Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

A new metric for measuring the conservancy of plant species

Andrew Gard[1] | **Glenn Adelson**[2] | **Ryan Sorrells**[3] | **Irene Luwabelwa**[3] | **Anna Evans**[3]

[1]Department of Mathematics and Computer Science, Lake Forest College, Illinois, U.S.A.

[2]Department of Environmental Studies, Lake Forest College, Illinois, U.S.A.

[3]Lake Forest College, Illinois, U.S.A

Correspondence
agard@lakeforest.edu

Abstract

The coefficient of conservatism, or C-value, is a widely-used measure of plant conservancy. Assigned on a 0-10 scale by experts in local flora, this metric is used in the United States and elsewhere in floristic quality assessment (FQA), a standardized process for rating the ecological value of natural areas based on the plant species found within them. The subjective and sometimes pejorative nature of C-values is a common basis for criticism of this technique. We introduce *co-occurrence analysis* and the *co-occurring native mean-C* measure as more robust tools for rating the conservancy of plants and demonstrate how such tools can be applied to identify potentially misclassified species. Co-occurrence analysis also provides a methodology for targeted data collection for under-represented species of interest in a local ecosystem, as we illustrate with field studies in the Chicago, U.S.A. region.

KEYWORDS
floristic quality assessment, coefficients of conservatism, evaluation of natural areas, quantitative ecology

1 | INTRODUCTION

As awareness of human impacts on natural ecosystems has grown over time, so has the demand for practical tools to assess those impacts at the local level. This demand takes many forms, including land management, regulation, conservation research, and grant-seeking (Spyreas, Meiners, Matthews, & Molano-Flores, 2012). While simple qualitative evaluation is obviously insufficient on its own for any of these purposes, quantitative assessment of natural sites is difficult given the inherently nebulous notion of ecological integrity and our incomplete understanding of the complex and confounded set of variables that feeds into it.

While some degree of subjectivity is unavoidable in any system of evaluation for natural areas, stability, consistency, and transparency are nonetheless achievable goals. *Floristic quality assessment (FQA)*, first developed in the late 1970s by Gerould Wilhelm (G. Wilhelm, 1977) and others in the Chicago, USA region, attempts to address each of these considerations. Each species known to be found in a particular ecological region is assigned a *coefficient of conservatism*, C, on a scale of 0-10, measuring the plant's perceived tendency to grow in undisturbed habitats. The practitioner gathers an inventory of plant species found at the site under consideration, then computes summary statistics using their C-values and related metrics. Chief among these is the *native mean-C*, or average C-value of all identified native species, followed by the *floristic quality index (FQI)*, which weights this mean-C by the square root of the total number of native species identified.

While no such measure is above reproach, this system at least, “places the subjectivity up front, in the a priori assignment of conservatism coefficients to each of our native plant species," according to Wilhelm (G. S. Wilhelm & Masters, 1995). "The goal was to design a system, based upon plants, that assesses natural quality repeatably and dispassionately, facilitates comparisons among sets of sites, and tracks changes in site quality over time." Whatever their flaws, C-values are at least stable, consistent, and transparent.

At a practical level, floristic quality assessment has several distinct advantages. As described by Spyreas (Spyreas, 2019), it combines simplicity and ease of use with flexibility and wide applicability. Anyone who can identify plant species (either by eye or with technological assistance) and who has access to a local database of C-values can perform an FQA, with no additional expertise or lab equipment needed.

Since its original conception, floristic quality databases of locally meaningful C-values have been compiled and put into use across the United States, Canada, and beyond (see for instance (Ladd & Thomas, 2015; Oldham, Bakowsky, & Sutherland, 1995; Gianopulos, 2014)). At the time of writing, the popular online repository UniversalFQA.org (Freyman, Masters, & Packard, 2016), discussed in Section 2.2, included a total of 85 such databases.

Floristic quality assessment is a fundamentally local, empirical, and biotic approach, based on the best experience of the botanists who have assigned C-values to plants in their home region. It makes no specific assumptions about the biological characteristics of high-conservancy plants, eschewing theoretical notions for simple human statements of the form, “I usually see this plant in remnant ecosystems" and, “this species even thrives in sidewalk cracks." This approach is both its strength and weakness.

Botanists assigning these C-values are subject to both random and systematic misjudgement. In flipping through any floristic quality database, a knowledgable reader will inevitably find individual C-values that miss the mark based on their own experience, sometimes radically. For instance, in the 2017 Chicago database (G. Wilhelm & Rericha, 2017) used in the analysis below, *Silene nivea*, snowy campion, carries a C-value of 10 but is quite often found in roadside swales filled with non-native weeds and native plants of low C-value. *Polanisia dodecandra*, clammyweed, on the other hand, has a C-value of 0 yet is quite often found as an integral part of remnant gravel-hill prairie ecosystems.

Ecologically irrelevant factors have sometimes prejudiced the assignment of C-values (Taft, Wilhelm, Ladd, & Masters, 1997). Rarer and more beautiful species tend to have higher C-values, not because they are more ecologically valuable but because practitioners have subconsciously assumed them to be more important. Conversely, common species are sometimes undervalued, even when they are known to be found in remnant ecosystems.

Each of these flaws in the initial assignment of C-values is mitigated by the averaging process of floristic quality assessment. A typical site evaluation includes a hundred or more distinct observed species, so a handful of misclassifications are unlikely to be influential in the final computation of a site's native mean-C or FQI. This is particularly true when a single site is considered longitudinally to track conservation or mitigation efforts.

