








RESEARCH ARTICLE

10.1029/2021JG006748

Unraveling Forest Complexity: Resource Use Efficiency, Disturbance, and the Structure-Function Relationship

Special Section:

Advances in scaling and modeling of land-atmosphere interactions

Bailey A. Murphy¹ , Jacob A. May² , Brian J. Butterworth^{3,4} , Christian G. Andresen² , and Ankur R. Desai¹ 

¹Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, WI, USA, ²Department of Geography, University of Wisconsin-Madison, Madison, WI, USA, ³Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA, ⁴NOAA Physical Sciences Laboratory, Boulder, CO, USA

Key Points:

- Vertical heterogeneity metrics are the most influential productivity drivers for heterogeneous temperate forests
- The structure-function relationship is mediated by resource use efficiency, and water use efficiency is a strong driver of productivity
- The mechanistic relationship between forest structure and function is dependent upon resolution used to calculate the structural metric

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

B. A. Murphy,
bamurphy5@wisc.edu

Citation:

Murphy, B. A., May, J. A., Butterworth, B. J., Andresen, C. G., & Desai, A. R. (2022). Unraveling forest complexity: Resource use efficiency, disturbance, and the structure-function relationship. *Journal of Geophysical Research: Biogeosciences*, 127, e2021JG006748. <https://doi.org/10.1029/2021JG006748>

Received 7 DEC 2021
Accepted 11 MAY 2022

Abstract Structurally complex forests optimize resources to assimilate carbon more effectively, leading to higher productivity. Information obtained from Light Detection and Ranging (LiDAR)-derived canopy structural complexity (CSC) metrics across spatial scales serves as a powerful indicator of ecosystem-scale functions such as gross primary productivity (GPP). However, our understanding of mechanistic links between forest structure and function, and the impact of disturbance on the relationship, is limited. Here, we paired eddy covariance measurements of carbon and water fluxes from nine forested sites within the 10 × 10 km CHEESEHEAD19 study domain in Northern Wisconsin, USA with drone LiDAR measurements of CSC to establish which CSC metrics were strong drivers of GPP, and tested potential mediators of the relationship. Mechanistic relationships were inspected at five resolutions (0.25, 2, 10, 25, and 50 m) to determine whether relationships persisted with scale. Vertical heterogeneity metrics were the most influential in predicting productivity for forests with a significant degree of heterogeneity in management, forest type, and species composition. CSC metrics included in the structure-function relationship as well as driver strength was dependent on metric calculation resolution. The relationship was mediated by light use efficiency (LUE) and water use efficiency (WUE), with WUE being a stronger mediator and driver of GPP. These findings allow us to improve representation in ecosystem models of how CSC impacts light and water-sensitive processes, and ultimately GPP. Improved models enhance our capacity to accurately simulate forest responses to management, furthering our ability to assess climate mitigation strategies.

Plain Language Summary The way that trees are arranged within a forest impacts the forest's ability to use light and water resources for photosynthesis. Forests that are arranged in more complex ways do a better job of using available resources, and have higher rates of photosynthesis, or productivity. By combining data that describes the complexity of the forest with data that describes how much photosynthesis is occurring, we can better understand which factors impact that relationship, and which types of forest complexity are the most important. We used data from nine temperate forest sites with a long history of management and found that vertical complexity was the most influential, and that the intensity of management had a large impact on the relationship between complexity and productivity. We also found that the relationship was controlled by how efficiently the forest used the available resources, and that the spatial resolution at which the data were examined changed the relationship. These findings will allow us to improve the mathematical models we use to test the impacts of forest management on forest productivity, which will enhance our ability to manage our resources in the face of climate change.

1. Introduction

Recent studies have indicated strong links between forest canopy structural complexity (CSC) and key ecosystem functions such as carbon and water cycling (Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018; Dănescu et al., 2016; Gough et al., 2019; Hardiman et al., 2011; Zhang et al., 2012). Mapping these links is a fundamental aspect of scaling measurements from the leaf to the landscape level and beyond, a preeminent challenge in the field of ecosystem ecology (Bonan, 2008; Fahey et al., 2019), yet the mechanisms underlying these links remain unclear. One approach to addressing this knowledge gap is through the pairing of high-frequency measurements of land-atmosphere exchange with high resolution measurements of CSC taken within the same spatial domain, to isolate mechanistic connections between forest structure and function. Forest CSC characterizes the three-dimensional arrangement of vegetation in a forest and includes variables such as rugosity, vertical complexity

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

index, and mean canopy height (Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018; McElhinny et al., 2005). Taken together, these variables constrain the ability of the forest to assimilate available resources, and thus the capacity for photosynthesis (Ehbrecht et al., 2021). The prevailing theory is that structurally complex forests are better able to optimize incoming light and water resources to assimilate carbon more effectively (Anten, 2016; Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018; Gough et al., 2016; Hardiman et al., 2011). It has been suggested that heterogeneous mixed forests with higher levels of CSC are tied to a heightened ability to capitalize on available resources, in part due to functional trait variability and niche differentiation (Dănescu et al., 2016; Hillebrand et al., 2018; Williams et al., 2016; Zhang et al., 2012).

Studies have shown that integrating information obtained from CSC metrics across spatial scales to describe overall CSC can serve as a powerful indicator of ecosystem-scale functions such as gross primary productivity (GPP), augmenting other commonly measured characteristics including species composition and diversity (Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018; Eitel et al., 2016; Fahey et al., 2019; Gough et al., 2019; Hardiman et al., 2011; Silva Pedro et al., 2017). Identifying not only which CSC variables have the greatest potential to predict GPP, but what potential controls or influential factors of the structure-function relationship might exist is a vital aspect of this effort. As well, relationships between productivity and CSC could provide mechanistic evidence for using these CSC metrics as predictors of forest carbon storage capacity and functionality. To understand the relationship between forest structure and ecosystem functions such as carbon and water cycling, structural complexity must be characterized in a reproducible way that can be easily incorporated into modeling and statistical analysis. Light Detection and Ranging (LiDAR) offers a way to robustly quantify aspects of structural complexity, and helps address critical knowledge gaps regarding our mechanistic understanding of how structure determines function (Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018; Camarretta et al., 2020) by enabling scientists to look at the relationship between forest structure and function through a more quantitative lens. This enhanced understanding results in improved process representation in ecosystem models, advancing our ability to predict ecosystem responses to human management and disturbance, as well as how that response interacts with other components of the earth system. However, as a community we are now faced with the challenge of deciphering which CSC metrics provide novel and relevant information related to ecosystem function, and how those metrics are impacted by spatial scale. Previous work has addressed pertinent issues related to classification and standardization of CSC metrics (Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018; Hardiman, Bohrer, et al., 2013; Hardiman, Gough, et al., 2013; Parker et al., 2004; van Ewijk et al., 2011), but relatively fewer studies have explored the issue of spatial scale in calculating and representing CSC metrics, especially when using aerial-based LiDAR systems.

Multiple LiDAR formats exist, including portable canopy, terrestrial laser scanning, spaceborne, and aerial LiDAR systems, and systems can be discrete-return or continuous-return recording. Each system is subject to different constraints, e.g., terrestrial laser scanning and aerial systems often encounter issues of canopy occlusion (Donager et al., 2021; Hardiman et al., 2018), and differences in how data are collected and used to calculate CSC metrics can impact CSC metric values. Synthesizing data from multiple LiDAR forms is one potential avenue to overcome the limitations associated with individual LiDAR formats (Hardiman et al., 2018) and advance scaling efforts. To successfully do this, an enhanced understanding of the impacts of scale on CSC data collected via different LiDAR formats is required. This study seeks to contribute to this effort by evaluating the mechanistic relationship between forest structure and function at multiple spatial resolutions ranging from 0.25 to 50 m, using Unoccupied Aerial System (UAS) LiDAR-derived CSC metrics to characterize forest structure, and high-frequency eddy covariance (EC) flux data of land-atmosphere CO₂ exchange to quantify ecosystem function. Fine spatial resolutions such as 0.25 and 2 m pixel sizes are included in this study to investigate the relationship between ecosystem function and fine-scale heterogeneity in CSC, which is closely tied to the determination of site microclimates (Ehbrecht et al., 2017), canopy light environments (Tang & Dubayah, 2017), and ultimately patterns of ecosystem functional response to changes in CSC (Smith et al., 2019). The ability of UAS LiDAR systems to capture three-dimensional profiles of stand structure is particularly useful in the mixed temperate forests of the upper Midwest USA, where once even-aged forests are undergoing a transition to more complex systems as they approach advanced stages of successional development following a long history of intensive disturbance (Bogdanovich et al., 2021; Frelich, 1995; Hardiman et al., 2011). In addition, three-dimensional profiles from LiDAR provide important information about the distribution of foliar traits that drive photosynthesis at the leaf level (Chlus et al., 2020; Kamoske et al., 2021).

The impact of human management and disturbance on structure-function relationships varies depending on the severity, frequency, spatial scale, and intensity of the event (Ehbrecht et al., 2021; Ford & Keeton, 2017). Smaller spatial scale and less severe disturbances such as the selective harvest of a given percentage of large trees within a stand tend to increase complexity by creating favorable conditions for understory trees to establish, which results in multilayered canopies (Wisconsin Department of Natural Resources, 2020). This amplified subcanopy growth occurs because disturbance drives a compensatory physiological response to more readily available light, which can also help sustain overall production even in the face of frequent low severity disturbances (Hardiman, Bohrer, et al., 2013; Hardiman, Gough, et al., 2013). In contrast, high intensity and severity disturbances that occur at broad spatial scales such as clearcutting or a high-mortality wildfire event tend to simplify CSC initially, leading to a temporary reduction in productivity (Gough et al., 2007), although stands often recover to predisturbance carbon uptake levels within the 10–20 years following a major disturbance event (Amiro et al., 2010).

Variability in disturbance legacies combined with a primarily mixed broadleaf-conifer forest composition and general landscape heterogeneity result in large variations in both carbon dynamics and stand complexity at the ecosystem scale. As CSC has been shown to be positively correlated with stand production, characterizing the mechanistic relationship between complexity and productivity will enable better representation of the potential impacts of these transitions in successional stage and complexity on carbon sequestration in Midwestern forests (Forrester et al., 2013). The study design of the 2019 Chequamegon Heterogeneous Ecosystem Energy-balance Study Enabled by a High-density Extensive Array of Detectors (CHEESEHEAD19) field experiment provided a unique opportunity to partially control for the influence of variability in climate, edaphic factors, and forest functional types on productivity, allowing for a more representative physiological understanding of the structure-function relationship than has been previously demonstrated.

The objective of this study was to identify mechanistic relationships between forest structure and function, explore potential controls or mediating factors on that relationship, and determine whether or not the structure-function relationship persisted when structural metrics were calculated at a variety of resolutions.

In pursuit of this objective, this project addressed four primary research questions:

1. Which CSC metrics are most influential for the prediction of stand primary productivity in mixed temperate forests with a high degree of heterogeneity and a long history of management?
2. How do management legacies impact these influential CSC metrics, and ultimately stand productivity?
3. Is the mechanistic relationship between forest structure and function direct, or is it mediated by other factors such as resource use efficiency (RUE)?
4. Is the mechanistic relationship between forest structure and function dependent upon the scale of structural metric calculation?

2. Methods

2.1. Experimental Design

This study utilized land-atmosphere exchange and CSC data collected at nine of the CHEESEHEAD19 field campaign study sites in Northern Wisconsin, USA. The CHEESEHEAD19 field campaign spanned from June to October 2019, during which 17 EC flux towers from the National Science Foundation Lower Atmosphere Observing Facility (LAOF) were deployed across a 10 × 10 km study domain. These 17 towers were in addition to the preexisting landscape-level AmeriFlux tall tower US-PFa situated within the study domain (Davis et al., 2003), and two additional temporary EC flux towers supported by Dr. Paul Stoy, bringing the total number of CHEESEHEAD19 EC towers to 20. The temporary CHEESEHEAD19 EC tower sites included 14 forested (primarily mixed conifer and broadleaf) sites, two tussock locations, one grass, and two lake locations. The primary research interests of CHEESEHEAD19 were to explore potential drivers behind the enduring lack of energy balance closure frequently observed over heterogeneous landscapes, and to address persistent challenges associated with upscaling surface energy fluxes (Butterworth et al., 2021). The study period reflects both the summer season land-atmosphere exchange as well as exchanges during the transition of vegetation into senescence. This observational period was chosen to support the energy balance related research interests of CHEESEHEAD19, as it

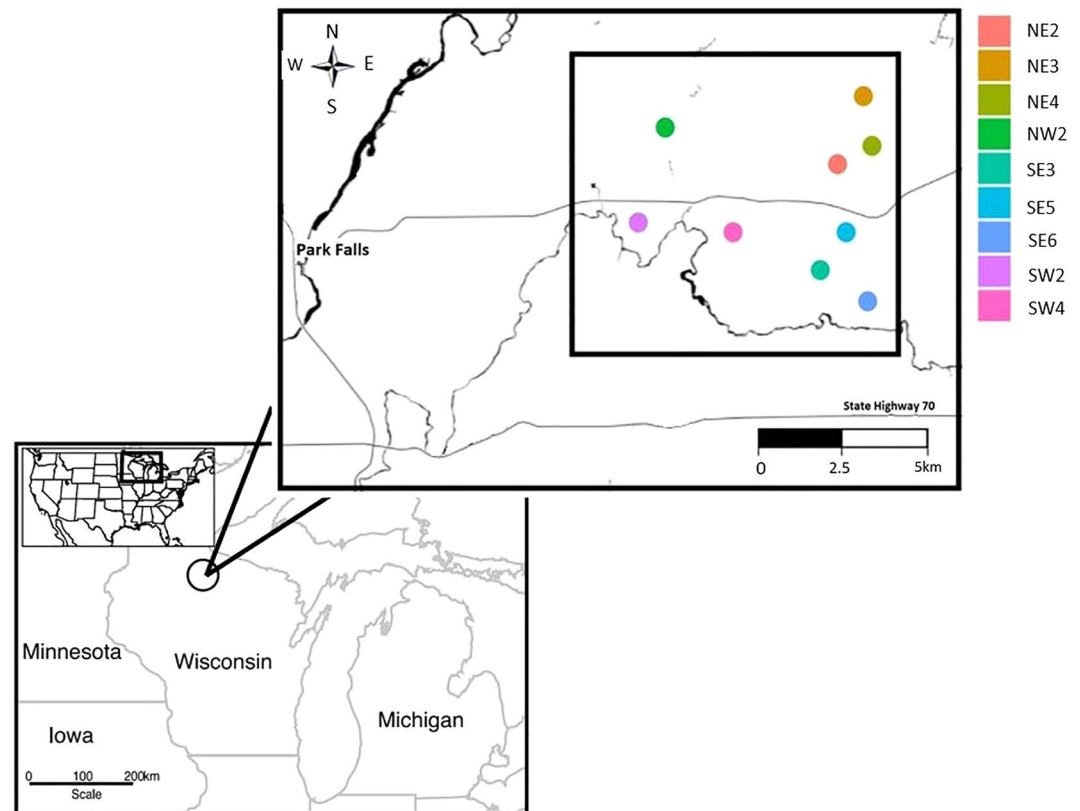


Figure 1. Map depicting the location of the study site within a regional and state context. The black circle on the state map depicts a 60-km radius around the location of the Park Falls, Wisconsin WLEF tall tower. Colored dots represent the nine sites within the 10×10 km CHEESEHEAD19 study domain (represented by the black square) selected for measurement of forest composition.

captures the shift in energy balance from a latent heat flux dominant landscape to a sensible heat flux dominant landscape (Butterworth et al., 2021).

