AN EXAMINATION OF TEMPORAL TRENDS, REGIONAL VARIATION, AND HABITAT-TYPE DIFFERENCES IN SITE-LEVEL FLORISTIC QUALITY, AND THEIR IMPLICATIONS FOR THE USE OF FLORISTIC QUALITY ASSESSMENT

BY

GREG SPYREAS

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Natural Resources and Environmental Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Associate Professor Brenda Molano-Flores, Chair, Director of Research
Professor Jeffrey D. Brawn
Professor James W. Dalling
Associate Professor Jeffrey W. Matthews
Associate Professor John B. Taft

Abstract

Floristic Quality Assessment (FQA) is a measure of site conservation value. It is premised on using an area's plant species composition and diversity to estimate its human disturbance and degradation levels. FQA metrics are increasingly popular and influential for making land conservation, restoration, and policy decisions, as well as for scientific research. While it has been demonstrated that FQA metrics accurately measure site degradation/disturbance levels, many other FQA metric properties are unknown, especially compared to other ecological metrics. For this research, I assessed three important properties of FQA measures that are not understood: their regional patterns, their variation among different habitat-types, and their trends over time. I used site-level vegetation data from an Illinois statewide habitat monitoring program (Critical Trends Assessment Program, CTAP) to characterize FQA metric properties across regions and habitat-types. I found that forests had higher Floristic Quality values on average than wetlands. However, a separate analysis of a select group of the state's most pristine habitats showed that the upper-range of forest Floristic Quality was equal to, or lower than, that of other habitat types. Therefore, the difference between wetlands and forests observed statewide was due to the greater relative abundance of highly degraded wetlands across the state. Across the state, Floristic Quality decreased with latitude overall, although the variation explained was not great. This relationship was stronger for forests, than wetlands, which showed a weak, quadratic latitudinal relationship. Forests were the only habitat that varied in richness, exhibiting a weak decline to the north. Temporal Floristic Quality trends were compared using a 50-year, old-field succession study. Values in all fields followed the same asymptotic pattern, reaching a peak after around 35-years. The consistency of FQA values over time show that when sites of different ages are compared with one another, an asymptotic trend in metric values should be considered the default trend that is likely over early- to mid-successional timeframes. In summary, these studies found FQA value differences among regions and habitat-types. However, there was little evidence that these differences reflect an *inherent* property to FQA values such that they would confound general use of these metrics. Instead, variation in Floristic Quality across habitat-types and regions was either found to be very small, or it was probably a reflection of human

degradation levels across sites. Nonetheless, users must consider that some variation in FQA values could be attributable to the natural ecological characteristics of regions or habitat-types (i.e., not all variation in values was attributable to anthropogenic degradation/disturbance), and depending on users' goals, variation may need to be accounted for. Specifically, these results highlight that site comparisons across very large latitudinal gradients, or ones that compare certain specific habitat-types with one another (e.g., upland versus floodplain forest), or ones where only very high-quality, pristine habitats are being compared, are the instances where FQA score comparisons should be done the most cautiously.

Acknowledgements

I am professionally and personally thankful for many more people than I can name here. Nevertheless, I would like to thank my advisor, Brenda Molano-Flores, for her steadfast guidance and navigation through the PhD process. I would also like to thank my committee members, John Taft, Jim Dalling, Jeff Brawn, and Jeff Matthews for their comments on the chapters. With regards to the writing and analysis in these particular chapters, I thank T.J. Benson, David N. Zaya, Scott Chiavacci, Dennis Skultety, Mary Colleran, and Tara Beveroth. With regards to the data used in this research, I am grateful to CTAP botanists past and present, especially James Ellis, Connie Carroll-Cunningham, and Tim Rye, and thanks also to Loy R. Phillippe, John Ebinger, Allen Plocher, Paul Marcum, Jeffrey W. Matthews, Dan Busemeyer, Connie Carroll-Cunningham, and Michael Wallace for help sampling reference sites. Data used in chapter 3 was supported by NSF grant DEB-0424605 to Scott Meiners. I am grateful to the Illinois Department of Natural Resources for funding the Critical Trends Assessment Program (CTAP).

