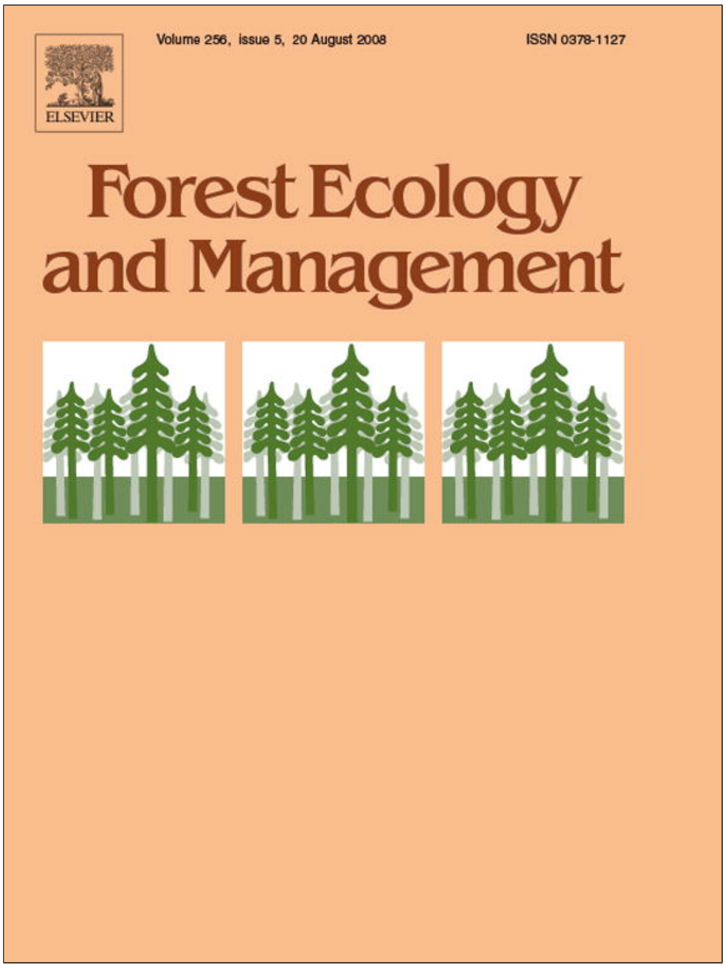


Volume 256, issue 5, 20 August 2008

ISSN 0378-1127

ELSEVIER

Forest Ecology and Management

Forest Ecology and Management 256 (2008) 872–882

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

A framework to categorize forest structure concepts

Nalini M. Nadkarni [a], Anne C.S. McIntosh [b,*], Judith B. Cushing [c]

[a] Lab II, The Evergreen State College, Olympia, WA 98505, USA
[b] Department of Renewable Resources, University of Alberta, 751 General Services Building, Edmonton, Alberta, T6G 2H1, Canada
[c] Lab I, The Evergreen State College, Olympia, WA 98505, USA

ARTICLE INFO

Article history:
Received 19 January 2008
Received in revised form 22 April 2008
Accepted 13 May 2008

Keywords:
Forest structure
Forest canopy
Classification
Categorization
Perception
Synthesis
Structure/function relationships

ABSTRACT

Forest structure affects ecosystem composition, dynamics, and function. The complexity of forest structure demands that researchers study only particular components of it, such as leaves, branches, or the medium of air that exists in between whole trees. Deciding which components to measure and how to analyze and portray them are determined by how a researcher's perception of the forest is colored by his or her particular questions of interest and by the measurement tools available. We have developed a conceptual framework to categorize how ecologists apparently perceive forest structure when they design and carry out their studies, with the objective of developing a better understanding of forest structure itself. The framework consists of a hierarchical categorization scheme that encompasses as many configurations of forest structural components used by researchers as possible. We first identified forest structural components examined by researchers and separated these into three major *representations*: groups of components, networks of components, and continuous components. Second, we applied three *descriptors* to each representation: dimensionality (four types), spatial referencing (three types), and reactiveness (two types). This created 72 potential categories (12 of which were impossible, leaving 60 possible categories). Third, we populated our framework by assigning forest structural components from each of 500 forest structure studies to these categories. Certain categories were much more heavily used than others; only 9 of the 60 possible categories were not populated by any studies. Potential applications of this framework include helping to combine data across categories; exploring associated functional attributes of each category to discern patterns in structure/function relationships; and prioritizing the development of ecoinformatics tools for the most commonly used categories or category combinations. This framework constitutes a new method for conceptualizing perceptions of forest structure and could also be applied to synthesis work, integrating forest structure data with data from other fields, or as a conceptual model for other fields where the structure of constituent components is complex.

1. Introduction

1.1. Importance of forest structure and its categorization

To understand and manage forests, we need to describe and categorize their complex and dynamic structural and spatial components (Robertson, 1987; Groffman and Tiedge, 1989; Martens et al., 1991). Forest structure – the quantities and spatial arrangements of forest components, including leaves, stems, and air – has been a central topic of forest ecology. Of the 87,564 forest-related citations in *Forestry Abstracts* (1976–2003), 11% list "canopy structure" and/or "forest structure" as keywords (the modern definition of "canopy structure" has overlapping boundaries with "forest structure"). Structure influences: (1) the abiotic environment, including wind (e.g., Harrington and DeBell, 1996), and light (Emborg, 1998); (2) forest dynamics (Runkle, 1990); (3) production and movements of volatile compounds (Rinne et al., 2000); (4) evolution and behavior of canopy-dwelling vertebrates (Putz et al., 1983, Dial et al., 2004); (5) rates of ecosystem processes such as photosynthesis and nutrient cycling (Ellsworth and Reich, 1993; Nadkarni and Sumera, 2004); and (6) hydrology (Clark et al., 1998). Many questions that concern forest structure require that a researcher select only a subset of all the components present, which means he/she must filter out or deduce certain components of forest structure. Thus, ecologists not only need tools to measure physical structures themselves, they also need a framework to categorize the ways in which they apparently perceive forest structure when they design and carry out their studies.

---

* Corresponding author. Tel.: +1 780 492 4135; fax: +1 780 492 4323.
E-mail addresses: nadkarnn@evergreen.edu (N.M. Nadkarni), amcintos@ualberta.ca (A.C.S. McIntosh), judyc@evergreen.edu (J.B. Cushing).

0378-1127/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.foreco.2008.05.021

Existing systems categorize forest structure, but these have tended to focus on particular attributes of the forest: (1) tree physiognomy (the shape of individual crowns, Grubb et al., 1963); (2) tree architecture (the forms of stems and growth patterns, Hallé et al., 1978); (3) height diversity (vertical distribution of canopy components, MacArthur and Horn, 1969); (4) stratification (predictable vertical separation into distinct horizons, Parker et al., 2002), (5) structural diversity (Franklin et al., 2002), and (6) fractals (a fragmented geometric shape that can be subdivided in parts, each with scale-invariant statistical properties Zeide, 1998).

However, many answers to questions relating to studying forest structure remain unknown. Examples include effects of forest fragmentation on mobile animal populations, consequences of global climate change on wood production and landscape-level disturbance regimes, and habitat needs of endangered species. Answers to these have been elusive for four reasons. First, some of the spatial components (e.g., outer branch tips) are physically difficult to get to, though recent canopy access methods have improved such access. Second, new instrumentation (remote sensing, laser pulses) now provides vast amounts of forest structure data, but finding the data and relating them explicitly to other data or to the question at hand is not straightforward. Third, the response time of long-lived trees is slow relative to the life cycle of grants and of human researchers. Fourth – and the factor we explore in this paper – interpretation of forest structure can differ greatly among researchers, depending on their questions and approaches. This places a subtle but potentially large obstacle to sharing, comparing, and integrating forest structural data that might otherwise eventually allow us to understand forest function, scale-up or -down in space or time, and connect or combine spatial components of different types.