Forest CSC was measured at nine of the forested CHEESEHEAD19 study sites using UAS mounted discrete-return LiDAR (Figure 1). These nine sites were selected given their forested composition and representative forest type, as well as overlap with flux tower footprints. While climatic conditions and topography are shared across the nine sites, the selected sites span a range of successional stages, dominant vegetation types, management histories, and degrees of heterogeneity. Through pairing EC surface-atmosphere carbon and water fluxes with LiDAR-derived forest CSC metrics, mechanistic relationships between forest structure and function could be directly tested.

Mechanistic relationships were explored using best subsets regression for initial model selection and structural equation modeling (SEM), specifically path analysis. Best subsets regression is a variable selection technique where all possible combinations of predictor variables are explored, and a subset of predictive models are selected based on a suite of model fit and performance criteria (Hocking & Leslie, 1967). The top three models for each metric calculation resolution identified using best subsets regression were then evaluated through SEM to isolate the single best-fit model for each resolution. The application of SEM allows for the establishment not only of which CSC metrics are influential in predicting GPP, but the specific strengths, significance, and variability of their predictive power. In addition, SEM allows for the testing of variables that potentially serve as mediators of the relationship between CSC and GPP, through the comparison of reduced and saturated model designs (Fan et al., 2016). This study explored the viability of resource use efficiency (RUE) as a mediator of the structure-function relationship, as previous studies have demonstrated it to be a strong predictor of forest productivity (Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018; Gough et al., 2019). Both water use efficiency (WUE) and light use efficiency (LUE) were used to represent overall stand RUE.

RUE describes how well a forest stand captures and utilizes its available resources to fix carbon dioxide, with greater efficiency typically resulting in higher levels of biomass production (Anderson-Teixeira et al., 2021; Binkley et al., 2004). This paper focuses specifically on light and water as the primary limiting resources controlling the rate of photosynthesis, although other factors including the supply of CO₂, concentration of photosynthetic enzymes such as Rubisco, and availability of catalysts including nitrogen and phosphorous have been explored at length in other studies (Ainsworth & Long, 2005; Hardiman, Bohrer, et al., 2013; Hardiman, Gough, et al., 2013; Tang et al., 2018). In addition, these mechanistic relationships were inspected at different structural metric calculation resolutions to determine whether the relationships persisted with scaling, or if they were simply artifacts of the resolution at which metrics were calculated. Structural metrics were calculated from discrete LiDAR returns (heights) collected at spatial resolutions of 0.25, 2, 10, 25, and 50 m. These five spatial resolutions refer to the size of each pixel contained within the gridded site area, where the gridded area is trimmed to match the average flux footprint of each site. For example, for the majority of the metrics calculated in this study, a metric calculation resolution of 2 m would correspond to a site area broken up into a grid of 2 m × 2 m pixels. CSC metric values are calculated using the LiDAR returns contained within each pixel, and metric values from each pixel are averaged together into a single representative value for each metric at each site. This range of CSC metric calculation resolutions was selected to (a) investigate fine-scale heterogeneity in structural complexity and its impact on ecosystem function, (b) overlap with the resolution of satellite-derived data products from instruments such as the Global Ecosystem Dynamics Investigation (GEDI) full-waveform LiDAR mounted on the International Space Station, and (c) represent a broad enough range in spatial scale to explore the dependency of metric values and derived mechanistic relationships of forest structure and function on spatial scale.

2.2. Site Description

The study area is a 10 × 10 km domain located in the Chequamegon-Nicolet National Forest in Northern Wisconsin. Most of the region is heavily forested and trees are primarily broadleaf but a significant conifer presence exists as well. There is a high degree of heterogeneity representative of a typical midlatitude forest, displaying a diverse array of wetlands, meadows, streams, and lakes in addition to forest cover. Typical homogenous patches of land cover are generally around 20 ha or less (Desai et al., 2015). Heterogeneity is further accentuated by a long history of nonuniform forest management practices including thinning and clearcuts, resulting in increased variability in stand age and structure. Forests in Northern Wisconsin typically have an age distribution centered around “middle age,” or 40–90 years (Birdsey et al., 2014; Wisconsin Department of Natural Resources, 2019). This age pattern is reflective of the fact that the majority of the forested land was logged in the mid-19th to early 20th century to clear land for agricultural purposes (Desai et al., 2008; Gough et al., 2007; Rhemtulla et al., 2009), which was followed by subsequent periods of agricultural land abandonment, reforestation, fire suppression, and intensive timber harvest (Birdsey et al., 2006). In addition to human management, the region experiences natural disturbance due to windstorms, insect invasion, and occasionally fire (Gough et al., 2007). Fires were historically influential during times of land clearing and Euro-American settlement (Rhemtulla et al., 2009), but wind damage has had more consistent impacts on stand structure and carbon storage over time (Schulte & Mladenoff, 2005).

The study domain is of relatively consistent low-grade elevation and human population is minimal. Slight variations in terrain elevation in combination with significant precipitation in all seasons results in a mix of saturated (wetland) and unsaturated (upland) sandy loam soils (Davis et al., 2003). Upland forests comprise roughly 65% of the landscape (Wisconsin Department of Natural Resources, 2019) and broadleaf deciduous tree types include quaking aspen (*Populus tremuloides*), sugar maple (*Acer saccharum*), red maple (*Acer rubrum*), basswood (*Tilia americana*), beech (*Fagus grandifolia*), and several varieties of oak and birch. Coniferous tree varieties include balsam fir (*Abies balsamea*), red, white, and jack pine (*Pinus resinosa*, *Pinus strobus*, *Pinus banksiana*), and white spruce (*Picea glauca*). Wetlands are both forested and unforested and account for ~35% of the land cover (Wisconsin Department of Natural Resources, 2019). Wetland tree species include alder (*Alnus incana*), cedar (*Thuja occidentalis*), tamarack (*Larix laricina*), and black spruce (*Picea mariana*; Davis et al., 2003). The area has a Köppen climate classification of Dfb, and experiences a humid continental climate characterized by warm humid summers and cold snowy winters, with no significant difference in precipitation amount between seasons (Arnfield, 2021).

Table 1
LiDAR Footprint Size, Instrument Installation Heights, and Tree Height Metrics for Each of the Nine Selected Forest Plots

Site	LiDAR footprint (km ²)	EC flux instrument height (m)	Avg. tree height (m)
NE2	0.48	32	14.20
NE3	0.24	32	18.10
NE4	0.18	32	18.70
NW2	0.23	12	8.80
SE3	0.82	32	8.10
SE5	0.22	13	12.40
SE6	0.23	32	10.30
SW2	0.22	30	10.90
SW4	0.82	32	13.50

2.3. Measurements

2.3.1. Flux Towers

Exchanges of carbon, water, and energy between the atmosphere and the land surface were collected at a frequency of 20 Hz using an open-path infrared H₂O and CO₂ gas analyzer (Campbell Scientific EC150) and sonic anemometer to measure three-dimensional wind speed (Campbell Scientific CSAT3AW). In addition to flux-specific instrumentation, the nine selected sites were similarly outfitted with meteorological instruments including slow-response air temperature and humidity sensors (NCAR SHT), barometers (Vaisala PTB210), and 4-component radiometers (Hukseflux NR01). Gas analyzers, sonic anemometers, barometers, and radiometers were all mounted at the top of the EC towers above the local forest canopy, mounting heights are presented in Table 1. Additional instrumentation included tower-mounted air temperature sensors at two levels within the canopy (2 m and midcanopy, which varied by site), and soil sensors (NCAR 4-level Tsoil, Meter EC-5 Qsoil, REBS HFT Gsoil, and Hukseflux TP01 Csoil) buried near the base of each tower in the upper soil profile (0–5 cm). Instrument power was supplied via exchangeable batteries, which occasionally resulted

in minimal data loss due to limited recharging capacity at the field operations base. NR01 radiometer deployment was delayed for sites NW2, NE3, SW2, and SE5, therefore no data exists for approximately the first 25 days of the study period. Radiometer data was filtered for sensor wetness and cleaning periods. Gas analyzers were cleaned 2–3 times during the study, and data was filtered out for periods of significant nighttime dew formation, which resulted in sensor biases.

Turbulent fluxes of carbon, water, and energy were calculated every 30 minutes from high-frequency (20 Hz) EC measurements. Prior to gap filling, a friction velocity (u^*) threshold calculation was performed using the approach outlined in Wutzler et al. (2018), where the u^* threshold is estimated with the moving point test. u^* is a reference wind velocity that represents the shear stress arising through movement across the land surface. Below the u^* threshold, turbulent mixing is weak enough that flux measurements are considered nonrepresentative of the actual flux state, and thus net ecosystem exchange (NEE) flux data are filtered out during those time periods. Gap filling and filtering of flux data were performed using the software REdDyProc (Wutzler et al., 2018). Prior to gap filling, an average of 37% of NEE values were missing across all nine sites, with individual site missing values ranging from 26% (SW2) to 61% (SE5). A visual representation of EC quality control results for the four checks performed (sonic diagnostic, infrared gas analyzer diagnostic, stationarity, and integral turbulence characteristics) is shown in Figure S3 in Supporting Information S1. Missing data occurred to some degree at every site, although the reasons for missing data (equipment malfunction or cleaning, temporary power loss, moisture interference, etc.) varied. GPP was approximated from NEE using the flux partitioning method described in Reichstein et al. (2005) and was calculated using both the nighttime and the light response curve methods for respiration (Reichstein et al., 2012).

2.3.2. Drone-Based LiDAR

To characterize three-dimensional forest structure, we employed a Routescene © discrete-return LiDAR onboard a UAS hexacopter DJI M600 Pro to collect high-density 3D scans (~600 points m⁻²). Over the span of 25–29 June 2019, we surveyed the footprints of the nine selected flux tower sites and areas ranging between 0.25 and 1 km² per site (Table 1) with a flight footprint of approximately 500 × 500 m. Autonomous flights (with a duration of ~20 min each) were programmed using Universal ground Control Software (UgCS) v3.2.113. Flights were performed at a speed of 6 m s⁻¹, 60 m above ground level, and 60 m side distance between parallel flight lines. Raw data were boresight calibrated, filtered and LiDAR heights were *.laz exported using Routescene proprietary software LidarViewer ©. Points within 1-mm radius were filtered and a box range filter centered on the sensor for each scan (scan rate 10 Hz) of 120-m width, 180-m height, and 120-m length was applied, ensuring each flight line would have complete overlap with other flight lines. Random noise was addressed using a statistical outlier removal filter and combined (only for multiple flights per site) in CloudCompare v2.10 (2019).

Table 2
Age, Disturbance, and Management History Data for the Nine Selected Forested Sites

Site	Average stand age	Minimum stand age	Maximum stand age	Management and disturbance history
NE2	56.77	92	18	Selective harvest, thinning, clearcut, blowdown, planting
NE3	71.29	108	41	Clearcut
NE4	108.5	150	76	Thinning, harvest
NW2	44.08	111	7	Blowdown, clearcut, thinning, planting, selective harvest
SE3	42	64	22	Hail damage, clearcut
SE5	55.67	106	7	Clearcut, shelterwood harvest, planting, thinning
SE6	49.5	92	19	Hail damage, blowdown, clearcut, thinning, harvest
SW2	63.5	124	26	Clearcut, planting
SW4	76.27	100	39	Blowdown, clearcut, harvest

2.3.3. Stand Age and Disturbance

Stand age and disturbance history data were obtained from the publicly available United States Department of Agriculture Forest Service Geodata Clearinghouse. All sites had multiple distinct age classes present, representing a range of successional statuses (Figure S1 in Supporting Information S1). The majority of the sites were dominated by stands in the young to middle age classes (Table 2), although regeneration saplings younger than 5 years were not specifically accounted for. The young age class corresponds to the stand initiation and stem exclusion successional stages (Odum, 1969), and the middle age class, defined by Pan et al. (2011) as roughly 40–100 years, corresponds to the understory reinitiation stage. Two sites (NE4 and SW2) contain stands that fall within the old growth successional stage, characterized in the temperate Lake States (Minnesota, Wisconsin, and Michigan) by the presence of long-lived tree species that are at or greater than 120 years of age and exist in an advanced stage of structural development (Frelich, 1995). Forest Inventory Analysis data show that the oldest forests sampled in the temperate Lake States region are between 200 and 210 years old (Birdsey et al., 2014).

Several sites have experienced significant disturbance in the form of clearcutting and harvest (Table 2), with the most recent harvest taking place in 2016 (SE6), and the most recent clear cut occurring in 2013 at stands in sites SE5 and NW2. Harvest is broadly defined here to include selective and shelterwood cuts as well as any harvest that is not stand replacing, whereas a clear cut specifies a stand replacing harvest occurring within the last 50 years. In addition to anthropogenic disturbance, sites SE6 and SE3 experienced substantial hail damage in the year 2000, and large-scale defoliation resulting from Forest Tent Caterpillar infestation occurred across the domain in 2001 (Wisconsin Department of Natural Resources). Blowdown due to wind stress has also been noted at sites SW4, SE6, NW2, and NE2, with the damage being most substantial at site SE6. Neither wildfire nor prescribed burning management activities have been a significant disturbance factor at any of the study plots. Species-specific planting has occurred at sites SW2, SE5, NW2, and NE2. The sites included in this study incorporate a range of management and disturbance histories that are broadly representative of temperate forests in the regional upper Midwest, but the effects of disturbance on forest structure and function explored here are largely qualitative, as neither stand age nor disturbance are expressly controlled for.