Table of Contents

Introduction: understanding ecological properties of site-level Floristic Quality scores

Conservation practitioners and ecologists are tasked with evaluating the naturalness, Biological Integrity, and conservation value of natural areas in a way that is as objective as possible. Vegetation-based measures are frequently employed in this capacity. But, the history of such measures' use has been one of only limited effectiveness and utility (Smith and Theberge 1986, Mace 2005). Among the ways that such measures are often lacking is that they must not only be straightforward in their meaning and application, but they must also be flexible. Ecological indicators and metrics would ideally allow natural areas to be compared in a standardized, repeatable, and easily interpreted fashion, across different habitats and over space and time. If they do not meet these criteria, they will not be widely used (Niemi and McDonald 2004).

It is in this context that Floristic Quality Assessment (FQA) has come to prominence as a measure of site conservation value and biological degradation; it is especially popular across North America (Appendix- Figure A.1). Floristic Quality is premised on a simple property— plant species composition may be used to rank sites by their human disturbance/degradation levels. This property is also an excellent measure of site conservation value, because undegraded habitats and their species have become rare. The rapid and widespread adoption of FQA has been attributable to its ease of use, as well as its measurement of site- or habitat-patch properties that are of keen conservation interest. From a performance perspective, studies have conclusively shown that site Floristic Quality measures are highly robust for their stated objective— namely, to use the composition of plant species at a site to accurately measure its levels of biological degradation and accrued human disturbances (Appendix- *Site FQA scores measure human disturbance and biological degradation*). However, one of the most important, but largely overlooked factors in FQA's widespread utility is that site Floristic Quality scores are readily compared with one another. Furthermore, their comparison assumes that values are standardized in their meaning. In other words, one site score of 20 means the same thing as another score of 20 in a given FQA state. This comparability of values and equivalence in meaning has allowed FQA to be used in a variety of policy and regulation contexts; for conservation study and research; to guide habitat assessments for their protection or acquisition; and in guiding habitat restoration strategies and management (Appendix- *How has FQA been used?*). Thus, FQA scores now yield a surprising amount of influence.

The assigned species weights (*C*-values) that constitute site Floristic Quality scores are relatively simple in their premise. Species *C*-values are based on the likelihood of each plant species' occurrence in remnant versus degraded habitats. Thus, FQA has no ecological theory underlying it. It relies on the expertise of the botanists who scored each state's flora— namely, their personal knowledge of regional plant species occurrences. Consequently, a curious aspect of FQA is how polarizing of a measure it is. Many conservation practitioners, applied ecologists, agency scientists, and even policy makers have enthusiastically adopted its use, because of the dearth of comparably flexible and meaningful measures of its kind. At the same time, many basic ecology researchers have dismissed it due to its perceived subjectivity or lack of grounding in ecological theory. While some of these criticisms are unfounded, (nearly any conservation valuation will incorporate some subjectivity), others may be warranted (Appendix). From the perspective of site-level Floristic Quality, the assumption that site values can be compared equivalently is premised on a basic understanding of their ecological properties (e.g., how do scores vary across space and time?). Similarly, the appropriate uses and comparability of other ecological metrics are guided by an understanding of their ecological properties (e.g., species richness comparisons across regions). In reality, however, a general understanding of site-level Floristic Quality value properties is largely absent. Therefore, the net effect of FQA's rapid adoption over the last several decades is that even though it backs a variety of substantial, real-world, conservation decisions, it lacks the critical assessment that is warranted by an ecological metric with such influence. This has no doubt been exasperated by the lack of serious consideration and study it has received by research ecologists.

Without research validating its assumptions and properties, erroneous and overly expansive uses of FQA have sometimes resulted (Appendix- *Species Conservatism*). Therefore, the general goals of this dissertation are twofold. The first is to assess the most critical, and largely untested, ecological properties underlying site-level Floristic Quality assumptions and use. The second is to highlight uses and assumptions that are unfounded in light of these results.

This dissertation's main chapters examine three critical properties of FQA scores by asking: how does site Floristic Quality vary among habitat-types, among regions, and over time? Previous study relevant to these topics is quite limited. This is often because information has come from studies that have not been explicitly designed to answer these particular questions, and their broaching of the topics has been a secondary research aim. Alternatively, when these topics have been the primary focus, the data available to address them has been severely limited in its inferential ability. In many ways, this dissertation represents both the first explicit, and the most comprehensive, tests addressing site Floristic Quality patterns and properties. This introduction briefly reviews any previous information or studies that are relevant to these three research topics— variation in Floristic Quality among habitat-types, regions, and over time. Following this introduction are the three study chapters. Finally, the last section, Appendix, provides a comprehensive review of the topic of Floristic Quality Assessment. It is the first of its kind. Its aim is to provide a comprehensive review of all literature on the subject, to synthesize Floristic Quality's essential concepts, and to highlight the area's most in need of study.

