged within the Work.

By granting this license, I acknowledge that I have read and agreed to the terms of this agreement and all related University of Michigan and Deep Blue policies.

Opus Author: Eric Bartis ______________________ Date 8/11/2018
Opus Author: ______________________ Date ______
Opus Author: ______________________ Date ______
Opus Author: ______________________ Date ______
Opus Author: ______________________ Date ______
Opus Author: ______________________ Date ______

Opus Type: ☐ Project ☒ Thesis ☐ Practicum

Please submit this signed form along with your final opus and signed opus verification form.