1.2. Multiple perceptions of forest structure

Our central point is that prior to collecting data that describe forest structure, a researcher consciously or subconsciously selects forest components relevant to his/her question. The kernel of this idea appeared in Parker (1995):

> Separate study objectives require different operational representations of canopy structure, even for the same forest. For an investigation focused on organisms inhabiting woody surfaces (e.g., epiphytes), the canopy may be conceived as a network of connecting limbs (e.g., Nychka and Nadkarni, 1990). The canopy has been conceptualized as a community of leaves and studied demographically (Parker et al., 1989) ... or as a three-dimensional porous medium, having both passive and active surfaces. This will be a subset of the components that are physically present, which gives rise to a particular conceptual "view" that best describes the structure of the forest to answer the particular question of a given researcher.

These differing views of forest structure evoke the childhood poem, *The Blind Men and the Elephant* (J.G. Saxe, *The Oxford Treasury of Children's Poems*). Each blind man examined a different part of the elephant – the tusk, the tail, the leg, or the trunk – and each came to wildly different conclusions about what the elephant looked like based on his particular experience and perceptual model. Just as each blind man failed to understand the topologically complex elephant because he did not recognize that his perception was only partial, so forest ecologists have sometimes failed to recognize that they are "blind" to those structural components of the forest that do not relate to their own study. One researcher, for example, might measure attributes for one component of interest (e.g., the photosynthetic capacity of leaves of one tree species), while another researcher might focus on attributes of another component of interest (e.g., the air movement over the boundary layer of collective foliar surfaces of the entire forest). Both appear to be studying leaves. However, the first researcher's perception of the forest is that it is a collection of individual objects (leaves), each one fixing an amount of carbon over time. The second researcher's view of the forest is that it is a continuous medium of air passing over the collective surface of the leaves, which functions as a solid object. Neither researcher's perspective takes the entirety of forest structure into account. However, both researchers are partly correct because their respective conceptual views capture an aspect of forest structural complexity. However, both researchers are also incorrect because neither's conceptual view encompasses nor explicitly relates to the other's view of forest structure. Such disconnects make it difficult for the two researchers to effectively compare or exchange data or visualizations, even though they have both studied the same leaves in the same forest.

We posit that a conscious awareness of which conceptual view(s) of forest structure exist(s) in the mind and measurements of the researcher, and understanding how different conceptual views relate to one another might help to compare and integrate forest structure studies. In the example above, categorizing the two seemingly identical components of interest (leaves) and their measured attributes in this example (photosynthetic capacity of individual leaves vs. turbulence created by air movement) within a framework that clearly identifies how each category relates to the other can potentially enable comparative work, and ultimately, ecological synthesis.

We further posit that forest ecologists lack a common framework to describe and categorize their *perceptions* of forests. To address a wide range of possible research questions, a categorization method must incorporate concepts and data from many different sampling components and spatial scales. For example, to study the movements of a small-bodied sedentary invertebrate, a single branch might constitute a suitable structural component. To study movement of a wide-ranging raptor, parcels within a large forest/pasture landscape would be the more appropriate components. If the flow of pollen is of interest, then the sampling scheme must treat the air and pollen contained therein as a continuous medium.

1.3. Models from other fields

Other fields use established frameworks to effectively and directly categorize components of interest. For example, in the airline industry, "lost luggage charts" streamline the process of identifying lost baggage, using a simple set of visual examples of different types of luggage and characteristics of their closures (e.g., zipper or clasp?), wheels (i.e., do they have them?), and body (e.g., what color?). Second, jewelers use gemology charts that define the attributes of a gem based simply on type, size, and number and angle of facets. This allows them to categorize any precious stone with a group of numbers and letters that describes its structure. Third, soil scientists use Munsell charts to explicitly associate soil color with numerical values to identify the hue, chroma, and value of each soil sample, thus characterizing any soil sample in a replicable way. These frameworks facilitate the rapid, standardized, repeatable, and efficient identification of complex objects, and can be extended to easily find a particular object among many other objects or within a complex system (e.g., a traveler can pinpoint his own luggage from a large pool of lost baggage, a jeweler can search for a particular size and shape of gem on the international market, a soil scientist can quantitatively describe a soil type to any other soil scientist).

The lack of a unified framework analogous to those above for identifying ecologists' perceptions of forest structure is puzzling, given the critical importance of comparing and integrating forest and canopy structural studies. Historically, one approach to improve the ability to compare datasets was to impose standardized protocols for field measurements of forest structure. However, standardizing ecological methods is difficult for both sociological and scientific reasons. Individuals prefer to determine their own protocols because of their past experiences and because methods change as instrumentation or scientific questions change. Previously collected data can only rarely be modified retrospectively to conform to standardized data.

We have developed a unified framework to categorize how ecologists appear to perceive forest structure. This was done in conjunction with our development of informatics and visualization tools for canopy scientists (Cushing et al., 2007; McIntosh et al., 2007). Our framework encompasses the perceptions of all types of forests and can be used to hierarchically categorize them into a finite number of specific and repeatable categories, just as soil scientists can for soil colors. In this paper, we present the conceptual framework and suggest that ecologists could use it to abstract, analyze, and compare the way they and others perceive forest structure, even where field data collection methods differ. We anticipate that this framework will help researchers become more aware of their own views of forest structure and enhance their ability to compare and synthesize their data with other datasets.

2. Methods

Our approach comprised three steps. First, we developed a preliminary framework by identifying categories to hierarchically organize forest structure data collected across a wide spectrum of sites, spatial dimensions, and protocols. Second, we populated the categories with examples from the literature on forest structure. Third, we analyzed studies that did not fit the categorization and drew conclusions about our framework and its applications.

We first briefly define six terms that have specific meanings in our system, which we expand upon throughout the paper:

- A forest *component* is a particular structural element of interest, e.g., individual tree branch, crown, or temperature gradient.
- An *attribute* is a quality, characteristic, or object that is measured and is inherent in or ascribed to a component (e.g., height, diameter, and leaf area).
- A forest *unit* is an aggregation of components and attributes that relates to those components, e.g., a tree – which is an aggregation of branches – and its diameter and length.
- A *category* is a defined division in our system of categorization.
- A *representation* is the first hierarchy of our categorization framework.
- An *entity* is each representation in a given study, along with its three *descriptors* (i.e., dimensionality, spatial referencing, and reactiveness).

To develop our preliminary conceptual framework, we first considered the range of structural components and attributes of forest structure that are typically included in forest ecology studies. In collaboration with other forest canopy researchers (http://canopy.evergreen.edu/workshop02/), we listed ways that forest *units* (components and associated attributes) could be conceptualized, the measurements that produced data within each of these conceptualizations, and the types of analyses and insights that each conceptualization yielded. We then developed a preliminary framework, consisting of three major *representations* that could encompass these units and concepts, and created three descriptors that hierarchically refine each of those three representations.

Second, we populated the preliminary framework with units from 500 papers on forest structure that were drawn from peer-reviewed literature. We drew upon an existing database containing 7100 citations relevant to forest ecology maintained by the International Canopy Network (www.evergreen.edu/ican; Nadkarni and Parker, 1994). These references have been compiled since 1994 from systematic weekly searches of *Current Contents on Diskette* (CCOD, 2008), and include citations from over 150 journals relevant to forest ecology, published between 1964 and 2003. Each citation was linked to one of 19 keywords (e.g., forest structure, nutrient cycling, forest management, ecosystem processes—see http://canopy.evergreen.edu/citations for complete list).

We chose papers to populate our framework in two ways. The first 120 papers represent a broad range of forest types, conceptualizations, and units of forest structure. The remaining papers were selected randomly from the entire pool of studies that contained keywords "canopy structure", "forest structure", or "remote sensing", and also contained explicit methods and quantitative data on one or more aspects of canopy or forest structure and/or function. Purely theoretical papers were excluded. For each paper, we noted which components of forest structure (e.g., stems, branches, and snags) and which composite or functional attributes were measured, calculated, or estimated (e.g., leaf area index, and basal area). We then took these components and attributes as the basic "unit" or units of forest structure in which the researcher was interested. We assigned each unit to a forest structure category. Unit category assignments were recorded in a Microsoft Access database and tagged to the original literature source.