Comparing site-level Floristic Quality scores among habitat-types

Site-level FQA scores are said to be comparable regardless of habitat (e.g., grassland versus forests) or plant community-types (Wilhelm and Ladd 1988, Swink and Wilhelm 1994b, Andreas and Lichvar 1995, Nichols et al. 2006, McIndoe et al. 2008, Mortellaro et al. 2012). This means that the local ecological factors (soil, hydrology, aspect, etc.) and the particular plant communities that these factors support are assumed not to affect the underlying FQA value of a site. Therefore, barring differences in anthropogenic disturbance levels between sites of different habitat-types, their scores will be equivalent. This inter-habitat FQA assumption has allowed studies to compare large numbers of sites without needing to control for habitat type, as scores are assumed to be absolute reflections of the human disturbance level, conservation value, biological quality, management, or whatever else is being studied (e.g., Lopez and Fennessy 2002, Cohen et al. 2004, Balcombe et al. 2005, Jog et al. 2006, Mack 2006, Miller and Wardrop 2006, Spyreas and Matthews 2006, Johnston et al. 2008).

However, there is some evidence to suggest that communities may differ in their inherent Floristic Quality. Two studies found that one of their wetland community types had higher Mean *C* values relative to others (anecdotally compared in, Rentch and Anderson 2006, significant means comparisons in, Bried and Edinger 2009). (For descriptions of FQA's two metrics, Mean C and FQI see Appendix- *Terminology & Metrics*). At the same time, these higher Mean *C* communities (acidic sphagnum bogs, Rentch and Anderson 2006, vernal ponds with acidic nutrient-poor soils, Bried and Edinger 2009), have shown lower richness, leading to parity in community FQI values. These authors attributed higher Mean *C*-values to greater representation by species adapted to harsh abiotic environments, which are assumedly the more Conservative species in those communities (see also, fens scoring higher than riparian shrublands in, Rocchio 2007). However, neither study could rule out that human disturbance legacies accounted for observed differences. Andreas et al. (2004) even more forcefully attributed community type differences to greater numbers (or ratios) of Conservative species:

> "…plant communities with a naturally high proportion of habitat specialists will score higher than an equally intact plant community that is naturally dominated by more generalist species… [site FQA scores may not be] due to differences in floristic integrity but from differences in plant community types."

They single out some of the same higher scoring wetland community types identified in the studies above:

> "…because marshes contain relatively few habitat specialists; [and] many of the plants in marsh communities will grow in other types of wetland habitats, whereas fens and bogs have many habitat specialists…If the goal is to identify high quality marsh habitats for acquisition or protection, the FQAI [Mean *C*] score from a particular marsh should be compared to other marsh scores, not to a scores from a fen or bog."

They concluded with a strong warning that inter-community comparisons are improper uses of FQA that may be "meaningless for conservation purposes."

Unfortunately, species Conservatism in this study was equated with species specialization, which is problematic (Appendix- *Definitions of Conservatism*), and calls to question their results. Their statistical power was also limited by a small sample size. Nonetheless, more comprehensive studies have found somewhat similar results. Matthews et al. (2005) found FQA differences among broad habitat types from their sample of 231 sites. Forested and shrub wetlands scored higher than emergent wetlands and ponds. Unfortunately, they made no formal means tests and their site's human disturbance histories could not be controlled for, so the implications of their results are not clear. The most extensive study thus far on the subject used 1,000 historic vegetation surveys across Wisconsin (USA) and found significant differences in Mean *C* scores among several of 16 plant community types (Rooney and Rogers 2002). This study was particularly informative because the authors compared relatively undisturbed reference sites, allowing them to better isolate differences in Mean *C* attributable to the inherent characteristics of these communities and their floral assemblages. Despite its extensiveness, this study also had limitations, as certain regions of the state were comparatively over-sampled. Thus, spatial differences in the state could not be ruled out as having influenced community-type results. For example, northern communities tended to score higher than southern ones (bogs, northern lowland forest, boreal forest, northern upland forest, beech forest, southern upland forest, respectively), suggesting the potential for a latitudinal gradient (Johnston et al. 2009, see also “poor fens” in, Johnston et al. 2010). Furthermore, sample methods may have also varied among sites (Curtis 1959). Finally, even though sites were said to be selected as representing high-quality areas in the 1940-1950’s, site disturbance histories were not well documented and may have been uncontrolled. Thus, differences could not be exclusively attributed to natural/inherent community assemblages in these studies.