Floristic quality assessment has been validated by numerous authors using a variety of methodologies. For instance, Bourdaghs, et al (Bourdaghs, Johnston, & Regal, 2006), Cohen et al (Cohen, Carstenn, & Lane, 2004), and Miller & Wardrop (Miller & Wardrop, 2006) have demonstrated correlations between mean-C and established measures of overall site quality in Michigan, Florida, and Pennsylvania, respectively. Bried et al (Bried, Jog, & Matthews, 2013) further show an association between mean-C and site condition across surveys and settings in upstate New York. Other studies (Zinnen, Spyreas, Zaya, & Matthews, 2021; Bauer, Koziol, & Bever, 2018) have connected C-values with relevant biological properties of species, with interesting results.

Most important of all is the work of Matthews, Spyreas, and Long (Matthews, Spyreas, & Long, 2015), which validates C-values by considering co-occurring species. "If subjectively assigned species' C-values carry meaningful information about plant assemblages and the conservation value of particular habitats, then individual species should tend to co-occur with species of similar C," they write. They find statistically significant evidence ($P < 0.001$) against a null hypothesis that C-values carry no information about co-occurring species in Illinois. On its own, this result is not surprising. More interesting and useful is their approach, which we build on here.

Co-occurrence analysis describes a plant's conservancy using the overall distribution of C-values of plants that it has been identified alongside. By taking into account hundreds or even thousands of co-occurring species among multiple instances of each target species, co-occurrence analysis allows plants and their habitats to speak for themselves. This methodology builds on the strengths of the mean-C metric (stability, consistency, and transparency), which has already been repeatedly validated, while addressing the flaws in individual assignments. A misclassified species will inevitably reveal itself through the plants that it associates with.

2 | MATERIALS AND METHODS

2.1 | Overview

Just as a tract of land can be assessed based on the plant species found there, so individual species can be assessed based on the plants found in proximity. Broadly speaking, one should expect to find highly-conservative species alongside plants with high C-values and less-conservative species alongside plants with low C-values. A plant whose C-value differs substantially from those it tends to co-occur with may be misclassified and should be reconsidered.

In this section, we present a specific and reproducible methodology for computing a measure of plant conservancy, the *co-occurring native mean C-value (CNMC)*, as well as a more qualitative tool, the *native co-occurrence profile*, the full numerical distribution of such C-values. These promise to be more robust measures of conservancy than assigned C-values of plants themselves.

The correct interpretation of the C-value is a topic much debated in the literature, with different authors ascribing drastically different meanings. See Zinnen, et. al. (Zinnen et al., 2021) for a summary of these perspectives, which we do not explore here. It is to be hoped that co-occurrence analysis will ultimately inform our understanding of C-values and of plant species themselves rather than being shaped by them.

2.2 | UniversalFQA.org

Co-occurrence analysis requires extensive data on observed associates of plant species in an ecological region, a need met by the online Universal FQA Calculator (`https://universalfqa.org/`). Established in 2013 by Openlands of Northeastern Illinois, this tool streamlines data management for field practitioners, allowing them to input the plant inventories observed during site assessments and have database information (including C-values) added automatically. Important site-level metrics such as mean-C and FQI are returned as well. Data submitted to this website is public by default, providing access to thousands of site assessments from nearly 100 regional databases.

Data from Universal FQA is of course imperfect. A tremendous diversity of practitioners using a wide variety of methodologies are represented. Some assessments there are culminations of deep and careful fieldwork, while others represent only cursory glances. A few are simply tests of the system of itself, where a prospective user has created a fictional assessment as a way to learn the tool.

None of these shortcomings need trouble us greatly. Test inputs typically consist of only a few observations representing a negligible numbers of co-occurrences, while CNMC-values are typically computed over hundreds of observations. The averaging process provides substantial protection against occasional data entry errors, whether those be deliberate or unintentional.

Co-occurrence analysis is largely indifferent to techniques used in data collection. As long as two species are found simultaneously by a practitioner who chooses to include them both in a single site assessment, they are taken to be co-occurring for the purposes of this study. This respect for the expert judgement of the local practitioner is a particular strength of the CNMC metric. While a spacial analysis might sound attractive, the practical issues involved are daunting.

For this project, we elected to focus on the Chicagoland region, where C-values were first developed. The updated version of that original database, *Flora of the Chicago Region* (G. Wilhelm & Rericha, 2017), has been used in over 400 separate inventories on universalFQA with over 21,000 observations and millions of total co-occurrences. While other databases are available in that region, the FCR remains the standard among experienced botanists in the region. We expect that an analysis of any overlapping database would yield very similar results due to the highly robust nature of co-occurrence analysis.

2.3 | The `fqar` package

Data analysis was performed using the R programming language (R Core Team, 2022) and the `fqar` package (Gard, Myers, & Luwabelwa, 2024), which enables efficient programmatic access to `universalFQA.org`. This package includes four categories of functions, all of which are relevant to the current project.

- *Indexing* functions provide summary information about databases and assessments available on `universalFQA.org`.
- *Downloading* functions import assessments individually and collectively according to specified search criteria.
- *Tidying* functions convert those assessments to a standard format suitable for data analysis.
- *Analytic* functions build co-occurrence summaries and compute co-occurrence profiles.

In this project, we downloaded every public assessment on `universalFQA.org` that used the 2017 Flora of the Chicago Region database. Two different inventory formats are included in that website (depending on whether the practitioner wished to include quadrat-level data or not), both of which were included in our consideration. After data preprocessing, we constructed a single data set including details of every observed species co-occurrence. The resulting set of co-occurrences is available at [removed for anonymity]. Full code for replicating this analysis can be found on that site as well as in the appendix to this paper.

The analysis described below includes only native plants which were included in at least three separate assessments. While consideration of non-native species is an important area for research, as discussed in section 4, the blanket numerical assignment of a C-value of zero to such taxa introduces complications that place it beyond the scope of the current analysis.

This filtered data set is also available on the site listed above. Data from `universalFQA.org` used in this project was accessed on 21 January 2025.