Third, after assigning each unit to a forest structure category, we identified studies that did not fit any category and determined whether certain categories within our framework might only be theoretical (i.e., it is physically impossible an actual forest unit fits that category). We then drew conclusions about both our framework and the applications for which it might be used in the future.

3. Results

3.1. Representations of forest structure

Our preliminary framework established three major *representations* (Fig. 1), into which each unit fit. The first representation is "Groups of Components" (GC) (Fig. 1A). Individual components within any given group of components are not interconnected, i.e., their physical connection to other parts of a tree or forest is not relevant. For example, Fujimori (1971) quantified the surface area of leaves, using measurements from multiple leaves, without noting the inter-connectedness among their locations. He was interested only in total photosynthetic area, regardless of the leaves' relationship to particular branches or stem portions.

The second representation is "Networks of Components" (NC) (Fig. 1B), in which components are considered inter-connected. The location of each component with respect to others is important to the question(s) being posed. For example, a researcher interested in the flow of water through a tree's vascular system needs to know which components are closely connected to each other, and which are more distantly linked in order to understand source–sink relationships for fixed carbon. For example, Sumida et al. (2002) quantified the distances and effects of branches from a neighboring tree on the orientation of branches within a tree.

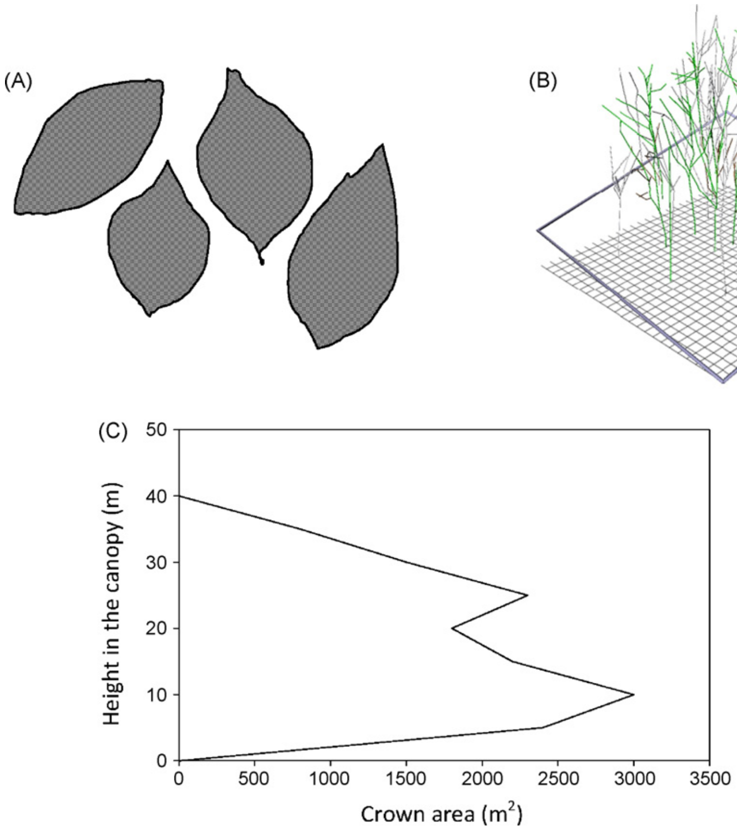


Fig. 1. Iconic visualization of forest structure components characterized as the three structural representations—(A) *Groups of components*: a group of leaves represents a group of surfaces that function to capture sunlight (Fujimori, 1971); (B) *Networks of components*: inter-connected stems and branches within individual trees represent components where the links between them are critical to the questions involved (Sumida et al., 2002); (C) *Continuous component*: a 2D representation of the vertical height profile of the cross-sectional crown area of a forest is an example of the forest structure being represented as a continuous component (Drake et al., 2002).

The third representation is "Continuous Components" (CC) (Fig. 1C), one that is continuous, rather than unconnected or inter-connected. For example, the upper surface of the canopy as recorded by measurements of tree height taken at intervals from the gondola of a canopy crane is then transformed into a surface map using contour lines. Even though the individual measurements might be initially perceived and measured as a group of components, if those measurements are refined and conceptually treated as a continuous component, we categorize the unit as the latter. For example, Nelson et al. (1984) used a laser profiling system to record ground height and uppermost canopy height continuously along the length of a flight path, which resulted in a cross-sectional area of the forest canopy.

3.2. Descriptors of forest structure representations

We created three descriptors that hierarchically refine each of the representations: *dimensionality*, *spatial referencing*, and *reactiveness*. Dimensionality reflects the number of dimensions of structural attributes collected or derived for a given unit of study within a representation, ranging from 0D to 3D (Fig. 2A). We describe a unit as zero-dimensional (0D) when structural attributes are not collected. For example, Gratani and Foti (1998) measured the photosynthetic activity of a subset of leaves, but did not take any structural measurements. One-dimensional (1D) units have structural attributes recorded along a single axis; e.g., Sumida et al. (2002) recorded branch lengths without collecting diameter information. Two-dimensional (2D) units are those where attribute measurements capture two axes; e.g., Fujimori (1971) traced the perimeter of selected leaves and then calculated surface area. Three-dimensional (3D) units incorporate structural attributes for all three spatial axes (*x*, *y*, and *z*); e.g., Sumida (1995) measured crown volume using the cylinder probe method. Note that for groups of components and continuous components, all levels of dimensionality are possible. However, networks of components are constrained to 1D through 3D because they are spatially inter-connected by definition—no zero-dimensional categories are possible.

The second refining descriptor for the representations is *spatial referencing* (Fig. 2B). There are three levels of spatial context in which information on given study units might be collected: *none* (NS), *relative* (RS), and *absolute* (AS). Relative spatial referencing occurs when information about a unit is provided only in terms of its location at a local scale; for example, when tree locations are mapped with respect to locally assigned *x* and *y* coordinates (e.g., Kneeshaw and Burton, 1997), or alternatively, when their location is given in terms of other components within the system (e.g., Kinerson and Fritschen (1971), who selected branches for sampling depending on the whorl in which they were located). Absolute spatial referencing provides information that allows a researcher to return to the given location and relocate the given component, (e.g., Imhoff et al., 1986). All three spatial referencing levels may occur among the groups of components and continuous components. Because of their inherent inter-connectedness, networks of components always include spatial referencing information (either RS or AS).

The third refining descriptor for each representation is *reactiveness* (Fig. 2C), which refers to whether or not a given unit has some functional capacity recorded for it. Units assigned to any of the three representations can be considered as either reactive