Some studies suggest that habitat-types do not inherently differ in Floristic Quality. Ervin et al. (2006) compared FQI values from different hydrogeomorphic (HGM) classes (i.e., wetland habitat types), and initially found differences. But, after statistically controlling for human disturbance and surrounding land use, FQA scores no longer differed. Miller et al. (2006) suggested that FQA was unique among the biological variables they tested in being neutral among HGM classes. These contradictory results highlight the need for further study across habitat types. Future study would need to control for human disturbance legacies. The goal is to make an inference about community types statewide and to avoid regional biases (e.g., over- or under-sampling certain regions). An ideal test would compare undisturbed remnants, from as wide a statewide coverage as possible, using a spatially unbiased site selection criteria (randomized, systematic, or stratified).

While Floristic Quality differences among habitat-types have thus far been discussed as reflecting a sites average Floristic Quality (i.e., its Mean *C* score), differences in FQI values could stem from inherent differences in habitat richness. For example, Francis et al. (2000) found that areas with different topography, aspect, moisture regime, vegetation structure, etc., within a forest, showed differences in their FQI scores because of their richness, but not their Mean *C*. Bernthal (2003) suggested caution when using FQI, as “some wetland types, such as temporary ponds, may have naturally low plant diversity.“ Similarly, wetland communities show different species-accumulation curves (beta-diversity), which could bias habitat specific FQI values depending on how they are sampled (Matthews 2003). Cohen et al. (2004) suggested Mean *C* as the more effective assessment tool when comparing among HGM classes because of potential richness differences among them.

In summary, a large part of the popularity and usefulness of FQA not only stems from its simplicity and ease of use, but it also comes from its flexibility. This flexibility includes the ability to answer questions using scores from sites spread over large areas and that includes different habitat-types (Allain et al. 2006, Medley and Scozzafava 2009). However, despite Swink and Wilhelm's (1994b) claim that FQA "assesses the aggregate conservatism of the plants inhabiting a site, *irrespective of community type(s)*", authors increasingly recognize that this assertion requires further empirical validation and study (Bried et al. 2012). A lack of inter-habitat comparability in site scores would hamper some FQA uses. Statewide Indices of Biotic Integrity (IBI) have dealt with similar concerns by relativizing scores by habitat-type (as well as by region) (Genet and Olsen 2006). This may be a future consideration for some FQA uses. Alternatively, if only a few communities are found to differ (e.g., bogs and fens above), only comparisons involving these habitats will warrant concern.

Comparing site-level FQA scores within states

As with the assumption of among habitat comparability, users frequently compare site-level scores with one another across large areas (e.g., across large states such as Illinois or the Dakotas (USA), Hargiss et al. 2007, Mack 2009, Spyreas et al. 2010). FQA scores on one border of a state are therefore assumed equivalent in meaning with those on the opposite border. Such comparisons rest on the assumption that users are isolating the effects from their study topic (e.g., a land-use disturbance, a type of management, effects from an invasive species), and that the underlying Floristic Quality of sites is otherwise neutral with respect to the inherent ecological factors that vary spatially across the state. Therefore, site scores should not show patterns related to biogeography, productivity, physiography, disturbance regime, climate, topography, edaphic conditions, glacial history, etc. under this assumption, and statewide patterns in scores should only reflect human disturbance legacies. If values are instead related to natural ecological variation across regions in a state, it could bias the use of FQA (see warning by Bernthal 2003). For conservation practitioners this could, for example, unintentionally disfavor land acquisition and protection in regions with a tendency towards inherently lower site Floristic Quality.