2.4 | Co-occurrence profiles and the CNMC

With this data set in hand, we were able to obtain a *co-occurrence profile* for every species that has been observed in the Chicago region, that is, a distribution of the C-values of all the plants that have been identified in proximity to each target species. For these purposes, each species is only counted once per assessment (so two species cannot co-occur multiple times at a single site) but can be repeated across the database (so two species can co-occur multiple times across different sites). This gives a fuller picture of each species' co-occurrence profile while also allowing for possible future consideration of pairwise occurrence trends. See section 4 for more on this.

As an example, the co-occurrence profile of *Solidago canadensis*, the Canada goldenrod, which is considered native in the Chicago region and has been assigned a C-value of 1, is shown in Figure 1.

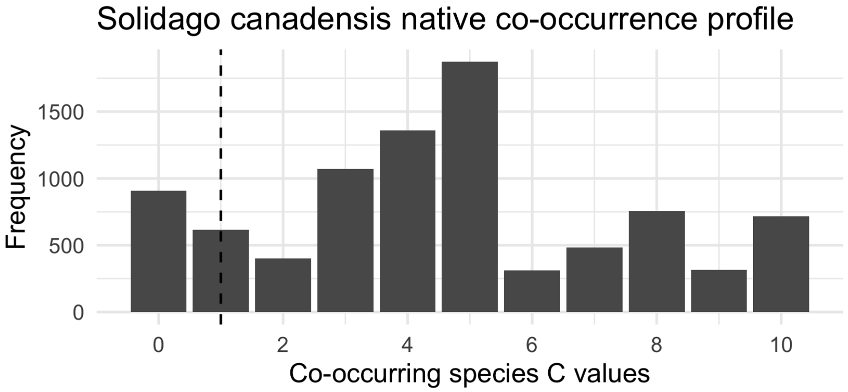

FIGURE 1 Co-occurrence profile of the Canada goldenrod

This particular profile represents sightings at 104 distinct assessments and reflects over 8,800 total co-occurrences. From the plot, it appears that *Solidago canadensis* is relatively indifferent to its neighbors and is nearly as likely to co-occur with 10's as with 0's, despite its assigned C-value of 1.

An even more pithy summary of these results is provided by the co-occurring native mean-C, or CNMC, the simple average of all C-values of plants that the species has been identified near. In the Chicago database, *Solidago canadensis* has a CNMC of 4.6. While the Canada goldenrod may not be considered a particularly conservative plant, it is frequently found near species that are.

2.5 | Targeted data collection

Of the 1876 native species listed in the FCR database on universalFQA.org, only 1380 (73.6%) had been identified in at least one public assessment at the time of writing. Only 926 (49.4%) had been identified in at least three. Clearly there are wide gaps in our knowledge about which plants are actually currently found in the Chicago region and what their associates tend to look like. Co-occurrence analysis suggests a methodology for filling such gaps.

A practitioner interested in a particular species that has so far not appeared in many floristic quality assessments can, upon finding that species in the wild, conduct a survey of nearby plants and begin to build a species profile from scratch. After locating several exemplars of the species with a reasonable number of associates, that profile can be considered to be valid. This approach is not without its drawbacks, of course. Potential sources of error are discussed in section 4.

As a part of this study, we conducted field work between May and August, 2023, documenting CNMC-values for species lacking robust data in the Universal FQA database. Throughout the season, we covered 30 taxa, with a mix of more- and less-frequently encountered species. The purpose was to develop a methodology for CNMC studies analogous to other standardized forms of data collection for plants, while attempting to fill some of the gaps in the Universal FQA datasets for the Chicago Region.

Our standard method was to first locate the target species in the wild, then to randomly select three individuals within different subpopulations not bordering hazardous obstacles or footpaths (the exception being *Castilleja sessiliflora*, which required data from two sites to reach a total of three plots). We placed these individuals at the center of a square one meter by one meter plot, physically represented by a PVC pipe square. Every identifiable species within the plot was noted, and the list was later transferred into Universal FQA and to the project's GitHub. This data provides a starting point for calculating the CNMC of these species in the Chicago Region.

The results presented in Section 3.4 show that our method can be used for targeted data collection within specific populations and could be usefully integrated into existing FQA and rare plant monitoring efforts, such as those at Midewin National Tallgrass Prairie (U.S. Forest Service, 2024). A more robust dataset collected annually could stand as an objective metric representing changes in habitat quality over time for plants of concern or related ongoing studies. Of the various habitats studied, the most difficult challenges encountered were hike-out time with materials and overlying woody plants, but neither of these issues represent obstacles difficult enough to warrant special procedures.

3 | RESULTS

The set of native co-occurrences in the FCR database described in section 2 consists of 21,216 observations of 1380 distinct native species from 443 site assessments, reflecting a total of 1,658,275 co-occurrences. There are 1876 native species listed in the Flora of Chicago Region, so this database includes approximately 73.6% of listed species. Some of these, like *Solidago canadensis*, are commonplace, while others have been identified only once or twice in public assessments. Reflecting this, we considered only taxa that appeared in at least three separate assessments, a necessary precaution in order to avoid making conclusions about any species based on potential outlier sightings.

After filtering, our data set consisted of 20,599 observations of 926 plant species, 49.4% of those listed in the FCR, with a total of 1,558,476 native co-occurrences. Addressing the gaps in this record is the subject of Section 3.4.

3.1 | Overall trends

As one would expect, plant species with high assigned C-values tend to have high CNMC values. We observed a coefficient of correlation of $r = .60$, which is significant ($p < 2.16 \times 10^{-16}$) against a null hypothesis that the values are uncorrelated.

This corroborates the results of Matthews, Spyreas, and Long (Matthews et al., 2015) but is not in itself important. It would be remarkable indeed if assigned C-values were truly uncorrelated with the C-values of co-occurring species.