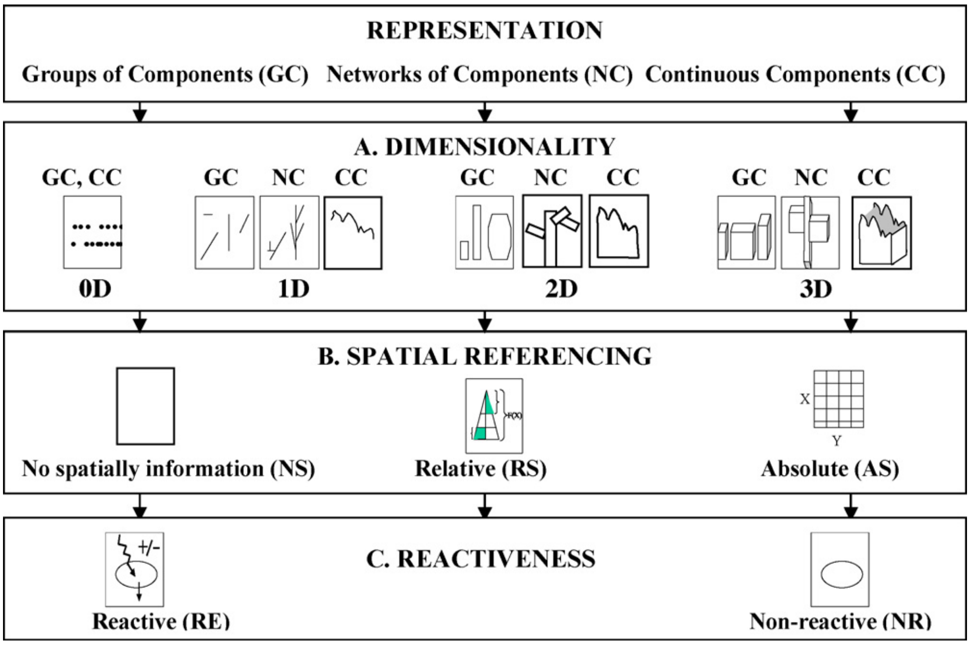


Fig. 2. Framework to categorize forest structure. The three representations that forest structural components (e.g., trees, branches, gaps) can encompass are groups of components (GC); networks of components (NC); and continuous components (CC). Three descriptors further refine these representations: (A) dimensionality (*x*D), (B) spatial referencing (NS, RS, AS), and (C) reactiveness (RE, NR). Dimensionality can range from 0- to 3D. Spatial referencing may occur at a local scale, with location information provided only with respect to other components within the stand (RS, e.g., branch position measured with respect to its location within the crown), or at an absolute scale when actual coordinates (AS, e.g., UTMs) are provided. A component is reactive (RE) or not reactive (NR), depending on whether or not it has some functional capacity recorded for it (e.g., reactive if sunlight absorption is recorded).

(RE) or non-reactive (NR). For example, leaves that have photosynthetic levels recorded for different levels of light are considered reactive. If those same leaves instead had their imperviousness to throughfall measured, they would be reactive, but with respect to foliar absorption of water rather than photosynthetic capacity. In general, if a unit only has structural attributes recorded (i.e., no functional attributes collected), then the unit is non-reactive. Units assigned to any of the three representations can be considered as either reactive (RE) or non-reactive (NR).

Each representation in a given study, along with its three descriptors is defined as an *entity*. Most studies included more than one *unit* (defined as an aggregation of a structure component and its attributes), resulting in 1050 total entity occurrences, exceeding the number of papers we examined (500 total papers). Each entity occurrence was considered as a single sample (Table 1).

We visually present the distribution of all entities by locating them within *categories*, which collectively creates our categorization space visualized within three cubes (Fig. 3). Each cube denotes a representation, and the axes capture its three refining descriptors, with each cube cell delineating an *entity category*. There are 60 possible categories, and each category for groups of components (24), networks of components (12), and continuous components (24) contains a unique "address" that identifies it. Just as Munsell Soil Chart codes uniquely identify and differentiate among soil colors (e.g., 5.5YR for a particular hue, chroma, and value of soil color), so do we designate a unique address for each *entity category* captured in a study, in the form of an alphanumeric string. The first two letters describe the representation, the number (at the start of the subscript string) describes the dimensionality, the next two letters in the string describe the spatial referencing, and the last two letters represent the reactiveness. For example, the address $GC_{1D\text{-}NS\text{-}NR}$ represents the category that is a group of components (GC) that is one-dimensional ($_{1D}$), non-spatially referenced ($_{NS}$), and non-reactive ($_{NR}$).

Table 1
Example occurrences of forest entities from existing forest research studies, categorized into our framework

Representation	Component	Dimensionality	Reactiveness	Spatial-referencing	"Address"	Source
GC	Needles	0D (no spatial measurements taken)	No	Relative: with respect to whorl within the crown	$GC_{0D\text{-}RS\text{-}NR}$	Kinerson and Fritschen (1971)
GC	Crowns	1D (crown length measured)	No	None	$GC_{1D\text{-}NS\text{-}NR}$	Valinger and Fridman (1997)
NC	Trunk sections	2D (length and width measured)	Yes (amount of epiphyte cover)	Relative: with respect to location on trunk	$NC_{2D\text{-}RS\text{-}RE}$	Pike et al. (1977)
NC	Forest patches	3D (timber volume calculated for each patch)	Yes (presence of flying squirrels)	Absolute: mapped with Landsat TM 5 satellite images	$NC_{3D\text{-}AS\text{-}RE}$	Reunanen et al. (2000)
CC	Cross-sectional region of the forest canopy	2D (ground and upper canopy heights along plane flightline using LIDAR)	No	None	$CC_{2D\text{-}NS\text{-}NR}$	Nelson et al. (1984)
CC	Forest canopy surface	3D (area of leaves by height in the canopy)	No	Absolute: *X*, *Y*, *Z* coordinates	$CC_{3D\text{-}AS\text{-}NR}$	Tanaka et al. (1998)

GC = group of components, NC = network of components, and CC = continuous component.