2.4. Statistical Analysis

2.4.1. Metric Extraction

LiDAR generated data sets were analyzed using the R programming language (R Core Team 2021; Version 4.0.4) package *lidR* (Roussel et al., 2020). The cloth simulation filter was used to identify ground points (Zhang et al., 2016) and triangulation was used to construct a digital terrain model from the ground points, which was then height-normalized. For each plot, 20 LiDAR metrics were calculated to describe tree height, arrangement, and stand complexity using the R programming language package *forestr* (Atkins, Bohrer, et al., 2018; Atkins,

Fahey, et al., 2018), a complete list of calculated CSC metrics and references is included in Table S1 in Supporting Information S1. *forestr* gives a comprehensive formulation of metrics for characterizing forest canopy CSC and arrangement using either portable canopy LiDAR or terrestrial laser scanning ground-based LiDAR platforms. Several metrics described in the *lidR* library were adapted for an area-based approach with a UAS platform. With the exception of “Rumple” and “VerticalDistMax,” each of the metrics were calculated by creating a raster of the site with a value for each pixel, then finding the average or standard deviation for all pixels within the site. For example, to find the average tree height, a raster of each site was first created where each pixel in the raster was assigned the average height of all the LiDAR returns within the pixel. For this metric LiDAR returns under 0.5 m were removed to exclude most ground points from the calculation. To summarize the data as a single number, the mean of all the pixels in the raster was used. Each raster-based metric was calculated at a resolution of 0.25, 2, 10, 25, and 50 m per pixel to check for resolution dependencies.

Some metrics require additional explanation. Rumple was computed by creating a canopy height model for each site and dividing its area by the projected ground area. VerticalDistMax was computed by finding the vertical distribution of all the points in a site and determining which height bin contained the most points. Vertical bins of 0.5 m and a lower cutoff of 5 m were used to prevent the ground cover and understory from influencing the result. Both of these metrics were calculated on a per site basis instead of a per pixel basis. Leaf area index (LAI) was also calculated using the formulation provided in the *forestr* library (Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018) and compared to LAI field measurements for verification, which showed a high correlation of $R = 0.78$ ($p < 0.05$).

LUE was calculated as the ratio of total daily GPP to total daily incoming photosynthetic photon flux density (PPFD), where PPFD is the incident flux density of photosynthetically active radiation (PAR), or the number of photons incident per unit time on a unit surface (Olson et al., 2004). PPFD is considered a synonym for incident PAR (IPAR; Olson et al., 2004). The exchange of carbon between the forest plots and the atmosphere was measured by the EC towers directly and partitioned into GPP and ecosystem respiration, R_{eco} (Reichstein et al., 2012). The site EC towers were only equipped to measure incoming and outgoing shortwave and longwave radiation as well as net radiation, as opposed to direct measurement of PPFD. Incoming shortwave radiation was converted to PPFD using a fraction of incoming solar irradiance in the photosynthetically active region of 0.50 (Knauer et al., 2018).

WUE describes the amount of carbon fixed per unit of water transpired (De Kauwe et al., 2013), and was calculated here as grams of carbon produced as biomass for every kilogram of water released through evapotranspiration (ET). ET is the sum of evaporation from the land surface and transpiration from vegetation, and is both the key process determining water use in forests (Fisher et al., 2017; Mathias & Thomas, 2021), and the primary process through which the carbon cycle is connected to and maintains the water cycle (Raupach et al., 2005). Since ET was not directly measured by this EC system, it was calculated from measured latent heat flux. Carbon values used in WUE calculations were drawn from EC tower measurements of GPP.

2.4.2. Model Determination

A suite of linear regression models was tested to evaluate the relationships between CSC metrics, RUE, and stand productivity. Nonlinear models were not tested, as previous studies exploring multiple nonlinear model representations have shown that although the relationships may in reality be nonlinear, nonlinear representations repeatedly failed to achieve statistical significance (Gough et al., 2019). The combination of CSC and RUE metrics that best predicted stand GPP was assessed using best subsets model selection. Model fit was evaluated using the Schwarz Bayesian Criterion (SBC), mean square error prediction (MSEP), and adjusted R^2 (R^2_{adj}), where the model with the lowest significant SBC ($p < 0.05$), lowest MSEP, and highest R^2_{adj} was selected as optimal. SBC was used as opposed to Akaike information criterion to account for the presence of multiple predictive variables and a relatively small sample size.

High multicollinearity was a significant problem in determining which CSC metrics were the most robust drivers of GPP. Several CSC metrics had intercorrelation values that exceeded 0.98 and thus were not included in the SEM. This included metrics related to the height at which a given quantile of returned energy was reached relative to the ground, such as the mean of the 25th, 50th, 75th, and 95th quantile of point heights. Variance inflation factors (VIF) were calculated for the best-fit models, and models were classified as having severe multicollinearity if the average VIF was greater than 10. Pearson's correlation coefficients were used to determine the strength

of pairwise interactions between variables for models where severe multicollinearity was a concern to determine which CSC metric was likely driving the observed multicollinearity, and that variable was subsequently removed and the resulting model was reevaluated. Pairwise correlations between variables selected for the final model formulation at each resolution are shown in Figure S2 in Supporting Information S1.

An SEM was used to ascertain the mechanistic relationship between stand productivity and the influential CSC metrics determined through best subsets selection, as well as whether or not the relationship was direct or was mediated by RUE. Path analysis, a subset of SEMs where models are created as a series of regressions to specify causal relationships between variables (Fan et al., 2016), was used to determine possible mediation effects of RUE through the comparison of reduced and saturated models. The reduced model allowed CSC metrics to predict WUE and LUE, and WUE and LUE to then predict GPP. The saturated model allowed for the same prediction pipeline, but CSC metrics could also bypass RUE and directly impact GPP (Figure 4). The existence of mediation in a relationship was determined by comparison of standardized beta coefficients in the uncontrolled path between predictor and response (the total effect) in bivariate models to beta coefficients representing the controlled path between predictor and response variables (the indirect effect) in multivariate models. The relationship between these controlled and uncontrolled pathway beta coefficients determines not only whether or not mediation exists, but also if it is partial or complete mediation. The strength of mediation is determined by the magnitude of the indirect effect.

SEM was performed at each of the five resolutions for LiDAR metric calculation to assess whether or not the mediation effect persisted with resolution changes. Reduced and saturated model fit were assessed using comparative fit index (CFI), standardized root mean square residual (SRMR), and SBC. CFI values closer to one indicate better model fit, so a threshold value of ≥ 0.80 was applied (Hu et al., 1992). SRMR represents the difference between observed and expected variable correlations, and a threshold value of ≤ 0.90 was applied, with a lower value indicating a better model fit. Maximum likelihood estimation was used to determine model fit, and parameter estimates were standardized across all observed variables. Bootstrapping was used to test the significance of indirect effects (and thus the significance of mediation) between CSC variables and productivity through LUE and WUE as well as for estimation of standard errors and bootstrap-based confidence intervals. One thousand draws were performed for each indirect effect evaluated. Significance testing of mediation was performed using the R programming language (R Core Team 2021; Version 4.0.4) package *lavaan* (Rosseel, 2012).

3. Results

3.1. Stand Productivity and RUE

Of the nine CHEESEHEAD19 sites examined here, eight were classified as net carbon sinks, where a negative flux value indicates a net flux of carbon into the ecosystem from the atmosphere. A single site (NE2) was classified as a net carbon source, albeit a minor one, with a net flux of 35 gC m^{-2} released to the atmosphere over the entire measurement period. In addition, at eight out of the nine sites greater variability in daily fluxes was observed for GPP than NEE, with an average variance of 28 gC m^{-2} for GPP compared to 7.8 gC m^{-2} for NEE. Across all sites average daily GPP ranged from 2.6 to 14 gC m^{-2} , and average daily fluxes of NEE ranged between -3.5 and 0.30 gC m^{-2} . Substantial variability was observed in daily total ecosystem respiration (R_{eco}), defined as the sum of both heterotrophic and autotrophic respiration, with an average variance of 20 gC m^{-2} . The highest productivity was observed at sites NE2, SW2, and SW4, with average GPP ranging from 10 to $14 \text{ gC m}^{-2} \text{ day}^{-1}$. Although NE2 has the highest productivity of the nine sites, it also has the highest average daily R_{eco} ($14 \text{ gC m}^{-2} \text{ day}^{-1}$), resulting in its ultimate classification as a slight net carbon source to the atmosphere, as $NEE = R_{eco} - GPP$. The three sites with the lowest productivity are NW2, SE5, and NE4. NW2 has a higher number of clearcuts than all other sites, several stands described as wet conifer bogs, and includes stands ranging in age from 7 to 111 years. SE5 includes a mix of aspen, pine, and upland hardwoods ranging in age from 19 to 92 years. NE4 is a considerably older site, with stand age ranging from 76 to 150 years, and consisting of mixed upland hardwoods, pine, and northern white cedar. Over the course of the June-October observational period, productivity peaked in June to mid-July and decreased into fall as leaves began to senesce, with an average change in GPP across all nine sites of 19 gC m^{-2} . Of the sites, NW2 exhibited the least seasonal change in productivity, with a total difference of only 5.9 gC m^{-2} between the start and end of the study period.

Both LUE and WUE varied between sites, with the across-site average LUE equaling 0.70 gC MJ^{-1} and WUE equaling $4.1 \text{ gC kg H}_2\text{O}^{-1}$. Average LUE variance was 0.19 gC MJ^{-1} and average WUE variance was $1.4 \text{ gC kg H}_2\text{O}^{-1}$. Site NE2 had the highest RUE overall, with a daily LUE of 0.96 gC MJ^{-1} and an WUE of $5.7 \text{ gC kg H}_2\text{O}^{-1}$. NE2 also had the highest variability in RUE, although this variability follows a clear pattern indicating the changes in RUE potentially emerge as a response to changes in temperature or other climatic variables. Site NW2 had the lowest overall RUE, with a daily LUE of 0.33 gC MJ^{-1} and an WUE of $2.9 \text{ gC kg H}_2\text{O}^{-1}$. Site NW2 had the lowest variability in LUE (0.11 gC MJ^{-1}), but the fourth highest variability in WUE ($1.4 \text{ gC kg H}_2\text{O}^{-1}$).

3.2. Classification of Structural Complexity

Of the 20 LiDAR metrics originally calculated, 12 unique metrics related to CSC were shown through best subsets selection to be both influential and statistically significant drivers of stand productivity when combined with RUE variables ($p \leq 0.05$), and thus were included in subsequent SEM testing (Table 3). LUE and WUE were present in all of the best-fit models regardless of spatial resolution, but the specific CSC metrics included in each of the five best-fit models varied depending upon resolution, although several overarching trends stood out. CSC metrics describing vertical heterogeneity were the most prevalent and existed in each of the five final model formulations. VCI_mean (van Ewijk et al., 2011) was the most frequently observed CSC metric, and was included in four of the five models. maxZ_sd, a metric associated with outer canopy heterogeneity, was present in three out of five models, and verticalDistMax, a metric associated with vertical heterogeneity, and LAI_sd were present in two out of five models. The remaining nine CSC metrics each only appeared in a best-fit model formulation a single time, and included rumple, meanZ_sd, sdZ_sd, LAI_sd, maxZ_mean, sdZ_mean, gap_fraction, canopy_ratio_mean, and LAI_mean. Of these nine CSC metrics, four are related to vertical heterogeneity (sdZ_sd, sdZ_mean, canopy_ratio_mean, and meanZ_sd), one to outer canopy heterogeneity (rumple), one is a measure of mean outer canopy tree height (maxZ_mean), two describe the area and density of vegetation distribution (LAI_sd and LAI_mean), and one describes the degree of canopy cover and openness (gap_fraction). Of these nine CSC metrics, three are only present in the 25-m resolution model and two are only present in the 50-m resolution model indicating that the larger resolution models have a greater departure from the other best-fit models. Fit metric ranges for the single best-fit model at each resolution displayed no significant differences by resolution. Average R^2_{adj} was 0.32 with a range of 0.05, average BIC was 4,418 with a range of 51, and average MSE was $16.40 \text{ gC m}^{-2} \text{ day}^{-1}$ with a range of $1.20 \text{ gC m}^{-2} \text{ day}^{-1}$. This suggests that CSC metric's viability as a driver of GPP is not restricted to fine or coarse resolutions.

All 12 of the site averaged CSC metrics varied depending upon the pixel size used in metric calculation. The majority of CSC metrics decreased in value as resolution became coarser, however, 5 of the 12 metrics (*MOCH*, *Vert_meanStd*, *VCI_AVG*, Θ , and *Canopy Ratio_AVG*) displayed the opposite trend (Figure 5). The observed shifts in metric values with changing resolution indicated that the overall mechanistic relationships between CSC metrics and productivity could be resolution dependent. The greatest differences with shifting resolution were observed in *VAI_maxheightmean* and *VCI_AVG*. *VAI_maxheightmean* decreased with decreasing spatial resolution, with values being reduced to 25–30% of the 0.25-m resolution value by the time a 10-m resolution was reached, and all sites had the same value (5 m) upon reaching the 25-m resolution. *VCI_AVG* increased with decreasing spatial resolution, with values increasing on average by 20% with each decrease in resolution, although the difference between 10 and 25 m was less pronounced, with an average difference of 5%, and metric values are relatively stable by 50-m resolution.

Vertical heterogeneity metrics describing the layering of the canopy such as σ_H , *Canopy Ratio_AVG* (Schneider et al., 2017), and *Vert_meanStd* generally had higher values at sites with distinct multilayered canopies such as NE2 and SE5, and lower values at sites with a more consistent single layered canopy, such as site NE4. Variability in *Vert_meanStd* values generally increased with decreasing resolution, with an increase in spread between sites of 45% from 0.25-m to 50-m resolution, whereas site-to-site variability decreased by ~46% with decreasing resolution for σ_H .

CSC metrics *VCI_AVG* and vertical variability (*Vert_DSDD*) offer insight as to the degree of variability in the distribution of vegetation within each vertical column. *Vert_DSDD* is similar to the metric “StdStd” described in Atkins, Bohrer, et al. (2018); Atkins, Fahey, et al. (2018), but it represents the standard deviation column variability of tree height, as opposed to mean leaf height. The highest *VCI_AVG* values were observed at site NE2 (9% higher than the average of the other eight sites, at a resolution of 0.25 m), and the lowest values were typically seen

at sites SE3 and NW2, depending upon spatial resolution. Variability between sites increased with decreasing resolution, largely in part to a widening spread between SE3 and NW2 and the remaining seven sites. The highest $Vert_{SDSD}$ values, indicating a less uniform vertical distribution of vegetation, were measured at sites with multi-layered canopies and multiple distinct age classes present, such as SW4 and NW2, which include stands ranging in age from 7 to 110 years.