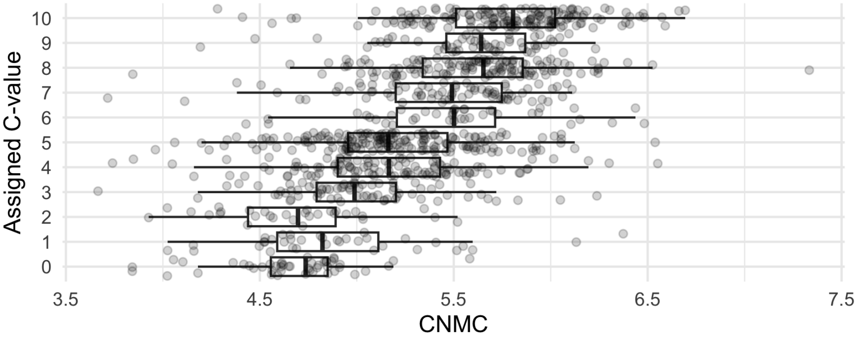

FIGURE 2 Higher C-values are associated with higher CNMC-values

On average, a single-unit increase in CNMC corresponded to an increase of approximately 3.26 in assigned C. Given that CNMC values are more tightly packed than assigned C-values due to the averaging process used to compute them, this is appropriate. In this data set, CNMC values ranged from 3.67 to 7.33 with a standard deviation of 0.54. By comparison, C-values for native species in the FCR range from 0 to 10 with a standard deviation of 3.11.

The overall trend line is given by the formula

$$C_{\text{adj}} \sim 3.26 \cdot \text{CNMC} - 11.45$$

While this formula is a useful rule of thumb for placing CNMC and native-C values on the same scale, individual species exhibit substantial variation and translating in this way should be done with caution. More on this in section 4.

3.2 | C-values of concern

We initially considered taxa whose assigned C-values differed substantially from the adjusted values, C_{adj}, predicted by their CNMC's, isolating the top 10 differences in each direction. We thus obtained a pair of lists of species tending to co-occur with species of notably different conservancy. These are displayed in tables 1 and 2.

While many of these species appeared only a small number of times, their respective co-occurrence counts (the total number of species occurrences they were identified near) are high enough for statistical consideration. Each of these species (and probably many more) should be looked at in more detail. The full database available on GitHub shows discrepancy values for all observed plants so the reader can easily consider any particular species of personal interest. Important considerations for doing so are included in section 4.

TABLE 1 Ten potentially overrated species

Species	C	CNMC	C_{adj}	Discrepancy	Co-occurrences	Sites
Helianthus giganteus	10	4.3	2.5	7.5	298	3
Carex crawei	10	4.4	2.9	7.1	699	3
Trillium sessile	10	4.6	3.4	6.6	611	6
Commelina erecta var. deamiana	10	4.7	3.7	6.3	580	5
Echinacea purpurea	10	4.7	3.8	6.2	5078	47
Trillium erectum	10	5.0	4.9	5.1	742	4
Carex lasiocarpa var. americana	10	5.1	5.1	4.9	1424	7
Echinacea pallida	10	5.1	5.2	4.8	4313	32
Matteuccia struthiopteris	10	5.2	5.3	4.7	926	5
Juglans cinerea	10	5.2	5.5	4.5	952	4

TABLE 2 Ten potentially underrated species

Species	C	CNMC	C_{adj}	Discrepancy	Co-occurrences	Sites
Turritis glabra	1	6.4	9.3	-8.3	1080	6
Smilax herbacea	1	6.1	8.5	-7.5	995	3
Epilobium ciliatum	0	5.6	6.7	-6.7	2193	15
Parthenocissus inserta	0	5.2	5.5	-5.5	3898	16
Lepidium virginicum	0	5.2	5.4	-5.4	2263	8
Equisetum arvense	0	5.0	5.0	-5.0	6639	38
Persicaria lapathifolia	0	5.0	4.8	-4.8	4405	22
Fragaria virginiana	0	4.9	4.6	-4.6	8651	75
Galium aparine	0	4.9	4.6	-4.6	6283	47
Eupatorium serotinum	0	4.9	4.5	-4.5	6047	39

3.3 | Co-occurrence profiles

No single metric can hope to fully capture the conservancy of plant species, and the CNMC is no exception. A co-occurrence data set empowers more granular examination via construction of *co-occurrence profiles* of the species it contains, that is, the numerical distribution of the C-values of plants found in proximity to the target species. In this section, we illustrate the value of this technique with a pair of examples drawn from tables 1 and 2.

The largest positive discrepancy observed in this data set corresponds to *Helianthus giganteus*, the giant sunflower, which has an assigned C-value of 10, a CNMC of 4.3, and an adjusted C-value of 2.5. By this metric, it is over-rated by 7.5 points.

One possible reason for the discrepancy is that *Helianthus giganteus* is a frequently misidentified species. Voss and Reznicek note that Helianthus, “is a notoriously difficult genus...with hybridization obscuring the differences between a number of species" (Voss & Reznicek, 2012). Many of the characteristics used to distinguish species from each other are highly qualitative and variable. *H. giganteus* is often confused with the less conservative *H. grosseserratus*, as both have relatively narrow leaves with shallow teeth and the tendency to have opposite leaves on the lower part of the stem and alternate leaves on the upper. Many keys distinguish them through qualitative statements about stem pubescence, which can be confusing late in the season when hairs have been physically abraded.

At the other extreme is *Turritus glabra*, tower mustard, which has an assigned C-value of 1, a CNMC of 6.4, and adjusted C-value of 9.3. By this metric, it is under-rated by more than 8 points.