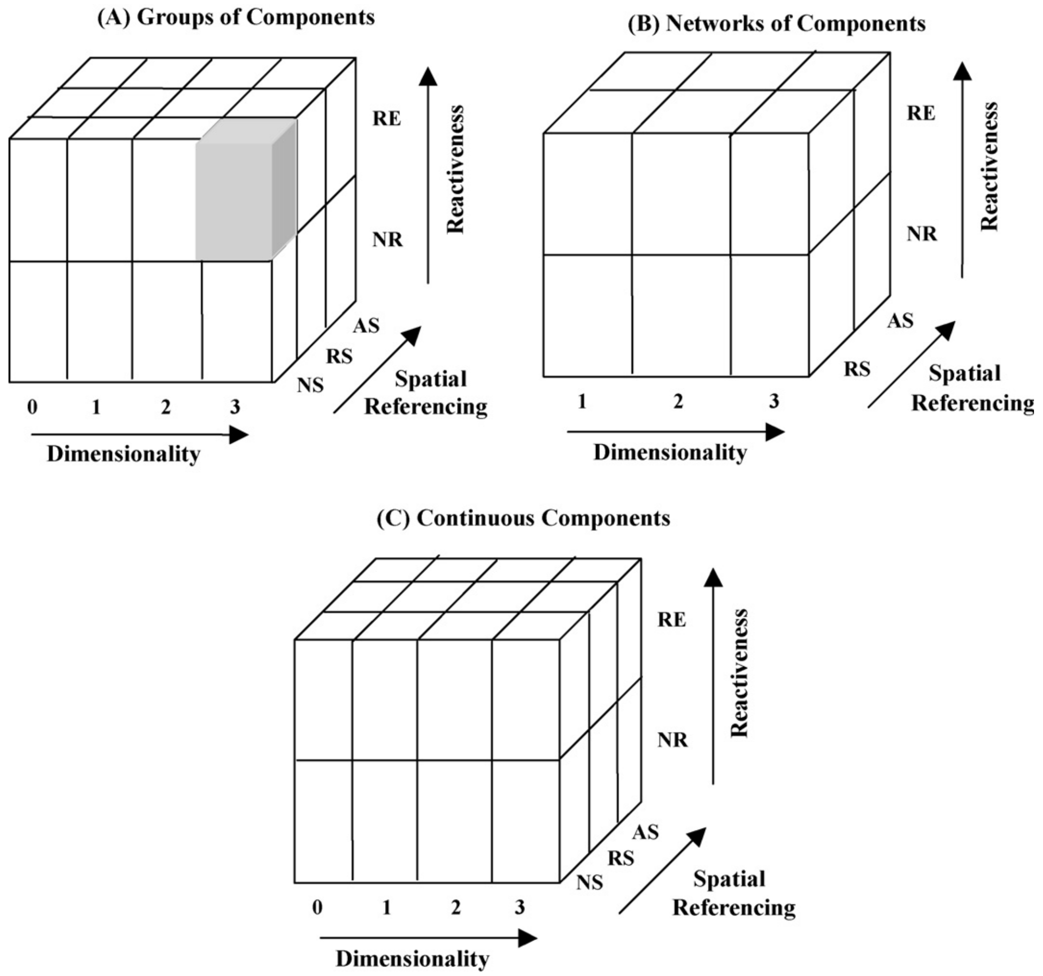


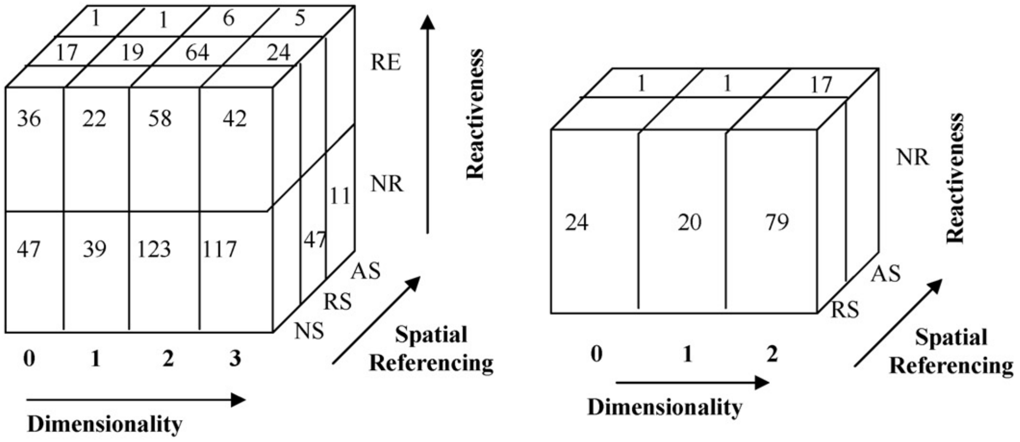
Fig. 3. (A–C) The 60 possible "addresses" for entity categories within our framework, using cubes to visualize each of the three general representations and their three associated descriptors. For example, the shaded address in (A) is $GC_{3D\text{-}NS\text{-}RE}$.

Table 2
The distribution of occurrences of entities within our framework

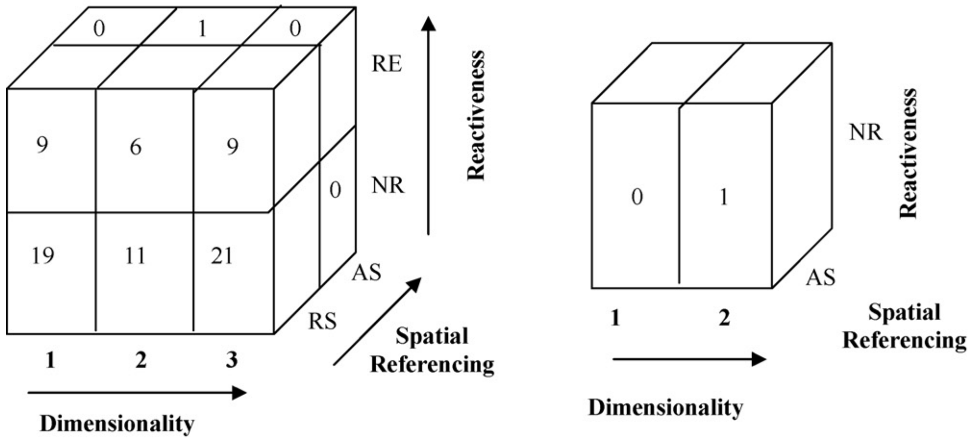
Descriptor	Representation			Percentage of entities of each descriptor for all representations
	GC (*n* = 821)	NC (*n* = 77)	CC (*n* = 152)	
Dimensionality				
0D (*n* = 131)	15.3	n/a	3.3	12.5
1D (*n* = 140)	12.4	36.4	6.6	13.3
2D (*n* = 428)	42.3	24.7	40.8	40.8
3D (*n* = 351)	30.0	39.0	49.3	33.4
Spatial referencing				
NS (*n* = 558)	59.1	n/a	49.3	53.1
RS (*n* = 384)	36.1	97.4	11.2	36.6
AS (*n* = 108)	5.2	2.6	41.4	10.3
Reactiveness				
NR (*n* = 654)	64.1	67.5	50.0	62.3
RE (*n* = 396)	35.9	32.5	50.0	37.7
Percentage of entities for all representations	78.2	7.3	14.5	

Columns contain our three forest representations and rows their three descriptors. We tallied 1050 entity occurrences from 500 published studies. Designations of representations are groups of components (GC), networks of components (NC), and continuous components (CC). Descriptors are dimensionality, spatial referencing, and reactiveness, as described in the text. Sample size (*n*) is the number of entities drawn from the literature and categorized into the representations and descriptors. Numbers in the body of the table are the percentages of entity occurrences by descriptors and representation types. The number in each cell is the percentage of entities for each descriptor as a proportion of the total number of entities in each of the representations. For example, the number in the uppermost cell on the left (15.3) shows that within the entities that were designated as Groups of Components (*n* = 831), 15.3% of them were zero-dimensional. The rightmost column is the percentage of each descriptor over all representations. For example, the uppermost right cell indicates that for all representations combined, 12.5% were zero-dimensional. The bottommost row is the summary of the percentage of each representation over all descriptors. The bottommost left cell indicates that 78.2% of all entities were categorized as Groups of Components. Note that this table does not categorize all 60-entity categories separately, because it is looking at each representation-descriptor group independently of the other descriptor groups—see Fig. 4 for distribution among each of the 60-entity categories.

(A) Groups of Components



(B) Networks of Components



(C) Continuous Components

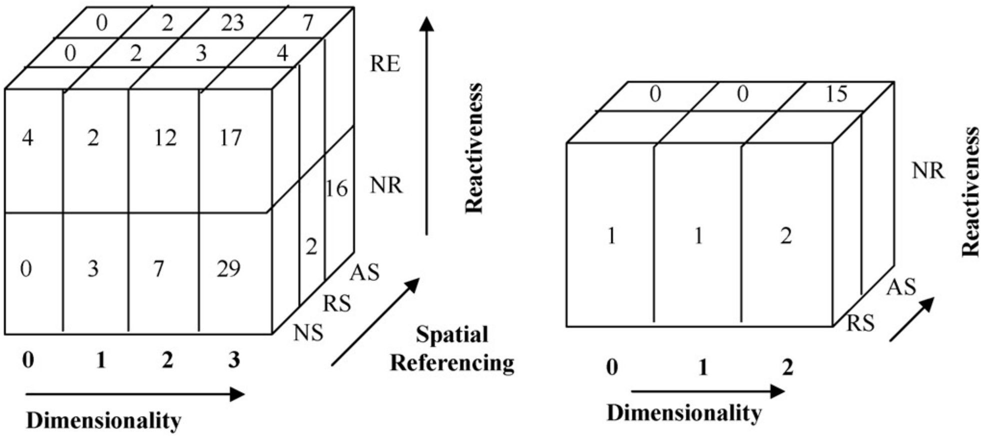


Fig. 4. Population of the 60 possible ‘‘addresses’’ presented in Fig. 3 with our 1050 occurrences of entities from the 500 studies illustrated by the number of entities assigned to each cube cell for (A) groups of components, (B) networks of components, and (C) continuous components. For each representation, the categories of the cubes that were hidden from view are presented to the right of the complete cube.

3.3. Distribution of entities

We explored the distribution of existing occurrences of entities within the framework by assigning each of the 1050 entity occurrences within our 500 papers to one of the framework's 60 possible categories (Table 2, Fig. 4). With the exception of only 2 of our 500 studies, our forest categorization framework categorized all of our units of study. The exceptions were units in two papers that estimated fractal dimensions of tree crowns, an approach that involved non-integer dimensions (Zeide and Pfeifer, 1991; Zeide, 1998). For our initial 120 studies, 17 categories within our framework were unpopulated. The number of unpopulated categories decreased to nine when we added 380 additional studies.