The final CSC metric addressing vertical heterogeneity is VerticalDistMax. VerticalDistMax is equivalent to the variable “mean height of vegetation area index maximum” ($VAI_{maxheightmean}$) described in the *forestr* package (previously referred to as VAI_{mode} in Atkins, Bohrer, et al. (2018); Atkins, Fahey, et al. (2018)). For resolutions 0.25 and 2 m, the highest $VAI_{maxheightmean}$ values are seen at sites NE3 and NE4, and the lowest values are seen at sites NW2 and SE3. For 10-m resolution only two values of $VAI_{maxheightmean}$ exist, 5 and 15 m, with three sites (NE2, NE3, and SW4) having values of 15 m and the remaining sites having values of 5 m. By 25 m model resolution, all sites have the same $VAI_{maxheightmean}$ value of 5 m, indicating that the metric calculation resolution has a significant impact on $VAI_{maxheightmean}$.

Influential CSC metrics representing outer canopy complexity include rumple and top rugosity (R_T). Rumble is defined as the ratio of the outer canopy surface area to the underlying ground surface area (Parker et al., 2004), where a higher value corresponds to a more complex canopy (Kane et al., 2010). Average rumple values were significantly impacted by metric calculation resolution, and substantially decreased at resolutions coarser than 0.25 m, indicating that at coarser resolutions the outer canopy surface appears artificially smoothed. Variability between sites also decreased with decreasing resolution, and at resolutions coarser than 2 m, differences in rumple values between sites were negligible. For context, in a Douglas-fir and western hemlock dominated 500+-year-old growth forest in Southern Washington (USA) with an extremely high level of outer canopy complexity, rumple values of 12 m were reported (Parker et al., 2004). R_T refers to the standard deviation of LiDAR column maximum return heights (Atkins, Bohrer, et al., 2018; Atkins, Fahey, et al., 2018). The highest values of R_T were observed at NE3 and NE4, with an average range of 3.3–8.2 m across all five resolutions.

Mean outer canopy tree height ($MOCH$) serves as a simple measure of vertical stand structure, by describing the maximum tree height averaged across all present species in a given stand. $MOCH$ values increase with decreasing resolution, presumably because taller trees dominate and skew the average when a larger field of view is utilized. LAI is the ratio of the (one-sided) total leaf area per unit of ground area, and describes the amount of leaf tissue exposed to ambient light in the forest canopy. The highest values of LAI_{AVG} were observed at sites NE3 and NE4, both among the oldest sites, and the lowest values were seen at sites SE3 and SE5, both fairly young aspen sites, with an overall range of 1.0–3.7 m across all five resolutions. LAI_{SD} describes the standard deviation, or the variability in LAI, and offers insight into how photosynthetic tissues are distributed in the forest canopy. The highest values were observed at site SW4, and the lowest values were seen at sites NE2 and NE3, with a total range of 0.34–1.4. LAI_{SD} values generally decrease with decreasing resolution, with a reduction in variability between sites (decrease in variability of 36% from 0.25-m to 25-m resolution). Values of Θ showed the greatest variability between sites at a resolution of 0.25 m, and at resolutions of 10 m and greater differences between sites became indistinguishable.

3.3. Structural Equation Modeling

Comparison of SEM models showed that the reduced model, where CSC metrics were restricted to influencing GPP through RUE as opposed to exerting direct influence over GPP, had a better overall fit than the fully saturated model. In other words, LUE and WUE actively mediate the mechanistic relationship between CSC variables and GPP, and changes in CSC result in changes in RUE and ultimately in GPP.

Across all five models, 38 mediation relationships were tested in total; 19 WUE mediated relationships and 19 LUE mediated relationships. The strength of mediation is determined by the magnitude of the indirect effect, where the indirect effect is calculated as the product of standardized beta coefficients representing the path between the predictor and mediation variables and the path between mediation variables and the response variable. Fourteen of the 19 WUE mediated relationships were significant (Figure 6), with all cases being partial mediation, complete mediation was not observed for either WUE or LUE. WUE as a mediator between GPP and the metrics $VAI_{maxheightmean}$ (present in 0.25-m and 10-m resolution models), LAI_{SD} (present in the 25-m resolution model), Θ , and $Canopy Ratio_{AVG}$ (both only present in the 50-m resolution model) were never shown

to be significant. Mediation strength, characterized by the magnitude of the indirect effect of a given CSC metric on GPP through WUE as a mediator was 0.10 on average, with a range of 0.14. Eight of the 19 LUE mediated relationships were significant (Figure 6), and metrics directly tied to light interception such as those related to LAI (LAI_{SD} and LAI_{AVG}) were always significantly mediated by LUE. LUE significantly mediated relationships between GPP and VCI_{AVG} , $MOCH$, R_T , and $Vert_{meanStd}$ as well, but the significance of mediation was not always consistent when a given CSC metric was present in different resolution models. LUE never significantly mediated relationships between GPP and $VAI_{maxheightmean}$, rumple, σ_H , $Vert_{SDSD}$, Θ , or $CanopyRatio_{AVG}$, regardless of spatial resolution. Mediation strength of LUE on the relationship between a given CSC metric and GPP was 0.03 on average, with a range of 0.08. WUE was shown to be a substantially stronger mediator between CSC and GPP than LUE, with a standardized mediation strength 290% larger than that of LUE when averaged across all nine plots. Averaged across all sites, the correlation between daily WUE and daily LUE was 0.40.

Mediation analysis by resolution showed differing trends between WUE and LUE as mediators. The significance of WUE as a mediator did not appear to be resolution dependent, whereas the significance of LUE as a mediator did appear to be dependent upon the spatial resolution of the model in question. For example, LUE was a significant mediator of R_T in the 25-m resolution model, but not in the 10-m resolution model. The presence of LUE as a significant mediator was more prevalent at coarser spatial resolutions (10 and 25 m) than at finer resolutions (0.25 and 2 m), but LUE only significantly mediated one of the five predictive variables included in the 50-m resolution best-fit model (VCI_{AVG}). Cases where the structure-function relationship was mediated by both WUE and LUE were observed more frequently at coarser resolutions than at finer resolutions, with the exception of the 50-m resolution model. In summary, for CSC metrics that experienced mediation (all but $VAI_{maxheightmean}$) the presence of a mediating factor in the overarching relationship between forest structure and function was consistent regardless of CSC metric calculation resolution, but which individual relationships were significantly mediated changed with resolution shifts when LUE was the mediating variable in question. Mediation effects were the least pronounced at a metric calculation resolution of 50 m, but this is potentially influenced by the higher prevalence of edge effects at such a coarse resolution, and the associated influence on metric uncertainty. Due to the variety of measurement units involved, beta coefficients were standardized to facilitate comparison and outliers were removed. Standardized beta coefficients show that at a resolution of 0.25 m, VCI_{AVG} and rumple were the strongest drivers of GPP ($\beta = 0.33$, $\beta = 0.11$), at 2 m VCI_{AVG} was the strongest driver of GPP ($\beta = 0.35$) followed by σ_H ($\beta = 0.16$), at 10 m VCI_{AVG} and H_{AVG} were the strongest drivers of GPP ($\beta = 0.33$, $\beta = 0.16$), at 25-m spatial resolution R_T and $Vert_{meanStd}$ were the strongest drivers ($\beta = 0.22$, $\beta = 0.18$), and at 50 m VCI_{AVG} and Θ were the strongest drivers ($\beta = 0.30$, $\beta = -0.11$). Additionally, all CSC metrics were stronger drivers of WUE than of LUE.

4. Discussion

Our findings support the emerging consensus that a positive mechanistic relationship exists between CSC and productivity in mixed temperate forests (Gough et al., 2016, 2019), but suggest that this is a multifaceted relationship impacted by additional factors such as species diversity and management history. As well, we found that this relationship is not direct but rather is mediated by the effective acquisition and assimilation of both light and water resources, and that RUE generally is enhanced by increasing CSC. Furthermore, we show that in a heterogeneous mixed temperate forest subject to disturbance, metrics describing the vertical profile of heterogeneity are the strongest drivers of productivity, as opposed to CSC metrics that are constrained to the outer canopy. Through analysis of the structure-function relationship at five structural metric calculation resolutions ranging from 0.25 to 50 m, we demonstrate that the scale of metric calculation has a significant impact on the metric values themselves, and thus on which CSC metrics are ultimately included in predictive models of productivity. We showed that shifting the spatial resolution also changes the dynamics of the relationship between RUE and CSC. Lastly, it was established that even in a study domain where sites have shared climatic and environmental conditions, differences in management and disturbance history as well as species diversity result in substantial variability in land-atmosphere exchanges of CO_2 . This is likely due to changes in forest composition and trait diversity in response to disturbance.

4.1. Structural Complexity

VCI_{AVG} was the most frequently observed CSC metric in the models, and consistently proved to be the most robust driver of both RUE and GPP in models where it was present, irrespective of spatial resolution. However,

CSC metrics related to outer canopy heterogeneity such as R_T and rumple were also imperative. Four out of the five best-fit models included CSC metrics related to both vertical and outer canopy heterogeneity, although vertical metrics were more prevalent in all cases. VCI_{AVG} describes how the vertical distributions of LiDAR returns differ from a uniform distribution, which is representative of the overall evenness of the vertical distribution of vegetation (Kane et al., 2010a, 2010b; van Ewijk et al., 2011). In an example presented by van Ewijk et al. (2011), a low VCI_{AVG} could correspond to the stand initiation stage, where the majority of point returns are congregated in the lowest vertical bins, whereas a mid to high VCI_{AVG} could correspond to a stand in the midst of understory reinitiation or even a transition into old growth, where vegetation is distributed between multiple height bins. The VCI_{AVG} values observed within the study domain are consistent with the relative dominance of stands in the young to midage classes.

The CSC metric $VAI_{maxheightmean}$ conveys important information about biomass allocation patterns. Models that did not contain $VAI_{maxheightmean}$ did contain CSC metrics related to LAI, suggesting that incorporating a variable that accounts for the complexity in arrangement of vegetative tissues is essential when describing a stand's ability to absorb incoming light. VAI is similar to the more commonly used LAI, but vegetative tissues include branches and stems in addition to photosynthesizing leaves (Scheuermann et al., 2018). However, it is worth noting that several studies have shown that the influence of LAI on production saturates in importance over time, but the same trend has not been observed in VAI (Hardiman et al., 2011), potentially making it a more reliable metric overall when describing the area-related distribution of vegetative tissues. σ_H is the standard deviation of the mean height of lidar returns for raster pixels, and builds on the canopy layering information provided by $VAI_{maxheightmean}$ by representing the variability associated with the height of greatest leaf density. High values (corresponding to a multilayered canopy) were observed at sites with a variety of age classes present, where harvest practices have resulted in patches with unique canopy features, such as site NE2 (Figure 3).

$Vert_{meanStd}$ is a reliable indicator of the spread between distinct canopy layers, high values were observed at sites such as NE2, which includes a dense canopy between 5 and 10 m tall with an additional canopy around 25 m tall, and a fairly sparse degree of vegetation between the two canopies (Figure 3). Pairing this metric with $Vert_{SDSD}$ illustrates the variability in vertical forest profiles, and offers insight into the arrangement of the understory. For example, high values of $Vert_{SDSD}$ were observed at sites with dense nonuniform understories, such as site SW4. In addition to conveying information about forest successional stage when combined with species information, $MOCH$ is important to consider when interpreting the significance of observed rumple values (Kane et al., 2010b), as rumple generally increases with increasing tree height. At first glance sites NE4 and NW2 could be classified as having similar levels of complexity, with rumple values of 3.4 and 3.5 m, respectively, at 0.25-m resolution. However, the large differences in $MOCH$ between the sites (19 m versus 8.8 m) draws attention to the fact that the variance in complexity between the two sites is more pronounced, as a similar rumple value for a stand with less than half the $MOCH$ of NE4 indicates that NW2 has a higher degree of CSC than is present at NE4.

The prevalence of vertical heterogeneity metrics focused on canopy layering and vegetation distribution in the models explored here further supports recent findings indicating that vertical complexity is a strong driver of productivity in mixed temperate systems (Fahey et al., 2019), that it plays an important role in determining seasonal dynamics in forest productivity (Smith et al., 2019; Tang & Dubayah, 2017) and emphasizes the role of vertical variation in driving biomass growth (Stark et al., 2012). All 12 influential CSC metrics explored here were sensitive to changes in metric calculation resolution, highlighting the need for consistency in the spatial resolution at which CSC metrics are calculated, and for the disclosure of metric calculation resolutions when reporting CSC metric values and interpreting the significance of findings. For most CSC metrics, values decreased as resolution became coarser (with $MOCH$, $Vert_{meanStd}$, and VCI_{AVG} as exceptions), as did variability between sites. Moreover, differences between sites became indistinguishable for rumple, $VAI_{maxheightmean}$, and Θ at resolutions coarser than 10 m. This suggests that for research questions centered around discerning differences in CSC between sites and the potential impacts of those differences on ecosystem function, a finer resolution should be used for CSC metric calculation. However, which sites are classified as most or least structurally complex overall is relatively consistent regardless of metric calculation resolution. Sites SE5 and NE2 consistently rank as the sites with the highest complexity, and sites SE3 and NW2 dependably rank as the sites with the lowest complexity. For some sites, such as NE3 and NE4, the comparative complexity ranking differs depending on which metric is being examined, e.g., both sites have very high complexity rankings in metrics LAI_{AVG} ,

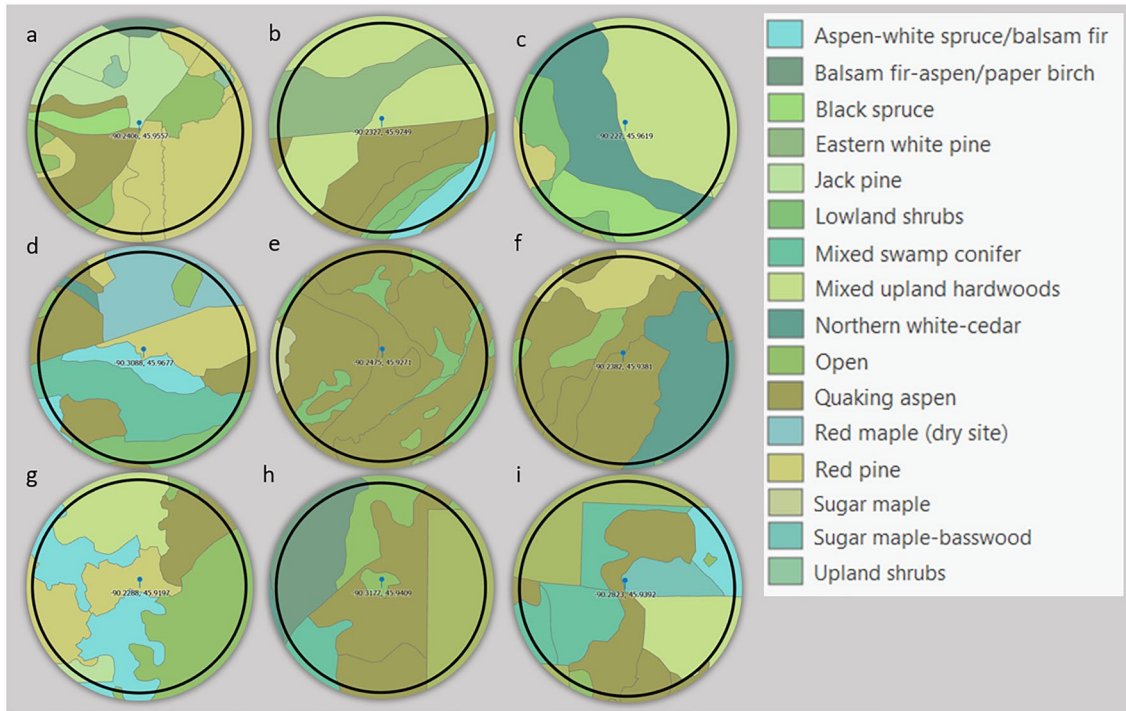


Figure 2. Vegetation coverage at each of the nine forested sites: (a) NE2, (b) NE3, (c) NE4, (d) NW2, (e) SE3, (f) SE5, (g) SE6, (h) SW2, and (i) SW4. Coverage is segmented by both vegetation type and stand age.