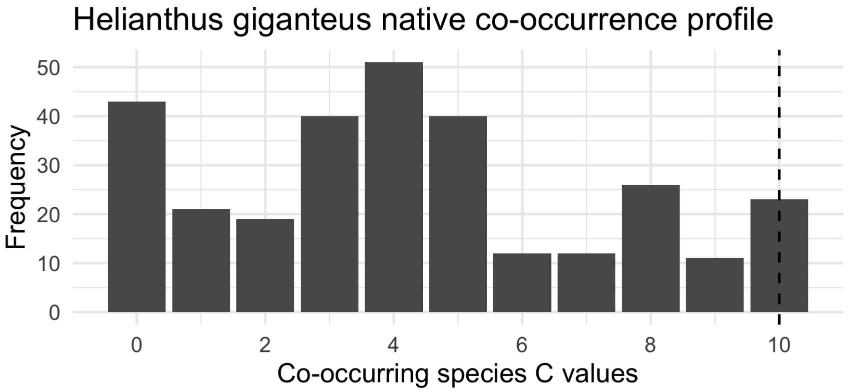

FIGURE 3 Co-occurrence profile of the giant sunflower

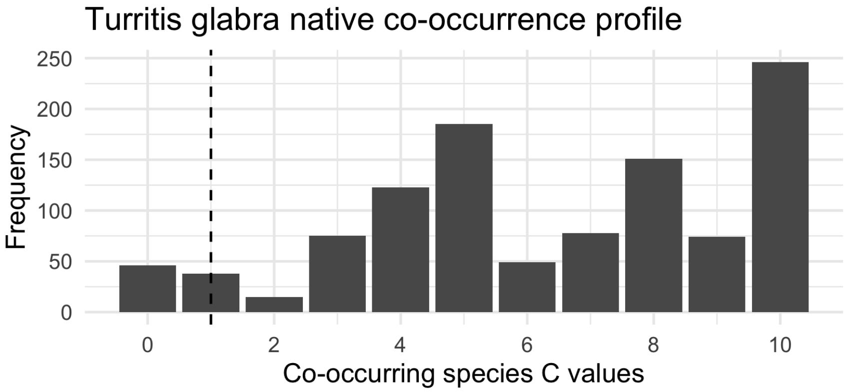

FIGURE 4 Co-occurrence profile of tower mustard

Turritis glabra may be mischaracterized for at least two distinct reasons. First, it is often found in prairie restoration seed mixes, which may have led botanists assigning its C-value to consider it a marker of low conservatism. Second, it tends to have two habitats: remnant sand prairies and open, sandy ground, and it's possible that observations in the database may have skewed toward the less weedy habitats. This phenomenon is discussed in Section 4.

3.4 | Data collection for previously under-documented species

Because our sample sizes are small and may represent aberrant cases, the following results from our 2023 study are to be taken as illustrative and not determinative. We hope these results spur others with knowledge of the locations of these species to perform similar tests and upload them to their FQA databases in order to increase the overall sample size and make the results more robust.

Data in Table 3 includes both our own observations and existing data on Universal FQA, following the standard practices described in Section 2. Adjusted C-values have been mapped to 0 in cases where the observed CNMC was below 3.33 and to 10 in cases where the observed CNMC was greater than 6.67 lest those adjusted values fall outside of the possible range of C-values.

TABLE 3 Results of targeted data collection

Species	Common Name	C	CNMC	C_{adj}	Discrepency	Co-occurrences	Total sites
Aletris farinosa	Colic root	9	5.7	7.1	1.9	814	5
Amorpha canescens	Leadplant	10	5.8	7.5	2.5	4888	32
Aralia racemosa	Spikenard	10	6.2	8.8	1.2	2984	15
Asclepias lanuginosa	Woolly milkweed	10	1.7	0.0	10.0	3	1
Cardamine pratensis var. palustris	Cuckoo flower	10	6.7	10.0	0.0	25	2
Carex crus-corvi	Raven's foot sedge	10	5.8	7.5	2.5	1724	13
Carex lupuliformis	False hop sedge	10	5.7	7.1	2.9	1014	5
Carex retrorsa	Retrorse sedge	10	6.6	10.0	0.0	569	2
Castilleja sessiliflora	Downy paintedcup	10	5.9	7.8	2.2	87	2
Cimicifuga racemosa	Black cohosh	10	5.6	6.8	3.2	1388	5
Conopholis americana	Cancer root	8	6.2	8.8	-0.8	1524	8
Cypripedium candidum	White lady's slipper	10	5.5	6.5	3.5	1356	4
Cypripedium parviflorum var. makasin	Small yellow lady's slipper	10	6.5	9.7	0.3	580	2
Cypripedium reginae	Showy lady's slipper	10	6.8	10.0	0.0	19	1
Dryopteris cristata	Crested wood fern	10	5.7	7.1	2.9	2058	9
Endodeca serpentaria	Virginia snakeroot	10	5.3	5.8	4.2	31	2
Euphorbia corollata	Flowering spurge	4	5.7	7.1	-3.1	4687	28
Glyceria septentrionalis	Floating manna grass	8	5.3	5.8	2.2	3230	18
Gymnocladus dioicus	Kentucky coffeetree	5	5.3	5.8	-0.8	1542	9
Iliamna remota	Kankakee mallow	10	5.9	7.8	2.2	396	3
Liparis loeselii	Fen orchid	4	5.4	6.2	-2.2	1896	5
Orobanche fasciculata	Clustered broomrape	10	6.2	8.8	1.2	32	2
Panax quinquefolius	Ginseng	10	5.9	7.8	2.2	1564	8
Platanthera psycodes	Lesser purple fringed orchid	10	6.5	9.7	0.3	582	2
Polygonatum pubescens	Hairy Solomon's seal	10	5.9	7.8	2.2	1980	11
Ranunculus flabellaris	Yellow water crowfoot	5	5.4	6.2	-1.2	3191	15
Sagittaria brevirostra	Midwestern arrowhead	7	5.2	5.5	1.5	619	5
Silene virginica	Fire pink	10	6.0	8.1	1.9	1888	10
Veronica scutellata	Marsh speedwell	9	5.8	7.5	1.5	1779	10
Viola subsinuata	Early blue violet	5	3.4	0.0	5.0	21	1

This data provides important preliminary insights into the ecology of the species tested and their associates and suggests many avenues for further research, as the following natural history information on several of the interesting results will illustrate.