Table 3
Hypothetical studies that could populate some of the “empty” entity categories

Representation	Component	Dimensionality	Reactiveness	Spatial referencing	“Address”
Network of components	Crown sections	1D (1-m vertical intervals of crown)	Yes, presence of arthropods	Absolute, recorded for GPS-surveyed trees	$NC_{1D\text{-}AS\text{-}RE}$
Network of components	Crown sections	1D (1-m vertical intervals of crown)	No, measured biomass	Absolute, recorded for GPS-surveyed trees	$NC_{1D\text{-}AS\text{-}NR}$
Continuous component	Upper surface of the canopy	1D (Height of the canopy from a LIDAR flight line)	No	Absolute, recorded GPS coordinates along LIDAR path	$CC_{1D\text{-}AS\text{-}NR}$

The distribution of the entities within the representations provided insights into how forest ecologists appear to perceive forest structure (Table 2, Fig. 4). First, the distribution of the entity occurrences indicates that most researchers (78.2%) saw the forest as groups of components, whereas fewer than 10% viewed them as networks of components. Second, nearly three-quarters of the entity occurrences were categorized as being of either 2- or 3-dimensions, with smaller and nearly equal numbers as 0- or 1-dimensions (12.5 and 13.3%, respectively). Over half of the entity occurrences were not spatially referenced, and only 10.3% were absolutely referenced (Table 2).

In terms of dimensionality, the representations of groups of components and continuous components had similar distributions, with fewest as 0D and 1D, and three-quarters as 2D and 3D. Continuous components, however, had a much larger proportion of 3D studies than did groups of components (49.3% vs. 30.0%). For networks of components, there were a disproportionately large number of 1D entity occurrences relative to groups of components or continuous components (39.0% vs. 12.4% and 6.6%, respectively).

In terms of spatial referencing, the distribution of entity occurrences in the representation of groups of components contrasted with continuous components, with only 5.2% of the former vs. 41.4% of the latter being absolutely referenced. The distribution of spatial referencing of entity occurrences in networks of components was more similar to groups of components (2.6%). For the descriptor of reactiveness, entity occurrences for all three representations had fairly similar distributions, with more than half being non-reactive.

4. Discussion

4.1. Analysis of the framework

Our framework successfully categorized forest structural units, as it captured >99% of units within 500 forest structure studies that measured attributes of forest structure at varying scales and locations. Our framework was not completely exhaustive, as two of our study units (that involved fractals) did not fit within any of our categories. Also, there were nine entity categories into which none of our study units were categorized. Explanations for the empty categories were (1) the particular subset of studies we identified might not have included all possible categories (Table 3); and (2) certain configurations and combinations might simply not be appropriate delineators of forest structure, just as certain forms of luggage might not exist.

Within the continuous components representation at least two categories are unlikely to ever be populated by real-world forest study units. In general, 0D, non-reactive, continuous components are improbable because it is unlikely that one would treat a continuous component as dimensionless if it were not studied in conjunction with functional data. However, if an estimate of biomass were of interest, and if the forest was considered as one unit, then data from such a study would populate this category. On the other hand, it is most likely that some structural information would be collected to derive the biomass estimates at such a large scale, where biomass is likely to be estimated from allometric equations rather than through collection of samples for weighing. Thus, we predict that two categories ($CC_{0D\text{-}NS\text{-}NR}$, and $CC_{0D\text{-}AS\text{-}NR}$) will remain empty no matter how many studies are examined.

Although, we recognize the limits of using a sub-sample of studies to examine how forest structure studies distribute into our categories, these studies were not chosen with any preconceptions of their distribution within our framework. One clear pattern that emerged from this process was the predominance of studies treating forest structure units as groups of components. This is not surprising, because the most straightforward way to measure structure is to treat individual components independently. This also allows for the greatest flexibility in terms of statistical analysis.

There are logistical limitations that are often involved in recording absolute positions (e.g., getting GPS coordinates under dense forest cover, resolution of GPS units in the vertical plane), which may explain why absolute spatial referencing was far less common (10.3% of the entity occurrences) than no referencing or relative referencing among our study units (53.1% and 36.6%, of the study units, respectively). In general, absolute location might not be critical to a study, given the number of studies based on a single sampling at one point in time. For these types of studies, there is no need to relocate the exact same sample unit and so the work entailed in recording its absolute position is not warranted. In contrast, if a researcher were interested in re-measuring the dynamics of specific structural components over time, as with studies of tree and forest dynamics (e.g., Matelson et al., 1995), then absolute spatial referencing would be critical. However, a grid system not related to any external features could be used to relocate and resurvey as a substitute for absolute spatial referencing. We expect the number of studies with absolute spatial referencing to increase in the future, given increasing affordability, precision, and accuracy of technology, and the importance of long-term monitoring.

The sizes and numbers of the sample units also influence spatial referencing. A researcher might collect absolute spatial data for a site location using GPS, but not for the location of each needle or twig. Absolute spatial referencing might be more readily applied in larger scale studies that involve, for example, a LIDAR flight path along which GPS coordinates are recorded that track the path. However, in some situations, the absolute positioning for small or abundant sampling units might be warranted. For example, when locating dwarf mistletoe infections at a forest canopy crane site, the absolute location of each infection was recorded so that follow-up visits could be conducted (Shaw and Weiss, 2000).

4.2. Use of the framework by forest ecologists

We developed a set of questions for researchers to answer when placing their own entities into our framework, and here we provide an example of how the protocol works for a researcher interested in the effects of mistletoe infection on tree productivity. Let's say she maps sample trees in her stand on a geo-referenced *x*, *y* coordinate system established on the forest floor. Using a laser range-finder on the ground that measures distances to the point of branch attachment, and using binoculars to document which branches are infected with mistletoe, she generates data that quantify, for each branch, the height at which it is attached to the bole, and whether it is infected or not.

To categorize the entities that result from her study in our framework, she would answer the following six questions, in this order: (1) *What is your research hypothesis*? (the distribution of mistletoe infections is not related to height, i.e., branches with mistletoe are randomly distributed with respect to height. (2) *What is/are the specific component(s) of forest structure that relate to your research questions*? (boles, branches). (3) *Which representation best fits your components*? (boles – groups of components, as stems are independently measured; branches – networks of components, as individual branches are linked to individual trees and their relative vertical location is relevant for the research hypothesis); (4) *Which dimensionality best fits your component(s)*? (boles and branches are both zero-dimensional, i.e., no structural data recorded for them) (5) *Which spatial referencing best fits your component(s)*? (boles – absolutely spatially referenced (geo-referenced *x*, *y* ground coordinate point, branches – relatively spatially referenced with respect to vertical position along the bole)). (6) *Which reactiveness best fits your component(s)*? (boles – non-reactive; branches – reactive with respect to presence of mistletoe infection). Her responses will allow her to match each of her entities into one of the three cubes in Fig. 4 (boles – $\mathrm{GC}_{\mathrm{0D\text{-}AS\text{-}NR}}$; branches – $\mathrm{NC}_{\mathrm{0D\text{-}RS\text{-}RE}}$). From this process, she would be able to reflect upon how common or uncommon her entities are relative to the larger body of other forest structure studies that have already been categorized. This would also inform her of other studies with entities that occupy the same category or comparable categories, to which she could connect her own study.

4.3. Insights in forest ecology research

Our framework to categorize perceptions of forest structure is different from previous systems to understand forest structure (e.g., Webb et al., 1967; Hallé et al., 1978). Our framework helps us understand how researchers perceive and understand forest structure in five ways.

First, many past approaches required ecologists to make standardized field measurements of forest components, which potentially constrained their resources and proclivities. Ecologists generally remain faithful to a known method rather than accepting a new protocol even if they have access to the same equipment or technology proscribed by another researcher. Our approach provides awareness of and identifies the array of choices a researcher has available.

Second, researchers who are aware of how components of a forest of interest have been categorized by other studies may be better able both to design studies that are comparable to others and to carry out comparative studies themselves. Cataloging existing and current study entity occurrences within a unique category makes it easier for researchers to find relevant and semantically comparable datasets, and to know how to transform data to make the datasets comparable (see below).