$VAI_{maxheightmeans}$, and $MOCH$ regardless of resolution, but consistently rank low in metrics R_T , LAI_{SD} , rumple, and $Vert_{SDSD}$.

Ultimately CSC cannot be encapsulated by a single metric, and a select set of metrics will provide a more comprehensive representation. For instance, pairing a variable like σ_H that offers insight as to whether a canopy is single or multilayered with a variable like LAI_{AVG} that describes the density and arrangement of photosynthetic tissues will reveal more about a stand's potential productivity than either variable in isolation could. However, which metrics should be included in predictive models of productivity is not a one size fits all situation, as shown here it is contingent upon spatial resolution.

4.2. The Structure-Function Relationship

Here, we showed that a positive mechanistic relationship exists between CSC and forest productivity in mixed temperate forests, and that CSC metrics which describe the vertical profile of heterogeneity are better predictors of GPP than metrics that are limited to the outer canopy alone. This is potentially due to vertical complexity metrics providing greater information content in terms of describing a forest's successional stage and ability to capture light as it moves beyond the outer surface of the canopy and penetrates into the forest below (Zimble et al., 2003). As early successional species overtake forest gaps created by disturbance to establish multicanopied stands, the more biodiverse forest with greater structural complexity and range of shade tolerances will make the forest more resource efficient under variable light conditions, increasing net carbon uptake (Hardiman et al., 2011; Hardiman, Bohrer, et al., 2013; Hardiman, Gough, et al., 2013; Hooper et al., 2005). For example, NE2, which has the highest GPP, WUE, and LUE, also exhibits high levels of CSC across the majority of the metrics evaluated. NE2 is predominantly pine, with aspen and paper birch intermixed (Figure 2). Due to a history of timber harvest and replanting (Table 2) there is a significant secondary pine canopy (Figure 3) with an average age of 22 years. This multilayered canopy is captured in the second highest values of VCI_{AVG} observed across all nine sites (at a spatial resolution of 0.25 m $VCI_{AVG} = 0.35$, 10% higher than the following seven sites), while $MOCH$ and rumple were also comparatively high, at 10% higher than average and 4.4% higher than average, respectively.

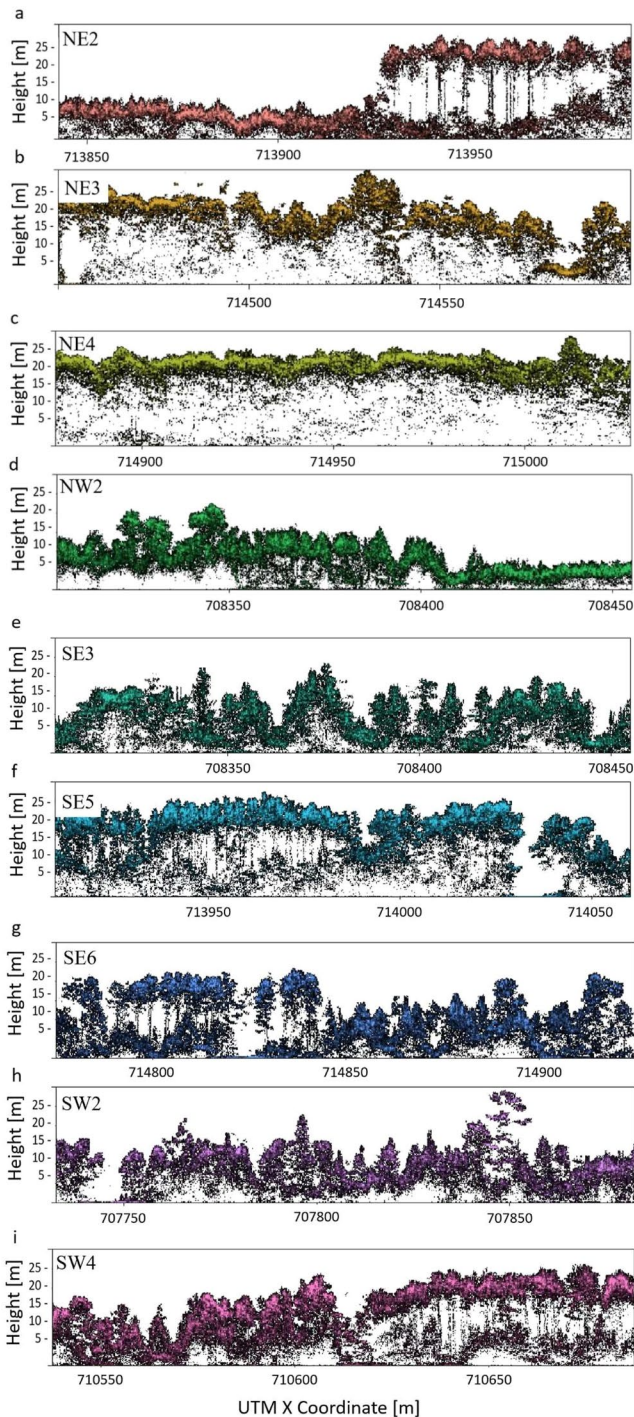


Figure 3. Light Detection and Ranging (LiDAR) point return 150 m transect images for the nine forested sites: (a) NE2, (b) NE3, (c) NE4, (d) NW2, (e) SE3, (f) SE5, (g) SE6, (h) SW2, and (i) SW4. Color saturation represents the relative number of returns at a height interval. The *x*-axis represents longitudinal coordinates in meters, expressed at 50-m intervals, and the *y*-axis is height above ground in meters.

Mediation of the structure-function relationship by RUE existed at all five CSC metric calculation resolutions, but was the least pronounced at a resolution of 50 m, where mediation was present in fewer than half of the mediation pathways that were tested (Figure 6). At a resolution of 50 m, CSC metrics are calculated within a 2,500 m² pixel. With such a large pixel size, edge effects are more pronounced, and can affect metric values, which can ultimately impact the mechanistic relationships that are derived using those values. Pixels located at the edge of a site contain a portion of the available LiDAR data, but a portion of the total pixel area lies outside of the available LiDAR data range, meaning that data collected in these edge pixels is essentially weighted higher than data collected across the rest of the site, because the pixels contain less LiDAR data but still cover the same total area. These edge effects manifest as uncertainty in CSC metric values. While negligible at fine resolutions such as 2 or 10 m, at larger scales the added uncertainty is amplified.

SEM highlighted WUE as a considerably stronger driver of GPP than LUE, but it is important to pause here and consider that the temperate mixed forests of Northern Wisconsin are not water limited ecosystems, and previous studies have shown that stand-scale productivity is predominantly a function of the capacity to harvest light and fix carbon (Reich, 2012), so why does WUE show up as highly influential when predicting GPP? The answer lies primarily in the relationship between WUE and LUE. The tiny stomata covering the leaf surface exist in a constant tradeoff between opening and sacrificing water for the chance to take up CO₂, both of which are necessary ingredients for photosynthesis (Monteith, 1965). Regardless of available light, when plants are water stressed, stomata close in an attempt to conserve existing resources, at the cost of reducing CO₂ uptake and thus photosynthetic capacity (Hatfield & Dold, 2019; Kukul & Irmak, 2020). However, when a plant has a steady supply of water, stomata will more readily open and a greater amount of atmospheric CO₂ can be fixed per unit of incident light (Binkley et al., 2004). A recent study by Ehbrecht et al. (2021) examining climatic controls on CSC at the global scale found that CSC was strongly correlated with water availability across all biomes examined, and that the relationship between water availability and use and CSC can be tied to mechanisms determining tree size. This is because water availability effectively controls functional diversity and shade tolerance as well as tree size following the hydrological limitation hypothesis (Ehbrecht et al., 2021). Shade tolerant trees are found in greater abundance in systems where growth is not limited by factors other than light, such as the nonwater limited systems present in Northern Wisconsin. All three of these factors (functional diversity, shade tolerance, tree size) contribute to CSC (Thom et al., 2021). In addition, a recent study by Smith et al. (2019) showed that vertical heterogeneity in particular plays an important role in modulating seasonal responses to water availability. Although the Smith et al. (2019) study took place in a tropical forest, comparable relationships may also exist in temperate forests such as our site, which experiences a warm, humid growing season and is not typically water limited.

However, the importance of the relationship between CSC and LUE cannot be understated, as it shows that the functional diversity driven by complexity is able to better capitalize on available resources (Penone et al., 2019; Williams et al., 2016). As well, although this study was limited in duration, other studies such as the Zhang et al. (2012) global meta-analysis of diversity productivity relationships showed that almost 30% of the variation in productivity between monocultures and polycultures was explained by heterogeneity

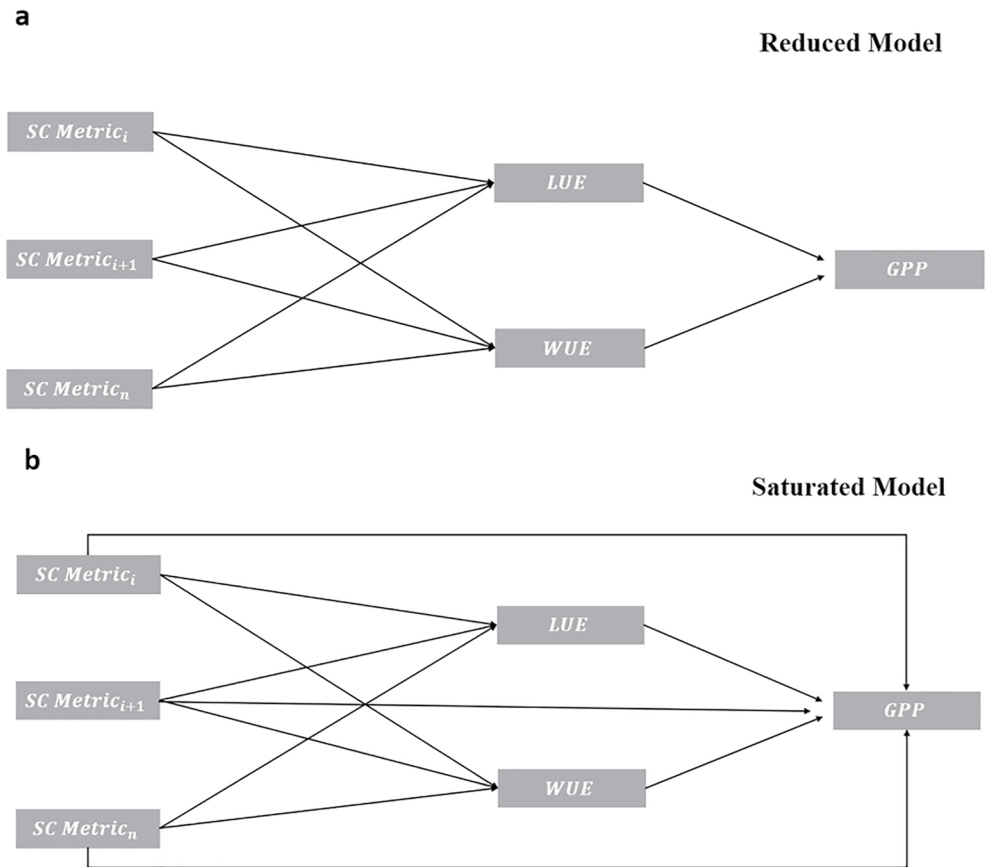


Figure 4. Conceptual figure outlining the (a) reduced and (b) saturated structural equation modeling (SEM) model designs. The reduced model (a) restricts canopy structural complexity (CSC) metrics to influencing the dependent variable, gross primary productivity (GPP), indirectly through their effect on light use efficiency (LUE) and water use efficiency (WUE), whereas the saturated model (b) allows CSC metrics to affect GPP both directly and indirectly through LUE and WUE. Arrows indicate the direction of influence from one variable to the next.

of shade tolerance, and that high shade tolerance variation within a community is likely one of the most important life-history traits, leading to more efficient resource use when scaled to the ecosystem level (Stark et al., 2012).