Asclepias lanuginosa, which shares the largest overall discrepancy (10) with several other species, has unique ecological characteristics. Studies have shown that, at least in the southern portion of its range (which includes our study area) the species rarely sets viable seed and is likely a candidate for natural ecological extirpation. The site we tested was a degraded gravel hill prairie bordering a remnant that would have encompassed the studied plants at one point, but does not include them anymore, the likely reason for its unusually low CNMC (Kim, Zaya, Fant, & Ashley, 2015).

Iliamna remota, which has a discrepancy of 2.2, is another interesting case, being globally limited to a single site, a 20-acre island that, prior to restoration efforts, had been initially grazed and then overrun by Amur honeysuckle, *Lonicera maackii*. The restoration of that site proceeded by removing the honeysuckle and other native and non-native weedy species, but without introduction of any native species other than three or four species of grasses necessary to facilitate burning. Our observations are that *Iliamna remota* survived the decades of degradation by germinating on the island slopes inhospitable to grazing and honeysuckle, but once the habitat had been restored, it was capable of germinating and surviving in most of the island's upland habitats. One of the issues in determining C-values is whether a plant grows in a habitat that is a remnant as opposed to a restored area, and a case like this one shows that the range of habitats does not fall in to such a simple dichotomy.

Interestingly, our studied woodland plant species all had moderate to low CNMC values. Whether this is informative of the ecological state of woodlands in the Chicago Region would require additional study, but these values do provide an interesting juxtaposition to the high CNMC values of remnant wetlands and beach ridge plain species, such as *Sagittaria brevirostra* and *Aphyllon fasciculatum*.

4 | DISCUSSION

4.1 | Possible sources of error and abuse

No single metric can fully capture the ecological value of a plant species, which is by nature complex and mutable. While the CNMC is an important new measurement tool, it is subject to both misinterpretation and abuse. This section includes cautions about its limitations and suggestions for its appropriate use.

First, while it may be tempting to directly compare CNMC and C-values for individual species, this requires care. The former are *averages* of the latter, and so naturally occupy a more narrow range of values. Reversion to the mean is to be expected, and plants with assigned C-values near 0 or 10 will naturally have less extreme CNMC-values. *CNMC-values should not be mistaken for C-values, and should not be interpreted as such.* When direct comparisons are desired, adjusted C-values like those described in Section 3.1 should be used. When more focused evaluation of individual species is the goal, full co-occurrence profiles should be considered.

Species co-occurrence profiles and CNMC-values are computed using data collected by field practitioners at targeted sites, typically managed preserves, sites under remediation, or other areas of interest to local ecologists. Such locations likely have higher levels of conservancy, in general, than highway meridians or abandoned parking lots, for example. This represents a form of convenience sampling, an inevitable source of statistical bias, with the effect that plants that might sometimes be found in un-interesting locations are likely to show slightly elevated CNMC values. Co-occurrence analysis and the CNMC account for such selection by considering only *relationships* between species rather than *absolute* characteristics. Still, we must acknowledge that there must be trends to the sorts of relationships which field botanists do not tend to record and that these trends will make their way into any co-occurrence analysis.

Although targeted data collection, described in section 2.5, is also susceptible to such bias, it's not clear that the problem will be exacerbated by the targeting process. If anything, previously undocumented species are *less* likely to be found in highway meridians and abandoned parking lots than more common species, so their observed co-occurrence profiles may actually more faithfully reflect their true natures. More study is required as we work to understand such questions.

Of greater concern is the documented tendency for certain *kinds* of plants to have systematically higher or lower C-values. The process of averaging across many sites when computing co-occurrence profiles and CNMCs is of limited usefulness in reconciling C-values between species with little or no territorial overlap. As with floristic quality assessment more generally, caution must be exercised when comparing sites from different regions of fundamentally different ecological character. This flaw in C-value assignment is largely inherited by co-occurrence analysis and will need to be addressed in future iterations of species databases.

Finally, it must be noted that CNMC values are fundamentally mutable. As databases grow and change, so may the co-occurrence profiles of the species found within them. While such evolution will generally be gradual and modest, it still presents practical questions that the ecological community will need to answer as the CNMC metric is adopted. At a minimum, longitudinal studies should fix their reference databases at the start and not allow them to change over time.

4.2 | New directions for ecological research

The co-occurrence methods introduced here open many new avenues for ecological research. The most obvious is the potential for systematic re-assessment of assigned C-values for species with large discrepancies. As C-values are traditionally assigned based on the expertise of local botanists, working alone or in a small team, they are subject to mistakes caused by a range of factors. First, every botanist necessarily has incomplete knowledge of the range of habitats in which a species grows in any region, and these habitat preferences can change over time. Second, although rarity, beauty, and ecological uniqueness are explicitly to be disregarded in assigning C-values, botanists are human and cannot help but be swayed by these factors. For instance, *Iliamna remota*, the Kankakee mallow, discussed above and noted as a member of the “degradation resistant” type in Table 4, shows ecological characteristics and CNMC values of a species that should probably be assigned a C of 5, but, given its global rarity and the odd fact that it has never to anyone’s knowledge migrated off its single riverine island home, anyone would be strongly tempted to assign it a 9 or 10. Third, C-values of particular species can change over time. As an example, in the 4th edition of *Plants of the Chicago Region* (1994), *Napaea dioica*, glade mallow, was assigned a C-value of 10. In the *Flora of the Chicago Region* that followed, the same species was assigned a C-value of 4. Habitat managers had found that in the intervening years, the glade mallow had become more aggressive in the region and had greatly increased its ecological amplitude. Our method can be used to target species that show signs of having a mis-assigned C-value and promote further research into the distribution and population dynamics of such species.