Third, placing forest structure studies into standardized categories might help researchers understand the relationships between forest structure and forest functions such as interception, retention, modification, and transfer of energy and materials through the canopy to the forest floor. For example, by identifying an extensive group of interception studies and assigning our framework's addresses to all of them, patterns in the way in which interception studies treat the forest canopy may emerge and lead to new insights about relationships between forest structure and its functional roles in ecosystem processes. For example, Nadkarni and Sumera (2004) examined the literature on the effects of forest structure on throughfall volume, and found that most researchers treated the forest as groups of components beneath which throughfall was collected. They then used this approach in designing their study and documented that the amount and types of structural components above a given ground location affect the volume of throughfall at that point.

Fourth, by helping a researcher identify which approaches are most often used (and for which the most literature is available), students or novice researchers can follow those approaches and be fairly certain of finding references and existing data to help design a successful study, and connect theirs to others. Because this framework also points out the less common approaches, a scientist who chooses a less commonly used category would be aware he/she is taking more of a gamble, but one which might lead to new insights because fewer publications have explored those designs and the questions they address.

Fifth, this framework may help ecoinformatics researchers building software and visualization tools for canopy researchers (e.g., Nadkarni and Cushing, 2001; Cushing et al., 2007; http://canopy.evergreen.edu). Functionality for calculating certain derived measurements or aggregate values could be provided for particular categories, and researchers would see those values automatically. For example, data collected from a 3D spatially explicit group of components (such as foliage vertically arrayed in a deciduous forest) would be amenable to the direct calculation of leaf area index, whereas data from a non-spatially explicit dataset would not. Formulae and calculations that one researcher develops for his or her study units could potentially be readily integrated with other researchers' datasets that fall within that category.

By inspection, we identify which forest space categories can be most easily linked to others. For example, we recognize that it

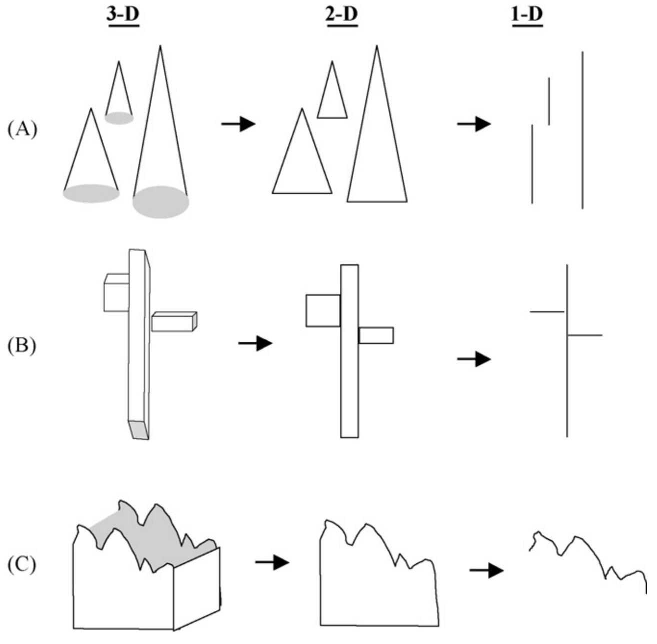


Fig. 5. Representation of examples of how dimensionality can be "collapsed." If a study has information collected with a greater number of dimensions, then there is the potential to collapse the information to a reduced number of dimensions to compare it with other studies of reduced dimensionality. This can be done for (A) a group of components, (B) network of components, and (C) continuous components.

would be fairly easy to “collapse” data from a 3D continuous component to be congruent with data collected in a 2D continuous component (Fig. 5). However, one cannot readily transform data in the converse direction. Similarly, collapsing a spatially referenced set of data by transforming it to spatially non-referenced data allows for a good match, but going in the opposite direction will usually not work. This is also possible on a case-by-case basis for reactiveness (e.g., a reactive group of components transformed to a non-reactive group of components). Although less intuitive, there may be opportunities where study entities could actually shift “upwards” (e.g., 2D to 3D) rather than collapse to a smaller dimensionality study (e.g., 3D to 2D), if a researcher is willing to introduce some uncertainty when comparing two data sets. For example, a researcher who collected tree diameter data and had allometric equations for heights could use the diameter values to produce estimates of tree heights. These data could then be compared with other datasets where 3D height and diameter data were both recorded.

The distribution of studies within our framework can be used to prioritize the implementation of database packages, visualization and statistical tools by targeting the categories used by the greatest number of researchers. Better tools, applicable to a large range of real-world studies, with faster development cycles, would improve the ability to manage and understand the complex spatial and temporal data related to forest structure and the ability to conduct synthetic research with data from multiple sites, time scales, or researchers.

5. Summary and conclusions

We present a categorization framework for the ways that ecologists perceive forest structure and use it to categorize how forest structure is reported in the literature. The framework encompasses a wide spectrum of spatial dimensions and research approaches, but we have focused on the conceptual aspects of forest structure. Other considerations are how the researcher takes actual structural measurements, and the ways in which these measurements are visualized or even analyzed and reported. Thus the collected data and certain information artifacts might constitute different categories than the conceptual view of the researcher’s unit of interest. For example, Parker et al. (1989) conceived the upper envelope of the canopy as a 1D continuous component, but they actually took the measurements as a group of points for which they recorded the heights, which would fall into the category of a 1D group of components. The visualization falls into the category of a 2D spatially referenced continuous component. Thus, in further refinements of this framework, we might assign the measurement and the visualization or conceptual view of each study component to separate categories.

The scientific community recognizes that researchers must enhance their capacity to carry out synthetic work (e.g., Collins et al., 2003). Our categorization system might help implement this directive, by guiding researchers to the appropriate avenues to gather and exchange data, and to identify patterns at a variety of spatial and temporal scales. This can enhance understanding of what the data mean, and conceptualize how real-world entities of interest inter-relate. Without such a framework, scientists are left with collections of data with little meaning beyond the small sphere of the particular location or point in time in which they were collected. Our framework assists in particular with one step of synthesis, the gathering and exchange of data across sites, time scales, researchers, and forest types.

In the future, we will expand this framework in three ways: (1) explore the implications of how conceptual views of components evolve as a researcher moves through the research process; (2) further explore how easy the framework is for researchers to learn and use, and whether different researchers would categorize different studies the same way within our framework; (3) examine dynamic as well as static elements of forest structure. We continue to seek new examples of forest structure units that populate our empty categories or fall outside of our present framework, and ways in which we might adapt our framework to encapsulate units that currently fall outside of its bounds.

Acknowledgments

We thank Geoffrey G. Parker, whose ideas on forest canopy structure were a seminal contribution to this paper. Barbara Bond, Lois Delcambre, Roman Dial, and David Maier contributed to initial concepts. Discussions about these ideas were initiated with participants at the NSF-supported Forest Canopy Structure Workshop (25–26 April 2002). Michael Finch, Young Mi Kim, Steve Rentmeester, Janet Rhoades, Abraham Svoboda, and Lee Zeman provided technical help. Roman Dial and Carri LeRoy gave guidance on an earlier draft. We are grateful to three anonymous reviewers for valuable feedback on the paper. This work was supported by the National Science Foundation grants and Research Experience for Undergraduate Supplements: DEB 05-42130; BDI 04-17311, 03-019309, 99-75510; BIR 99-74035, 96-30316; and INT 99-81531. We recognize the Helen R. Whiteley Center at Friday Harbor Laboratories, University of Washington, for support.

References

CCOD, 2008. Current Contents on Diskette, Environment, Ecology and Agriculture Series, Institute for Scientific Information, Philadelphia, PA.

Clark, K.L., Nadkarni, N.M., Schaefer, D.A., Gholz, H.L., 1998. Atmospheric deposition and net retention of ions by the canopy in a tropical montane forest, Monteverde, Costa Rica. Journal of Tropical Ecology 14, 27–45.