For most CSC metrics examined here, increasing CSC is associated with increasing RUE, although the magnitude of the trend is dependent upon resolution. The exception is LAI_{SD} , which has a negative relationship with both WUE and LUE at all resolutions. The strongest positive relationship exists between VCI_{AVG} and WUE, and the weakest relationship exists between R_T and LUE. Mediation analysis showed that neither WUE or LUE significantly mediated the relationship between $VAI_{maxheightmean}$ and GPP (Figure 6), suggesting that either the relationship is direct, or additional unaccounted for factors play the role of mediator. The most complex sites (SE5 and NE2) have differing relationships to productivity. Site NE2 has the highest GPP of all nine sites, but also has the highest R_{eco} , resulting in its classification as a small net source of CO_2 to the atmosphere. Site SE5 has the second lowest seasonal GPP as well as the second lowest R_{eco} . The two least complex sites, SE3 and NW2, have among the lowest total seasonal GPP and R_{eco} . SW2 and SW4 have the second and third highest seasonal GPP, yet consistently display only moderately levels of CSC at all five spatial resolutions. However, both of these sites contain stands in a wide range of age classes (Table 2), indicating heterogeneity in successional stages, and both sites are noted as containing very wet areas, with older (>100 years) mixed conifer swamp stands.

4.3. Disturbance Impacts

With the exception of NE2, sites with a record of greater disturbance severity and intensity, presented as clearcutting or shelterwood harvest, exhibit lower levels of complexity across the majority of CSC metrics, and across all

Table 3
Canopy Structural Complexity Metrics Included in SEM, Isolated as Highly Influential Through Best Subsets Selection for Their Strength as Drivers of GPP

Resolution (m)	Metric	Symbol	Units	Complexity category
0.25	rumple	rumple	ratio	Canopy heterogeneity
	verticalDistMax	$VAI_{maxheightmean}$	m	Vertical heterogeneity
	VCI_mean	VCI_{AVG}	—	Vertical heterogeneity
2	VCI_mean	VCI_{AVG}	—	Vertical heterogeneity
	LAI_mean	LAI_{AVG}	—	Area and density
	meanZ_sd	σ_H	m	Height
10	verticalDistMax	$VAI_{maxheightmean}$	m	Vertical heterogeneity
	maxZ_sd	R_T	m	Canopy heterogeneity
	maxZ_mean	MOCH	m	Height
	VCI_mean	VCI_{AVG}	—	Vertical heterogeneity
25	maxZ_sd	R_T	m	Canopy heterogeneity
	sdZ_sd	$Vert_{SDSD}$	m	Vertical heterogeneity
	sdZ_mean	$Vert_{meanStd}$	m	Vertical heterogeneity
	LAI_sd	LAI_{SD}	—	Area and density
50	maxZ_sd	R_T	m	Canopy heterogeneity
	gap_fraction	Θ	ratio	Cover and openness
	VCI_mean	VCI_{AVG}	—	Vertical heterogeneity
	canopy_ratio_mean	$Canopy\ Ratio_{AVG}$	ratio	Vertical heterogeneity
	LAI_sd	LAI_{SD}	—	Area and density

metrics addressing vertical complexity. One reason for this could be that the harvests at NE2 were all selective harvests, and resulted in distinct structurally heterogeneous “patches” within the site at different successional stages and with a high degree of canopy cover. In contrast to the primarily broadleaf understory present at multiple other sites, several patches within NE2 feature a prominent conifer understory. As mixed conifers tend to show higher levels of vertical complexity than many purely broadleaf stands do (Ehbrecht et al., 2017; Pommerening & Murphy, 2004; Zenner, 2016), the presence of a developing conifer understory could be contributing to a higher overall VCI_{AVG} . This is supported by the presence of a substantial conifer understory at one other site, SW4, which exhibits the highest degree of VCI_{AVG} amongst the nine sites (0.35). Again, with the exception of NE2, sites with a record of more substantial disturbance had lower levels of productivity, and lower levels of RUE. For example, site NW2, which had the highest frequency of both clearcuts and harvest events, had the lowest GPP of all nine sites and also had the lowest average daily LUE (0.33 gC MJ^{-1}) and WUE ($2.9\text{ gC kg H}_2\text{O}^{-1}$) values.

More moderate disturbances such as thinning and selective harvest could be contributing to increased CSC within the study area, through assisting in the transition to uneven aged stands (Gough et al., 2021). This is observed at site SE6, which consists of a 19-year-old mixed aspen, white spruce, and balsam fir stand, a 22-year-old jack pine stand, a 75-year-old aspen stand, and a 92-year-old mixed upland hardwood stand (Figure 2). SE6 underwent species-specific commercial thinning to reduce stand density, which has been shown to impact stand growth and structure (Wisconsin Department of Natural Resources, 2020). SE6 also experienced salvage cutting to remove dead or damaged trees following a severe hail storm in 2000. Sites SE5 and NE2 consistently ranked as the most complex sites regardless of spatial resolution, and both sites have experienced moderate management disturbances such as thinning as well as manual planting.

4.4. Implications and Shortcomings

A 2018 review by Fahey et al. (2018) showed that although the prevalence of complexity terminology with respect to silviculture has increased over time, the actual incorporation of complexity metrics when designing

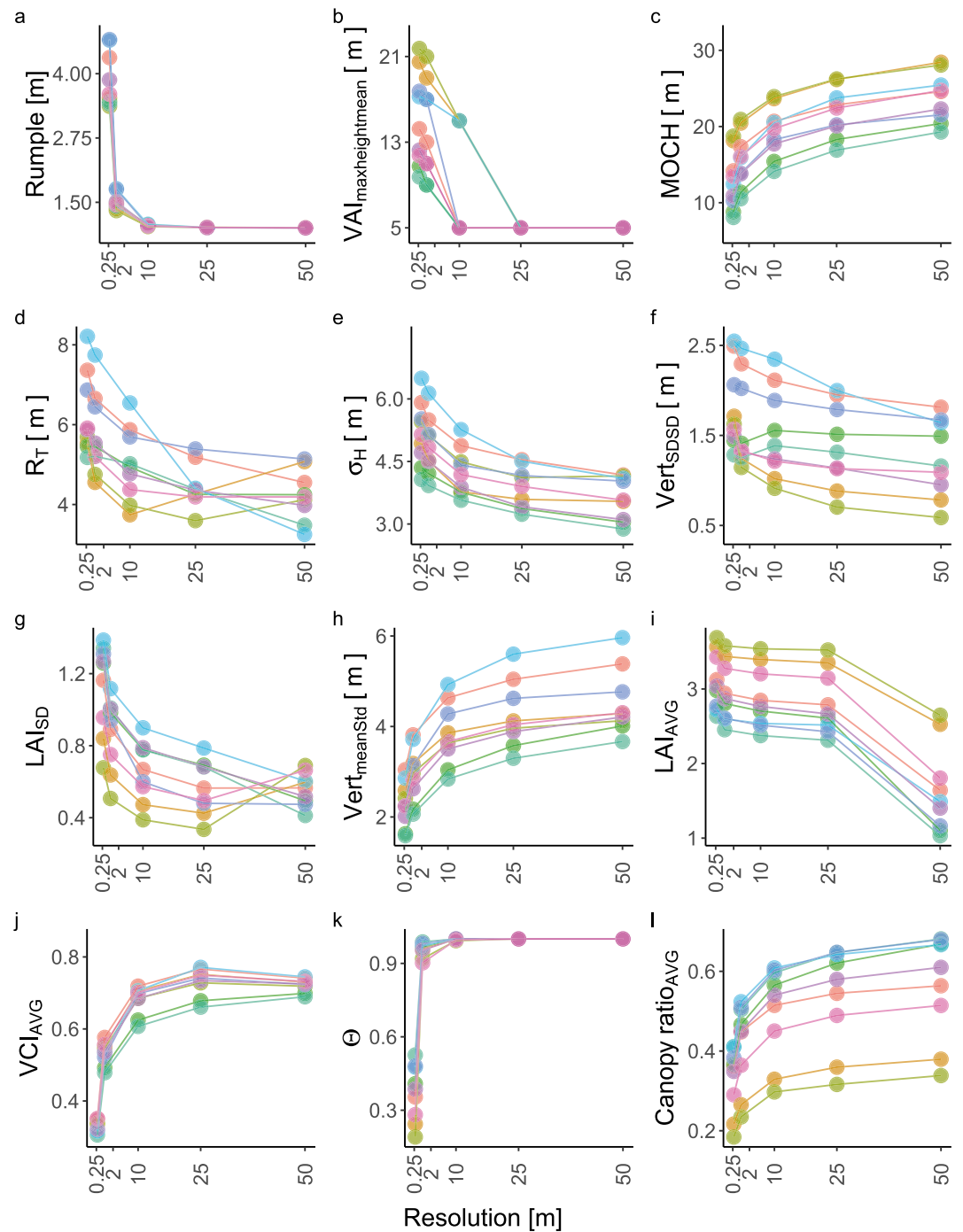


Figure 5. Canopy structural complexity (CSC) metric values by site at each of the five metric calculation resolutions explored, 0.25, 2, 10, 25, and 50 m.

long-term silviculture projects has “plateaued in the past decade or more.” This could indicate that although awareness about the importance of forest complexity has increased, a lack of understanding regarding the long-term impacts of managing to enhance complexity persists. Through the exploration of mechanistic relationships between forest CSC and function, this study highlighted which complexity metrics provide important information about RUE and productivity. These metrics can then be integrated as flexible structural parameters in mechanistic ecosystem models that simulate light and water-sensitive processes. Through this, we can improve the ability of models to mimic true ecosystem responses to management, from a biogeochemical perspective.

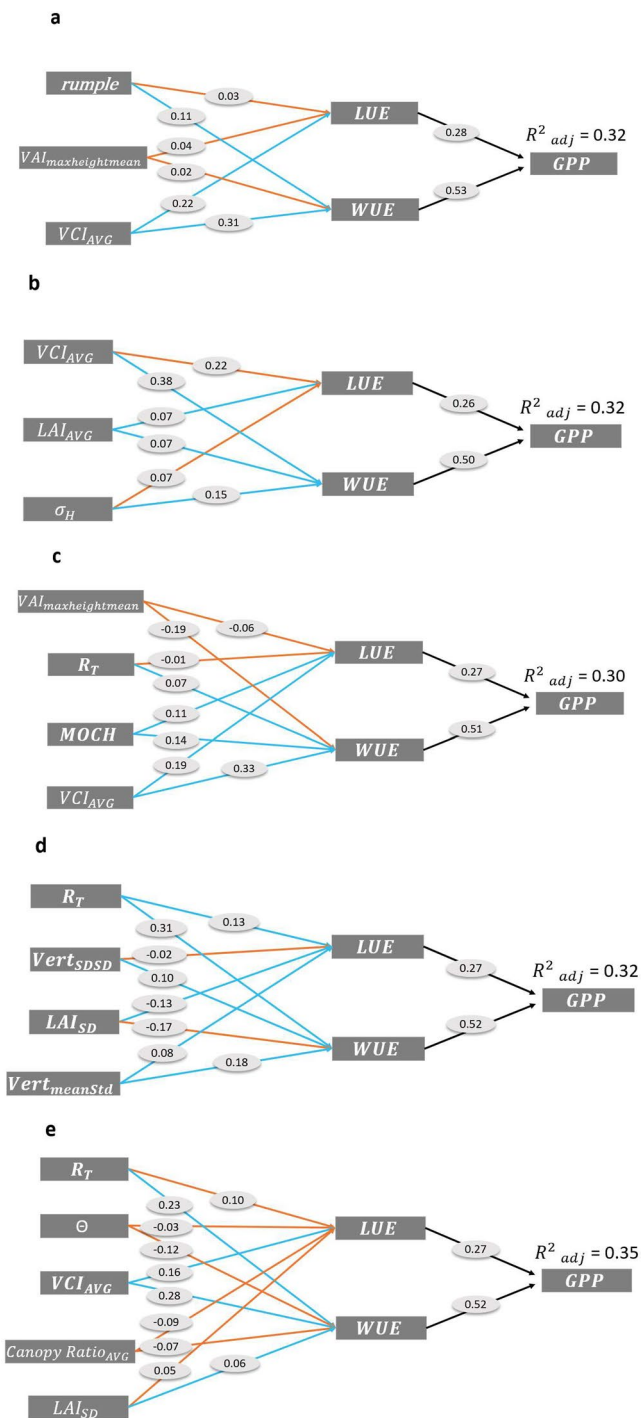


Figure 6. Reduced structural equation modeling (SEM) diagrams at canopy structure complexity (CSC) metric calculation resolutions of 0.25 m (a), 2 m (b), 10 m (c), 25 m (d), and 50 m (e). CSC predictor variables are on the far left, mediating variables (light use efficiency (LUE) and water use efficiency (WUE)) are in the center, and the response variable, gross primary productivity (GPP) is on the far right. Standardized regression coefficients are shown in gray circles. Blue lines indicate a statistically significant ($p < 0.01$) mediation effect between mediating and predictive variables, while orange lines indicate that mediation either does not exist or is not statistically significant.

This improved representation will allow us to explore the future response of forests to a variety of management regimes and representative concentration pathways, enhancing our ability to assess mitigation and adaptation strategies beyond direct observational studies, which often take many years to produce outcomes. This would facilitate more accurate predictions of the future of a suite of ecosystem goods and services including carbon storage potential, which could significantly impact the development of natural climate solutions in the regional Midwest.

The persistent superior performance of the reduced SEM, where the relationship between CSC and GPP is moderated by RUE, suggests that although specific CSC metric values change slightly with metric calculation resolution shifts, the existence of a mediation effect itself is not scale dependent. This indicates that the mechanistic relationships outlined here can be scaled up from the stand to the ecosystem level to provide novel insights into forest function and carbon storage potential. While this expands the utility of observational studies, it also provides new opportunities to validate and apply information obtained from satellites, such as the GEDI high resolution ecosystem LiDAR, which is capable of measuring global forest canopy height and vertical structure (Dubayah et al., 2020).

Although the EC method is the most well-established method for taking continuous measurements of energy and trace gas exchange (Desai et al., 2008), it is not without drawbacks. All measurements have associated uncertainty, and in the context of EC measurements these uncertainties can be segregated into several categories, including uncertainty due to instrument or calibration error, technological limitations of the instruments themselves, inadequate sample size, and environmental conditions that violate the assumptions at the core of EC theory (Richardson et al., 2012). Some of these errors are stochastic and appear as random noise in the data, while other errors are systematic and result in a bias that is relatively constant over time. Numerous other studies have explored these uncertainties at length (Hollinger & Richardson, 2005; Loescher et al., 2006; Massman & Lee, 2002; Richardson et al., 2006), but it is worth noting general trends in overall EC uncertainty here. Random error in 30-min fluxes ranges from 10% to 20% (Loescher et al., 2006), with annual estimates around 10% (Richardson et al., 2006), as error generally decreases with longer time series and averaging (Loescher et al., 2006). Flux uncertainty follows a strong seasonal pattern (uncertainty is generally higher during the growing season), and is sensitive to land cover type and wind speed (Hollinger & Richardson, 2005; Richardson et al., 2006). Error is also associated with the partitioning of NEE into GPP and R_{eco} and varies by partitioning method, but a survey of 23 methods conducted by Desai et al. (2008) showed that on average the difference in GPP was <10%, with additional uncertainty depending on the abundance of gaps in the data. In this study, there was an average of 37% gaps in measured NEE values across the nine sites.