More broadly, our analysis represents a first step toward a deeper understanding of how groups of species with the same C-values that behave differently from each other ecologically. These fall into several non-exclusive categories. Table 4, provides a first approximation of categories for species that deviate from the norm of species that have the same or similar C-values, and suggest these for further investigation. C-values are from *Flora of the Chicago Region (2017)*, except where noted.

TABLE 4 Categories of interest

Type	Description	Examples
Adaptive redistribution	Change to a given region or ecological system increases the viability of a species, bringing about increased ecological tolerance, expression, and/or geographic range of the taxon	*Mikania scandens* (C = 10) *Napaea dioica* (C = 10 (1994) → C = 4 (2017))
Co-occurrent genotypes	Co-existing morphologically identical lineages of both native and introduced genotypes render the provenance of any given individual(s) inconclusive and indeterminable	*Galium aparine* (C = 0) *Prunella vulgaris* (C = 0) *Achillea millefolium* (C = 0)
Cultivated pseudo-nativity	Widespread cultivation of a native species has obscured, overwritten, or made difficult to classify conservatism relative to its pre-settlement niche.	*Echinacea purpurea* (C = 10) *Pinus strobus* (C = 10) *Rubus odoratus* (C = 5 (1994) → Non-native (2017))
Degradation persistent	Species has a tendency to persist through degradation of its habitat, often appearing out-of-place in systems otherwise composed of much less conservative species. Its presence is not indicative of any ecological trend, despite bearing a high C-value.	*Asclepias lanuginosa* (C = 10) *Comandra umbellata* (C = 9) *Iliamna remota* (C = 10)
Non-causative niche	Species occupies an ecological niche that is not caused by conservatism trends, and is therefore unpredictably present in both low and high-conservatism situations.	*Orobanche uniflorua* (C = 5) *Liparis loesellii* (C = 4) *Lemna minor* (C = 5) *Cuscuta gronovii* (C = 5)
Schismatic preferences	Species is strongly divided between high-conservatism and low-conservatism habitats, with little to no in-between. Its presence is indicative of high conservatism in remnants, but must be assigned a much lower C-value due to expressed opportunism in low-conservatism habitats.	*Asclepias verticillata* (C = 1) *Geranium carolinianum* (C = 1) *Turritis glabra* (C = 2)
Foundation species	A taxon integral to development and expression of remnant ecosystems, and therefore a necessary presence in some remnants but unable to be granted high C-values. These species both do and do not predict remnants, and are therefore most often assigned C = 5 as a compromise.	*Andropogon gerardi* (C = 5) *Carex stricta* (C = 5) *Carya ovata* (C = 5) *Quercus alba* (C = 5) *Sorghastrum nutans*. (C = 5) *Typha latifolia* (C = 5)

To flesh out our thinking, it is instructive to compare *Orobanche uniflora* with other species with a C-value of 5. A value of 5 is often applied to species that are considered foundational to an ecosystem. The common oaks of the Midwest - *Quercus alba*, *Q. rubra*, and *Q. macrocarpa* - are the foundational species of the Oak-Hickory woodland, and along with the shagbark hickory (*Carya ovata*), all have C-values of 5. Big bluestem (*Andropogon gerardi*) and Indian grass (*Sorghastrum nutans*) are foundational species of the tallgrass prairie, and likewise have C-values of 5. The sedges, *Carex lacustris and C. stricta*, along with what's left of the broad-leaved cattail (*Typha latifolia*) are foundational species in various types of marshes, and all three of these have C-values of 5. A foundational species would generally not have a C-value too much lower than 5, as it would necessarily be found in all the remnant communities of its type. But it would also tend not to have a value too much higher than 5, as those communities could never have formed in the first place without some ability of the foundational species to survive and thrive in more degraded or ruderal ecosystems. It seems that 5 has become the default designation for a foundational species, but one that can be overridden - almost always to a lower value - when conditions necessitate it. One such example is switchgrass, *Panicum virgatum*, which has assigned C-value of 3. But all these foundational species are about as different ecologically from *Orobanche uniflora* as possible. The latter is a low density, low biomass, uncommon, and inconstant member of its ecosystem.

To put this in terms of the driver-passenger model of ecosystem organization (MacDougall & Turkington, 2005; Peterson, Allen, & Holling, 1998), the oaks and tall grasses just referenced are the archetypal ecosystem drivers, affecting the ecology of every member of their communities, and *Orobanche uniflora* is a paradigmatic passenger, leaving little effect on the members of its community as it passes in and out of view. This comparison reminds us that one C-value can encompass a wide range of ecological actors and some classes of these actors (such as those we've tentatively placed in the "non-causative niche" category) may be far more labile in their C-values, and thus more worthy of further investigation, than others.

Another avenue for further research comes from the ability to target particular habitats and pose larger ecological questions about them. The presence in particular habitats of species with large discrepancies in C-values can raise questions beyond the mis-assignment of individual C-values. It can direct our attention to changing ecological dynamics, which is especially important for conservationists in a time of species movement due to climate change. If several species that had been thought to have low C-values, but have higher observed CNMC-values, keep showing up in particular ecosystems, it can cause ecologists to look more carefully at those ecosystems to see if there are overlooked ecological changes or associations that explain them.