Collins, F.S., Morgan, M., Patrinos, A., 2003. The human genome project: lessons from large scale biology. Science 300, 286–290.

Cushing, J.B., Nadkarni, N.M., Finch, M., Fiala, A.C.S., Murphy-Hill, E., Delcambre, L., Maier, D., 2007. Component-based end-user database design for ecologists. Journal of Intelligent Information Systems 29, 7–24.

Dial, R., Bloodworth, B., Lee, A., Boyne, P., Heys, J., 2004. The distribution of free space and its relation to canopy composition at six forest sites. Forest Science 50, 312–325.

Drake, J.B., Dubayah, R.O., Knox, R.G., Clark, D.B., Blair, J.B., 2002. Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. Remote Sensing of Environment 81, 378–392.

Ellsworth, D.S., Reich, P.B., 1993. Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. Oecologia 96, 169–178.

Emborg, J., 1998. Understory light conditions and regeneration with respect to the structural dynamics of a near-natural temperate deciduous forest in Denmark. Forest Ecology and Management 106, 83–95.

Forestry Abstracts, 1976–2003. CAB International, London, U.K.

Franklin, J.F., Spies, T.A., Van Pelt, R., Carey, A.B., Thornburgh, D.A., Berg, D.R., Lindenmayer, D.B., Harmon, M.E., Keeton, W.S., Shaw, D.C., Bible, K., Chen, J., 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. Forest Ecology and Management 155, 399–423.

Fujimori, T., 1971. Analysis of forest canopy on the basis of a *Tsuga heterophylla* forest. Japanese Journal of Ecology 21, 134–140.

Gratani, L., Foti, I., 1998. Estimating forest structure and shade tolerance of the species in a mixed deciduous broad-leaved forest in Abruzzo, Italy. Annales Botanici Fennici 35, 75–83.

Groffman, P., Tiedge, J., 1989. Denitrification in north temperate forest soils: spatial and temporal patterns at the landscape and seasonal scales. Soil Biology and Biochemistry 21, 613–620.

Grubb, P.J., Lloyd, J.R., Pennington, T.D., Whitmore, T.C., 1963. A comparison of montane and lowland rain forest in Ecuador. I. The forest structure, physiognomy and floristics. Journal of Ecology 51, 567–601.

Hallé, F., Oldeman, R.A.A., Tomlinson, P.B., 1978. Tropical trees and forests: an architectural analysis. Springer-Verlag, New York, NY.

Harrington, C.A., DeBell, D.S., 1996. Above- and below-ground characteristics associated with wind toppling in a young *Populus* plantation. Trees 11, 109–118.

Imhoff, M., Story, M., Vermillion, C., Khan, F., Polcyn, F., 1986. Forest canopy characterization and vegetation penetration assessment with space-borne radar. IEEE Transactions on Geoscience and Remote Sensing Ge-24, 535–542.

Kinerson Jr., R.S., Fritschen, L.J., 1971. Modeling a coniferous forest canopy. Agricultural Meteorology 8, 439–445.

Kneeshaw, D.D., Burton, P.J., 1997. Canopy and age structures of some old sub-boreal *Picea* stands in British Columbia. Journal of Vegetation Science 8, 615–626.

MacArthur, R.H., Horn, H.S., 1969. Foliage profile by vertical measurements. Ecology 50, 802–804.

Martens, S.N., Ustin, S.L., Norman, J.M., 1991. Measurement of tree canopy architecture. International Journal of Remote Sensing 12, 1525–1545.

Matelson, T.J., Nadkarni, N.M., Solano, R., 1995. Tree damage and annual mortality in a montane forest in Monteverde, Costa Rica. Biotropica 27, 441–447.

McIntosh, A.C.S., Cushing, J.B., Nadkarni, N.M., Zeman, L., 2007. Database design for ecologists: composing core entities with observations. Ecological Informatics 2, 224–236.

Nadkarni, N.M., Cushing, J.B., 2001. Lasers in the jungle: the forest canopy database project. Bulletin of the Ecological Society of America 82, 200–201.

Nadkarni, N.M., Parker, G.G., 1994. A profile of forest canopy science and scientists—who we are, what we want to know, and obstacles we face: results of an international survey. Selbyana 15, 38–50.

Nadkarni, N.M., Sumera, M., 2004. Old-growth forest canopy structure and its relationship to throughfall interception. Forest Science 50, 290–298.

Nelson, R.F., Krabill, W., Maclean, G., 1984. Determining forest canopy characteristics using airborne laser data. Remote Sensing of Environment 15, 201–212.

Nychka, D., Nadkarni, N.M., 1990. Spatial analysis of points on tree structures: the distribution of epiphytes on tropical trees. University of North Carolina, Institute of Statistics Mimeograph Series No. 1971, Raleigh, NC, USA.

Parker, G.G., 1995. Structure and microclimate of forest canopies. In: Lowman, M.D., Nadkarni, N.M. (Eds.), Forest Canopies. Academic Press, Incorporated, San Diego, CA/Orlando, FL/Troy, MO, pp. 73–106.

Parker, G.G., Davis, M.M., Chapotin, S.M., 2002. Canopy light transmittance in Doulgas-fir-western hemlock stands. Tree Physiology 22, 147–157.

Parker, G.G., O'Neill, J.P., Higman, D., 1989. Vertical profile and canopy organization in a mixed deciduous forest. Vegetatio 85, 1–11.

Pike, L.H., Rydell, R.A., Denison, W.C., 1977. A 400-year-old Douglas-fir tree and its epiphytes: biomass, surface area, and their distributions. Canadian Journal of Forest Research 7, 680–699.

Putz, F.E., Coley, P.D., Lu, K., Montalvo, A., Aiello, A., 1983. Uprooting and snapping of trees: structural determinants and ecological consequences. Canadian Journal of Forest Research 13, 1011–1020.

Reunanen, P., Mönkkönen, M., Nikula, A., 2000. Managing boreal forest landscapes for flying squirrels. Conservation Biology 14, 218–226.

Rinne, J., Hakola, H., Laurila, T., Rannik, Ü., 2000. Canopy scale monoterpene emissions of *Pinus sylvestris* dominated forests. Atmospheric Environment 34, 1099–1107.

Robertson, G., 1987. Geostatistics in ecology: interpolating with known variance. Ecology 68, 744–748.

Runkle, J.R., 1990. Gap dynamics in an Ohio *Acer-Fagus* forest and speculations on the geography of disturbance. Canadian Journal of Forest Research 20, 632–641.

Shaw, D.C., Weiss, S.B., 2000. Canopy light and the distribution of hemlock dwarf mistletoe (*Arceuthobium tsugense* (Rosendahl) G.N. Jones ssp. tsugense) aerial shoots in an old-growth Douglas-fir/western hemlock forest. Northwest Science 74, 306–315.

Sumida, A., 1995. Three-dimensional structure of a mixed broad-leaved forest in Japan. Vegetatio 119, 67–80.

Sumida, A., Terazawa, I., Togashi, A., Komiyama, A., 2002. Spatial arrangement of branches in relation to slope and neighborhood competition. Annals of Botany 89, 301–310.

Tanaka, T., Yamaguchi, J., Takeda, Y., 1998. Measurement of forest canopy structure with a laser plane range-finding method—development of a measurement system and applications to real forests. Agricultural and Forest Meteorology 91, 149–160.

Valinger, E., Fridman, J., 1997. Modelling probability of snow and wind damage in Scots pine stands using tree characteristics. Forest Ecology and Management 97, 215–222.

Webb, L.J., Tracey, J.G., Williams, W.T., Lance, G.N., 1967. Studies in the numerical analysis of complex rain-forest communities. I. A comparison of methods applicable to site/species data. Journal of Ecology 55, 171–191.

Zeide, B., 1998. Fractal analysis of foliage distribution in loblolly pine crowns. Canadian Journal of Forest Research 28, 106–114.

Zeide, B., Pfeifer, P., 1991. A method for estimation of fractal dimension of tree crowns. Forest Science 37, 1253–1265.