This study primarily examined the influence of biotic forest factors, but the inclusion of prominent abiotic factors such as nutrient regimes could further enhance the study. Combining chemical analysis of leaves with the remote sensing of CSC and EC measurements of land-atmosphere carbon exchange would account for the influence of factors such as nitrogen availability in determining controls on RUE and productivity (Reich, 2012). Another limitation of this study is the relatively short window in which data were collected. Although this observational window supported the primary

Table 4
EC Flux Towers Included in This Study, With Unique DOI's

CHEESEHEAD tower name	Ameriflux tower name	Data set DOI
NW2	US-PFc	https://doi.org/10.17190/AMF/1717851
NE2	US-PFh	https://doi.org/10.17190/AMF/1717855
NE3	US-PFi	https://doi.org/10.17190/AMF/1717856
NE4	US-PFj	https://doi.org/10.17190/AMF/1717857
SW2	US-PFl	https://doi.org/10.17190/AMF/1717859
SW4	US-PFn	https://doi.org/10.17190/AMF/1717861
SE3	US-PFq	https://doi.org/10.17190/AMF/1717863
SE5	US-PFs	https://doi.org/10.17190/AMF/1717865
SE6	US-PFt	https://doi.org/10.17190/AMF/1717866

goals of CHEESEHEAD19 related to addressing issues of energy balance closure, from a carbon cycle perspective it failed to capture winter effects on net carbon budgets. Incorporating multiyear data sets would address this problem as well as allow for a more thorough examination of the influence of stand age on RUE and productivity, whereas here analysis was inconclusive. Moreover, as neither disturbance and management or stand age were expressly controlled for, the impacts of management and disturbance on the structure-function relationship explored here are largely qualitative. Although the high density of EC towers in a small study domain controlled for several factors such as differences in soil type, forest type, and mesoclimate, differences in microclimate still existed between sites. This is presented as variability in temperature, latent and sensible heat flux, and wind properties including turbulence. Although heterogeneity in land cover existed, there was very little difference in topography to drive variability in air circulation or relative humidity, so the observed differences in microclimate were likely due to diversity in vegetation type and density, as well as proximity to and abundance of water. Lastly, the somewhat small site sample size involved in this study suggests caution should be exercised when evaluating SEM fit statistics.

5. Conclusions

Quantifying mechanistic relationships between forest CSC and productivity is essential to advancing our ability to scale measurements from the leaf to stand to landscape level. This will greatly enhance our capacity to directly assess landscape-level ecosystem functions and implications for natural climate solutions. We approached this challenge using a combination of UAS LiDAR-derived CSC metrics from nine forested sites within a 10×10 km study domain, land-atmosphere exchange data from nine EC towers located within those forested sites, and SEM. Through employing a high density of EC towers across a relatively small spatial domain, we were able to separate variability in climate, soil fertility, and forest functional types from structural controls on productivity, allowing for a more representative physiological understanding than has been previously demonstrated. We conclude that (a) structural metrics describing the vertical complexity of a forest (specifically VCI_{AVG}) are the strongest drivers when predicting productivity in temperate mixed forests with a significant degree of heterogeneity and a long history of management; (b) variability in the type and intensity of management and disturbance legacies contribute to substantial differences in CSC metric values as well as productivity; (c) the relationship between forest structure and function is not direct, but is actively mediated by light and water RUE, with WUE being a stronger driver of GPP; and (d) CSC metric values change with shifts in the resolution of metric calculation, resulting in changes to the mechanistic relationship between forest structure and function. This emphasizes the need for consistency in the spatial resolution at which CSC metrics are calculated, and for the disclosure of resolutions of metric calculation when reporting CSC metric values and interpreting the significance of findings. These findings will allow us to improve mechanistic representation in ecosystem models of how CSC impacts light and water-sensitive processes, and ultimately GPP. This will strengthen the ability of models to mimic true ecosystem responses to management and disturbance, allowing for a more accurate assessment of the response of forests to various management regimes and representative concentration pathways, enhancing our ability to assess climate mitigation and adaptation strategies.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Drone LiDAR data are available at http://co2.aos.wisc.edu/data/CHEESEHEAD-incoming/Drone_LiDAR/. All flux and meteorological data utilized in this study, in addition to other data sets generated through CHEESEHEAD19, are publicly available through the CHEESEHEAD19 data repository hosted by the National Center

for Atmospheric Research's Earth Observing Laboratory at https://www.eol.ucar.edu/field_projects/cheesehead. Supporting information and photographs of the tower sites are available through the CHEESEHEAD19 project website: www.cheesehead19.org. EC tower data are also publicly available through the Ameriflux website, digital object identifiers (DOI) for the nine towers utilized in this study are presented in Table 4.

Acknowledgments

This project was financially supported by award AGS-1822420 from the National Science Foundation (NSF), the University of Wisconsin-Madison's Office of the Vice Chancellor for Research and Graduate Education (OVCERGE) fall competition award and NSF Emerging Frontiers Macrosystems Biology award DEB-1702996 supporting the Management and Disturbance in Forest Ecosystems (MANDIFORE) project. Jacob May received additional mentorship and funding from Dr. Philip Townsend of the University of Wisconsin-Madison. Brian Butterworth was additionally supported by the NOAA Physical Sciences Laboratory. We would like to acknowledge that the study domain is located in the ancestral homeland of the Ojibwe people. We honor the indigenous caretakers of these lands before us, today, and of generations to come.

References

- Ainsworth, E. A., & Long, S. P. (2005). What have we learned from 15 years of free-air CO₂ enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO₂. *New Phytologist*, *165*(2), 351–372. <https://doi.org/10.1111/j.1469-8137.2004.01224.x>
- Amiro, B. D., Barr, A. G., Barr, J. G., Black, T. A., Bracho, R., Brown, M., et al. (2010). Ecosystem carbon dioxide fluxes after disturbance in forests of North America. *Journal of Geophysical Research*, *115*, G00K02. <https://doi.org/10.1029/2010JG001390>
- Anderson-Teixeira, K. J., Herrmann, V., Morgan, R. B., Bond-Lamberty, B., Cook-Patton, S. C., Ferson, A. E., et al. (2021). Carbon cycling in mature and regrowth forests globally. *Environmental Research Letters*, *16*(5), 053009. <https://doi.org/10.1088/1748-9326/abed01>
- Anten, N. P. R. (2016). Optimization and game theory in canopy models. In K. Hikosaka, Ü. Niinemets, & N. P. R. Anten (Eds.), *Canopy photosynthesis: From basics to applications* (pp. 355–377). Springer. https://doi.org/10.1007/978-94-017-7291-4_13
- Arnfield, A. J. (2021). Koppen climate classification. Encyclopaedia Britannica. Retrieved from <https://www.britannica.com/science/Koppen-climate-classification>
- Atkins, J. W., Bohrer, G., Fahey, R. T., Hardiman, B. S., Morin, T. H., Stovall, A. E. L., et al. (2018). Quantifying vegetation and canopy structural complexity from terrestrial LiDAR data using the ForestR package. *Methods in Ecology and Evolution*, *9*(10), 2057–2066. <https://doi.org/10.1111/2041-210X.13061>
- Atkins, J. W., Fahey, R. T., Hardiman, B. H., & Gough, C. M. (2018). Forest canopy structural complexity and light absorption relationships at the subcontinental scale. *Journal of Geophysical Research: Biogeosciences*, *123*(4), 1387–1405. <https://doi.org/10.1002/2017JG004256>
- Binkley, D., Stape, J. L., & Ryan, M. G. (2004). Thinking about efficiency of resource use in forests. *Forest Ecology and Management*, *193*(1–2), 5–16. <https://doi.org/10.1016/j.foreco.2004.01.019>
- Birdsey, R., Pan, Y., Janowiak, M., Stewart, S., Hines, S., Parker, L., et al. (2014). *Past and prospective carbon stocks in forests of northern Wisconsin: A report from the Chequamegon-Nicolet National Forest Climate Change Response Framework (NRS-GTR-127, p. NRS-GTR-127)*. Northern Research Station: U.S. Department of Agriculture, Forest Service. <https://doi.org/10.2737/NRS-GTR-127>
- Birdsey, R., Pregitzer, K., & Lucier, A. (2006). Forest carbon management in the United States: 1600–2100. *Journal of Environmental Quality*, *35*(4), 1461–1469. <https://doi.org/10.2134/jeq2005.0162>
- Bogdanovich, E., Perez-Priego, O., El-Madany, T. S., Guderle, M., Pacheco-Labrador, J., Levick, S. R., et al. (2021). Using terrestrial laser scanning for characterizing tree structural parameters and their changes under different management in a Mediterranean open woodland. *Forest Ecology and Management*, *486*, 118945. <https://doi.org/10.1016/j.foreco.2021.118945>
- Bonan, G. B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, *320*(5882), 1444–1449. <https://doi.org/10.1126/science.1155121>
- Butterworth, B. J., Desai, A. R., Metzger, S., Townsend, P. A., Schwartz, M. D., Petty, G. W., et al. (2021). Connecting land-atmosphere interactions to surface heterogeneity in CHEESEHEAD19. *Bulletin of the American Meteorological Society*, *102*(2), E421–E445. <https://doi.org/10.1175/BAMS-D-19-0346.1>
- Camarretta, N., Harrison, P. A., Bailey, T., Potts, B., Lucieer, A., Davidson, N., & Hunt, M. (2020). Monitoring forest structure to guide adaptive management of forest restoration: A review of remote sensing approaches. *New Forests*, *51*(4), 573–596. <https://doi.org/10.1007/s11056-019-09754-5>
- Chlus, A., Kruger, E. L., & Townsend, P. A. (2020). Mapping three-dimensional variation in leaf mass per area with imaging spectroscopy and lidar in a temperate broadleaf forest. *Remote Sensing of Environment*, *250*, 112043. <https://doi.org/10.1016/j.rse.2020.112043>
- CloudCompare (version 2.10). (2019). GPL software. Retrieved from <http://www.cloudcompare.org/>
- Dănescu, A., Albrecht, A. T., & Bauhus, J. (2016). Structural diversity promotes productivity of mixed, uneven-aged forests in southwestern Germany. *Oecologia*, *182*(2), 319–333. <https://doi.org/10.1007/s00442-016-3623-4>
- Davis, K., Bakwin, P., Yi, C., Berger, B., Zhao, C., Teclaw, R., & Isebrands, J. G. (2003). The annual cycles of CO₂ and H₂O exchange over a northern mixed forest as observed from a very tall tower. *Global Change Biology*, *9*, 1278–1293. <https://doi.org/10.1046/j.1365-2486.2003.00672.x>
- De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Hickler, T., et al. (2013). Forest water use and water use efficiency at elevated CO₂: A model-data intercomparison at two contrasting temperate forest FACE sites. *Global Change Biology*, *19*(6), 1759–1779. <https://doi.org/10.1111/gcb.12164>
- Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., et al. (2008). Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. *Agricultural and Forest Meteorology*, *148*(6), 821–838. <https://doi.org/10.1016/j.agrformet.2007.11.012>
- Desai, A. R., Xu, K., Tian, H., Weishampel, P., Thom, J., Baumann, D., et al. (2015). Landscape-level terrestrial methane flux observed from a very tall tower. *Agricultural and Forest Meteorology*, *201*, 61–75. <https://doi.org/10.1016/j.agrformet.2014.10.017>
- Donager, J. J., Sánchez Meador, A. J., & Blackburn, R. C. (2021). Adjudicating perspectives on forest structure: How do airborne, terrestrial, and mobile lidar-derived estimates compare? *Remote Sensing*, *13*(12), 2297. <https://doi.org/10.3390/rs13122297>
- Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., et al. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, *1*, 100002. <https://doi.org/10.1016/j.srs.2020.100002>
- Ehbrecht, M., Schall, P., Ammer, C., & Seidel, D. (2017). Quantifying stand structural complexity and its relationship with forest management, tree species diversity and microclimate. *Agricultural and Forest Meteorology*, *242*, 1–9. <https://doi.org/10.1016/j.agrformet.2017.04.012>
- Ehbrecht, M., Seidel, D., Annighöfer, P., Kreft, H., Köhler, M., Zemp, D. C., et al. (2021). Global patterns and climatic controls of forest structural complexity. *Nature Communications*, *12*(1), 519. <https://doi.org/10.1038/s41467-020-20767-z>
- Eitel, J. U. H., Höfle, B., Vierling, L. A., Abellán, A., Asner, G. P., Deems, J. S., et al. (2016). Beyond 3-D: The new spectrum of lidar applications for Earth and ecological sciences. *Remote Sensing of Environment*, *186*, 372–392. <https://doi.org/10.1016/j.rse.2016.08.018>
- Fahey, R. T., Alveshere, B. C., Burton, J. I., D'Amato, A. W., Dickinson, Y. L., Keeton, W. S., et al. (2018). Shifting conceptions of complexity in forest management and silviculture. *Forest Ecology and Management*, *421*, 59–71. <https://doi.org/10.1016/j.foreco.2018.01.011>
- Fahey, R. T., Atkins, J. W., Gough, C. M., Hardiman, B. S., Nave, L. E., Tallant, J. M., et al. (2019). Defining a spectrum of integrative trait-based vegetation canopy structural types. *Ecology Letters*, *22*(12), 2049–2059. <https://doi.org/10.1111/ele.13388>

- Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: An updated review. *Ecological Processes*, 5(1), 19. <https://doi.org/10.1186/s13717-016-0063-3>
- Fisher, J. B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., et al. (2017). The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. *Water Resources Research*, 53, 2618–2626. <https://doi.org/10.1002/2016WR020175>
- Ford, S. E., & Keeton, W. S. (2017). Enhanced carbon storage through management for old-growth characteristics in northern hardwood-conifer forests. *Ecosphere*, 8(4), e01721. <https://doi.org/10.1002/ecs2.1721>
- Forrester, J. A., Mladenoff, D. J., & Gower, S. T. (2013). Experimental manipulation of forest structure: Near-term effects on gap and stand scale C dynamics. *Ecosystems*, 16(8), 1455–1472.
- Frelich, L. E. (1995). Old forest in the Lake States today and before European settlement. *Natural Areas Journal*, 15(2), 157–167.
- Gough, C. M., Atkins, J. W., Fahey, R. T., & Hardiman, B. S. (2019). High rates of primary production in structurally complex forests. *Ecology*, 100(10), e02864. <https://doi.org/10.1002/ecy.2864>
- Gough, C. M., Bohrer, G., Hardiman, B. S., Nave, L. E., Vogel, C. S., Atkins, J. W., et al. (2021). Disturbance-accelerated succession increases the production of a temperate forest. *Ecological Applications*, 31(7), e02417. <https://doi.org/10.1002/eap.2417>
- Gough, C. M., Curtis, P. S., Hardiman, B. S., Scheuermann, C. M., & Bond-Lamberty, B. (2016). Disturbance, complexity, and succession of net ecosystem production in North America's temperate deciduous forests. *Ecosphere*, 7(6), e01375. <https://doi.org/10.1002/ecs2.1375>
- Gough, C. M., Vogel, C. S., Harrold, K. H., George, K., & Curtis, P. S. (2007). The legacy of harvest and fire on ecosystem carbon storage in a north temperate forest. *Global Change Biology*, 13(9), 1935–1949. <https://doi.org/10.1111/j.1365-2486.2007.01406.x>
- Hardiman, B. S., Bohrer, G., Gough, C. M., & Curtis, P. S. (2013). Canopy structural changes following widespread mortality of canopy dominant trees. *Forests*, 4(3), 537–552. <https://doi.org/10.3390/f4030537>
- Hardiman, B. S., Bohrer, G., Gough, C. M., Vogel, C. S., & Curtis, P. S. (2011). The role of canopy structural complexity in wood net primary production of a maturing northern deciduous forest. *Ecology*, 92(9), 1818–1827. <https://doi.org/10.1890/10-2192.1>
- Hardiman, B. S., Gough, C. M., Halperin, A., Hofmeister, K. L., Nave, L. E., Bohrer, G., & Curtis, P. S. (2013). Maintaining high rates of carbon storage in old forests: A mechanism linking canopy structure to forest function. *Forest Ecology and Management*, 298, 111–119. <https://doi.org/10.1016/j.foreco.2013.02.031>
- Hardiman, B. S., LaRue, E. A., Atkins, J. W., Fahey, R. T., Wagner, F. W., & Gough, C. M. (2018). Spatial variation in canopy structure across forest landscapes. *Forests*, 9(8), 474. <https://doi.org/10.3390/f9080474>
- Hatfield, J. L., & Dold, C. (2019). Water-use efficiency: Advances and challenges in a changing climate. *Frontiers of Plant Science*, 10, 103. <https://doi.org/10.3389/fpls.2019.00103>
- Hillebrand, H., Langenheder, S., Lebrecht, K., Lindström, E., Östman, Ö., & Striebel, M. (2018). Decomposing multiple dimensions of stability in global change experiments. *Ecology Letters*, 21(1), 21–30. <https://doi.org/10.1111/ele.12867>
- Hocking, R. R., & Leslie, R. N. (1967). Selection of the best subset in regression analysis. *Technometrics*, 9(4), 531–540. <https://doi.org/10.1080/00401706.1967.10490502>
- Hollinger, D. Y., & Richardson, A. D. (2005). Uncertainty in eddy covariance measurements and its application to physiological models. *Tree Physiology*, 25(7), 873–885. <https://doi.org/10.1093/treephys/25.7.873>
- Hooper, D. U., Chapin, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., et al. (2005). Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs*, 75(1), 3–35. <https://doi.org/10.1890/04-0922>
- Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112(2), 351–362. <https://doi.org/10.1037/0033-2909.112.2.351>
- Kamoske, A. G., Dahlin, K. M., Serbin, S. P., & Stark, S. C. (2021). Leaf traits and canopy structure together explain canopy functional diversity: An airborne remote sensing approach. *Ecological Applications*, 31(2), e02230. <https://doi.org/10.1002/eap.2230>
- Kane, V. R., Bakker, J. D., McGaughey, R. J., Lutz, J. A., Gersonde, R. F., & Franklin, J. F. (2010a). Examining conifer canopy structural complexity across forest ages and elevations with LiDAR data. *Canadian Journal of Forest Research*, 40(4), 774–787. <https://doi.org/10.1139/X10-064>
- Kane, V. R., McGaughey, R. J., Bakker, J. D., Gersonde, R. F., Lutz, J. A., & Franklin, J. F. (2010b). Comparisons between field- and LiDAR-based measures of stand structural complexity. *Canadian Journal of Forest Research*, 40(4), 761–773. <https://doi.org/10.1139/X10-024>
- Knauer, J., El-Madany, T. S., Zaehle, S., & Migliavacca, M. (2018). Bigleaf—An R package for the calculation of physical and physiological ecosystem properties from eddy covariance data. *PLoS One*, 13(8), e0201114. <https://doi.org/10.1371/journal.pone.0201114>
- Kukal, M. S., & Irmak, S. (2020). Interrelationships between water use efficiency and light use efficiency in four row crop canopies. *Agrosystems, Geosciences & Environment*, 3(1), e20110. <https://doi.org/10.1002/agg2.20110>
- Loescher, H. W., Law, B. E., Mahrt, L., Hollinger, D. Y., Campbell, J., & Wofsy, S. C. (2006). Uncertainties in, and interpretation of, carbon flux estimates using the eddy covariance technique. *Journal of Geophysical Research*, 111, D21S90. <https://doi.org/10.1029/2005JD006932>
- Massman, W. J., & Lee, X. (2002). Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges. *Agricultural and Forest Meteorology*, 113(1–4), 121–144. [https://doi.org/10.1016/S0168-1923\(02\)00105-3](https://doi.org/10.1016/S0168-1923(02)00105-3)
- Mathias, J. M., & Thomas, R. B. (2021). Global tree intrinsic water use efficiency is enhanced by increased atmospheric CO₂ and modulated by climate and plant functional types. *Proceedings of the National Academy of Sciences of the United States of America*, 118(7), e2014286118. <https://doi.org/10.1073/pnas.2014286118>
- McElhinny, C., Gibbons, P., Brack, C., & Bauhus, J. (2005). Forest and woodland stand structural complexity: Its definition and measurement. *Forest Ecology and Management*, 218(1–3), 1–24. <https://doi.org/10.1016/j.foreco.2005.08.034>
- Monteith, J. L. (1965). *Evaporation and environment* (Vol. 19, pp. 205–234). Symposia of the Society for Experimental Biology.
- Odum, E. P. (1969). The strategy of ecosystem development. *Science*, 164(3877), 262–270.
- Olson, R. J., Holladay, S. K., Cook, R. B., Falge, E., Baldocchi, D., & Gu, L. (2004). *FLUXNET. Database of fluxes, site characteristics, and flux-community information* (ORNL/TM-2003/204). Oak Ridge National Lab. (ORNL). <https://doi.org/10.2172/1184413>
- Pan, Y., Chen, J. M., Birdsey, R., McCullough, K., He, L., & Deng, F. (2011). Age structure and disturbance legacy of North American forests. *Biogeosciences*, 8(3), 715–732. <https://doi.org/10.5194/bg-8-715-2011>
- Parker, G. G., Harmon, M. E., Lefsky, M. A., Chen, J., Pelt, R. V., Weis, S. B., et al. (2004). Three-dimensional structure of an old-growth Pseudotsuga-Tsuga canopy and its implications for radiation balance, microclimate, and gas exchange. *Ecosystems*, 7(5), 440–453. <https://doi.org/10.1007/s10021-004-0136-5>
- Penone, C., Allan, E., Soliveres, S., Felipe-Lucia, M. R., Gossner, M. M., Seibold, S., et al. (2019). Specialisation and diversity of multiple trophic groups are promoted by different forest features. *Ecology Letters*, 22(1), 170–180. <https://doi.org/10.1111/ele.13182>
- Pommerening, A., & Murphy, S. T. (2004). A review of the history, definitions and methods of continuous cover forestry with special attention to afforestation and restocking. *Forestry*, 77(1), 27–44. <https://doi.org/10.1093/forestry/77.1.27>

- Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., et al. (2005). Model-data synthesis in terrestrial carbon observation: Methods, data requirements and data uncertainty specifications. *Global Change Biology*, *11*(3), 378–397. <https://doi.org/10.1111/j.1365-2486.2005.00917.x>
- R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Reich, P. B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1736), 2128–2134. <https://doi.org/10.1098/rspb.2011.2270>
- Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., et al. (2005). On the separation of net ecosystem exchange into assimilation and ecosystem respiration: Review and improved algorithm. *Global Change Biology*, *11*(9), 1424–1439. <https://doi.org/10.1111/j.1365-2486.2005.001002.x>
- Reichstein, M., Stoy, P. C., Desai, A. R., Lasslop, G., & Richardson, A. D. (2012). Partitioning of net fluxes. In *Eddy covariance: A practical guide to measurement and data analysis*. Springer Science and Business Media.
- Rhemtulla, J. M., Mladenoff, D. J., & Clayton, M. K. (2009). Legacies of historical land use on regional forest composition and structure in Wisconsin, USA (mid-1800s–1930s–2000s). *Ecological Applications*, *19*(4), 1061–1078. <https://doi.org/10.1890/08-1453.1>
- Richardson, A. D., Aubinet, M., Barr, A. G., Hollinger, D. Y., Ibrom, A., Lasslop, G., & Reichstein, M. (2012). Uncertainty quantification. In *Eddy covariance: A practical guide to measurement and data analysis* (pp. 173–209). Springer Atmospheric Sciences.
- Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K. J., Flanagan, L. B., Katul, G. G., et al. (2006). A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. *Agricultural and Forest Meteorology*, *136*(1), 1–18. <https://doi.org/10.1016/j.agrformet.2006.01.007>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.
- Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., et al. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. *Remote Sensing of Environment*, *251*, 112061. <https://doi.org/10.1016/j.rse.2020.112061>
- Scheuermann, C. M., Nave, L. E., Fahey, R. T., Nadelhoffer, K. J., & Gough, C. M. (2018). Effects of canopy structure and species diversity on primary production in upper Great Lakes forests. *Oecologia*, *188*(2), 405–415. <https://doi.org/10.1007/s00442-018-4236-x>
- Schneider, F. D., Morsdorf, F., Schmid, B., Petchev, O. L., Hueni, A., Schimel, D. S., & Schaeppman, M. E. (2017). Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nature Communications*, *8*(1), 1441. <https://doi.org/10.1038/s41467-017-01530-3>
- Schulte, L. A., & Mladenoff, D. J. (2005). Severe wind and fire regimes in northern forests: Historical variability at the regional scale. *Ecology*, *86*(2), 431–445.
- Silva Pedro, M., Rammer, W., & Seidl, R. (2017). Disentangling the effects of compositional and structural diversity on forest productivity. *Journal of Vegetation Science*, *28*(3), 649–658. <https://doi.org/10.1111/jvs.12505>
- Smith, M. N., Stark, S. C., Taylor, T. C., Ferreira, M. L., deOliveira, E., Restrepo-Coupe, N., et al. (2019). Seasonal and drought-related changes in leaf area profiles depend on height and light environment in an Amazon forest. *New Phytologist*, *222*(3), 1284–1297. <https://doi.org/10.1111/nph.15726>
- Stark, S. C., Leitold, V., Wu, J. L., Hunter, M. O., deCastilho, C. V., Costa, F. R. C., et al. (2012). Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. *Ecology Letters*, *15*(12), 1406–1414. <https://doi.org/10.1111/j.1461-0248.2012.01864.x>
- Tang, H., & Dubayah, R. (2017). Light-driven growth in Amazon evergreen forests explained by seasonal variations of vertical canopy structure. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(10), 2640–2644. <https://doi.org/10.1073/pnas.1616943114>
- Tang, Z., Xu, W., Zhou, G., Bai, Y., Li, J., Tang, X., et al. (2018). Patterns of plant carbon, nitrogen, and phosphorus concentration in relation to productivity in China's terrestrial ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(16), 4033–4038. <https://doi.org/10.1073/pnas.1700295114>
- Thom, D., Taylor, A. R., Seidl, R., Thuiller, W., Wang, J., Robideau, M., & Keeton, W. S. (2021). Forest structure, not climate, is the primary driver of functional diversity in northeastern North America. *Science of The Total Environment*, *762*, 143070. <https://doi.org/10.1016/j.scitotenv.2020.143070>
- van Ewijk, K. Y., Treitz, P. M., & Scott, N. A. (2011). Characterizing forest succession in Central Ontario using lidar-derived indices. *Photogrammetric Engineering & Remote Sensing*, *77*(3), 261–269. <https://doi.org/10.14358/PERS.77.3.261>
- Williams, C. A., Gu, H., MacLean, R., Masek, J. G., & Collatz, G. J. (2016). Disturbance and the carbon balance of US forests: A quantitative review of impacts from harvests, fires, insects, and droughts. *Global and Planetary Change*, *143*, 66–80. <https://doi.org/10.1016/j.gloplacha.2016.06.002>
- Wisconsin Department of Natural Resources. (2019). *Wisconsin Wiscland 2 landcover database level 4*. Retrieved from <https://data-wi-dnr.opendata.arcgis.com>
- Wisconsin Department of Natural Resources. (2020). *Silviculture Handbook*. Forest Economics and Ecology. Applied Forestry Bureau.
- Wutzler, T., Lucas-Moffat, A., Migliavacca, M., Knauer, J., Sickel, K., Šigut, L., et al. (2018). Basic and extensible post-processing of eddy covariance flux data with REddyProc. *Biogeosciences*, *15*(16), 5015–5030. <https://doi.org/10.5194/bg-15-5015-2018>
- Zenner, E. K. (2016). Differential growth response to increasing growing stock and structural complexity in even- and uneven-sized mixed *Picea abies* stands in southern Finland. *Canadian Journal of Forest Research*, *46*(10), 1195–1204. <https://doi.org/10.1139/cjfr-2015-0400>
- Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., & Yan, G. (2016). An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sensing*, *8*(6), 501. <https://doi.org/10.3390/rs8060501>
- Zhang, Y., Chen, H. Y. H., & Reich, P. B. (2012). Forest productivity increases with evenness, species richness and trait variation: A global meta-analysis. *Journal of Ecology*, *100*(3), 742–749. <https://doi.org/10.1111/j.1365-2745.2011.01944.x>
- Zimble, D. A., Evans, D. L., Carlson, G. C., Parker, R. C., Grado, S. C., & Gerard, P. D. (2003). Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment*, *87*(2–3), 171–182. [https://doi.org/10.1016/S0034-4257\(03\)00139-1](https://doi.org/10.1016/S0034-4257(03)00139-1)