Co-occurrence research can be expected to help sharpen our conservation priorities. For example, *Turritis glabra*, tower mustard, is an uncommon plant in the Chicago region that has been assigned a C-value of 1. Our analysis of the published data (Figure 4) shows that it has a CNMC of 6.4 and adjusted C-value of 9.3. Conservation stewards who note that tower mustard is coming into an ecosystem may currently choose to disregard that ecosystem as a conservation priority due to falling native mean-C. But if a habitat contains tower mustard and a few other species that CMNC show have been underrated, conservation priorities may need to be reassessed and the ecology of the system more carefully studied.

Co-occurrence analysis can be applied to species which are not currently considered native to a particular ecoregion, some of which may yet have high conservancy and others of which might be problematic or harmful. By considering the sorts of plants that such species associate with, we may be able to systematically distinguish the one from the other. Given the inherently blurry nature of the native/non-native distinction, such an approach would be invaluable.

Finally, the system described in this paper leaves open the attractive possibility of spacial analysis. Considering plant communities at a more detailed level could offer insights beyond those made possible by co-occurrence analysis, though a number of practical and theoretical challenges will need to be overcome before such an approach can be put into practice.

References

Bauer, J. T., Koziol, L., & Bever, J. D. (2018). Ecology of floristic quality assessment: testing for correlations between coefficients of conservatism, species traits and mycorrhizal responsiveness. *AoB Plants*, *10*(1), plx073.

Bourdaghs, M., Johnston, C. A., & Regal, R. R. (2006). Properties and performance of the floristic quality index in great lakes coastal wetlands. *Wetlands*, *26*(3), 718–735.

Bried, J. T., Jog, S. K., & Matthews, J. W. (2013). Floristic quality assessment signals human disturbance over natural variability in a wetland system. *Ecological indicators*, *34*, 260–267.

Cohen, M. J., Carstenn, S., & Lane, C. R. (2004). Floristic quality indices for biotic assessment of depressional marsh condition in florida. *Ecological Applications*, *14*(3), 784–794.

Freyman, W. A., Masters, L. A., & Packard, S. (2016). The universal floristic quality assessment (fqa) calculator: an online tool for ecological assessment and monitoring. *Methods in Ecology and Evolution*, *7*(3), 380–383.

Gard, A., Myers, A., & Luwabelwa, I. (2024). The fqar package: R tools for analyzing floristic quality assessment data. *Journal of Open Source Software*, *9*(96), 6366. Retrieved from `https://doi.org/10.21105/joss.06366` doi: 10.21105/joss.06366

Gianopulos, K. (2014). Coefficient of conservatism database development for wetland plants occurring in the southeastern united states. *North Carolina Dept. of Environment & Natural Resources, division of water resources: wetlands branch. Report to the EPA, region*, *4*.

U.S. Forest Service

. (2024). *Midewin and its plants of concern.* Retrieved February 3, 2025, from `https://www.fs.usda.gov/wildflowers/Rare_Plants/conservation/success/midewin_plants_of_concern.shtml`

Kim, E. S., Zaya, D. N., Fant, J. B., & Ashley, M. V. (2015). Genetic factors accelerate demographic decline in rare asclepias species. *Conservation genetics*, *16*, 359–369.

Ladd, D., & Thomas, J. R. (2015). Ecological checklist of the missouri flora for floristic quality assessment. *Phytoneuron*, *12*(1), 1–274.

MacDougall, A. S., & Turkington, R. (2005). Are invasive species the drivers or passengers of change in degraded ecosystems? *Ecology*, *86*(1), 42–55.

Matthews, J. W., Spyreas, G., & Long, C. M. (2015). A null model test of floristic quality assessment: Are plant species' coefficients of conservatism valid? *Ecological Indicators*, *52*, 1–7.

Miller, S. J., & Wardrop, D. H. (2006). Adapting the floristic quality assessment index to indicate anthropogenic disturbance in central pennsylvania wetlands. *Ecological indicators*, *6*(2), 313–326.

Oldham, M., Bakowsky, W., & Sutherland, D. (1995). Floristic quality assessment system for southern ontario. *Ontario. Natural Heritage Information Center Report, Ministry of Natural Resources, Peterborough, Ontario, Canada.*

Peterson, G., Allen, C. R., & Holling, C. S. (1998). Ecological resilience, biodiversity, and scale. *Ecosystems*, *1*, 6–18.

R Core Team. (2022). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Spyreas, G. (2019). Floristic quality assessment: a critique, a defense, and a primer. *Ecosphere*, *10*(8), e02825.

Spyreas, G., Meiners, S. J., Matthews, J. W., & Molano-Flores, B. (2012). Successional trends in floristic quality. *Journal of Applied Ecology*, *49*(2), 339–348.

Taft, J. B., Wilhelm, G. S., Ladd, D. M., & Masters, L. A. (1997). *Floristic quality assessment for vegetation in illinois, a method for assessing vegetation integrity*. Illinois Native Plant Society Westville, Illinois.

Voss, E. G., & Reznicek, A. A. (2012). *Field manual of michigan flora*. University of Michigan Press.

Wilhelm, G. (1977). *Ecological assessment of open land areas in kane county, illinois* . Kane county urban development division.

Wilhelm, G., & Rericha, L. (2017). Flora of the chicago region: A floristic and ecological synthesis..

Wilhelm, G. S., & Masters, L. (1995). Floristic quality assessment in the chicago region and application computer programs. *Morton Arboretum, Lisle, IL*, 43212–1192.

Zinnen, J., Spyreas, G., Zaya, D. N., & Matthews, J. W. (2021). Niche ecology in floristic quality assessment: Are species with higher conservatism more specialized? *Ecological Indicators*, *121*, 107078.