A few studies have compared site scores to examine regional FQA variation. Nichols (1999, 2008) analyzed the floras of 554 ponds and lakes across Wisconsin (USA), and found significant regional differences in FQI, Mean *C*, and species richness from northern to southern regions. This study attributed variation to decreasing development and human disturbance with increasing latitude, and not necessarily to effects from a natural spatial-ecological FQA gradient. Furthermore, because lake ecologies differ across the state (e.g., primarily oligotrophic, gravel bottom lakes in the north to eutrophic, mud bottom lakes in the south), these authors also could not rule out effects from comparing different community types. Cohen et al. (2004) did not find differences in wetland FQA scores across a latitudinal gradient in Florida (U.S.A). However, their site scores only included species designated as “wetland” plants, allowing for the possibility that calculating scores from complete flora lists may have shown different results. Miller et al. (2006) found that several *other* biological metrics they tested for monitoring wetland Biological Integrity were not useful statewide because of significant regional variation; however, because there was no such pattern for Mean *C,* they highlighted its promise for statewide comparisons. Reiss (2006) found latitudinal variation in Mean *C* values across Florida’s forested wetlands (U.S.A), but they warned that further inquiry was needed to rule out the possibility that land-use and human disturbance explained these regional patterns.

Given that other site-level ecological metrics, such as those based on diversity, vary along regional ecological gradients, further testing for spatio-ecological Floristic Quality patterns is needed. Controlling for underlying human disturbance gradients is clearly a considerable challenge to such inquiries, as ecological versus human effects often co-vary and can be difficult to discern (Findlay and Houlahan 1997, Nichols 1999, Johnston et al.

2008). If natural spatial variation in Floristic Quality is found, it would be useful to know whether such gradients are gradual across landscapes (e.g., those based on latitude) or if they are localized or abrupt (as might be expected with topography or edaphic features). This would help FQA users to know when spatial contexts are a concern and need to be factored into comparisons. Nichols (1999) suggested that in order "to be useful for a variety of monitoring applications it [FQA] might have to be calibrated for local conditions. Local-scale controls would be difficult to account for and would probably impede the ready use of FQA. Gradual landscape-wide gradients however, would be easier to account for with general correction factors (e.g., Johnston et al. 2010).

Understanding long-term successional trends in FQA and comparing site-scores over time

The temporal dynamics in conservation metrics are often insufficiently understood, leading to their misuse (Niemi and McDonald 2004, Fleishman et al. 2006). FQA measures are no exception, as incorrect accounting for temporal trends in FQA scores may compromise research. Scoring individual species *C*-values based on their successional tendencies is problematic (see Appendix- *Successional status*). However, the assumption that site age is related to Floristic Quality is reasonably made by users. Given that many rare, disturbance-sensitive species are restricted to the oldest, least-disturbed habitat patches this certainly makes sense (where age is defined as the time since a major anthropogenic disturbance) (Peterken and Game 1984, Honnay et al. 1998, Kindscher and Tieszen 1998, Honnay et al. 1999). Therefore, the highest site Floristic Quality levels are typically equated with "mature", "late", "advanced", or "climax" successional stages in many FQA treatments. Furthermore, many FQA users assume that absent human disturbance, site values increase in an orderly, predicable manner over time and in concert with succession. However, while it is reasonable to assume that FQA scores in remnant habitats "are stable over time, at least provided that the management does not change" (Smart et al. 2011), the assumption that FQA scores will *inevitably* and *consistently* increase over time if developing habitats are protected from disturbance, and as succession proceeds, is incorrect (Matthews et al. 2009b).

How then do FQA scores behave in developing habitats and over time? Post-disturbance changes in Floristic Quality scores seem to exhibit varied trends. Time since logging has been correlated with higher Floristic Quality, as has forest maturity (Francis et al. 2000, Wallace 2001). Most study on this topic comes from habitat restorations. And, Floristic Quality's temporal behavior in the early stages of wetland and grassland restorations (aka recreations) has not been consistent or predictable (Mushet et al. 2002, Balcombe et al. 2005, Spieles et al. 2006, McIndoe et al. 2008, Matthews et al. 2009b, Middleton et al. 2010). Restorations often "fail", as Floristic Quality declines because of exotic species invasions and/or improper site conditions (e.g., hydrology, Matthews and Spyreas 2010). Studies attempting to understand FQA in restorations over time are often complicated by confounding variables among them, such as supplemental seeding or planting, site landscape context, and management regimes. A chrono-sequence of abandoned uplands (not active restorations), was studied to isolate the FQA-time relationship from confounding site management variables, and these showed a linear increase in scores between 1 and 50 years in age with a peak of ≈2.75 (Mean *C*) in the 50-year old plot (Rothrock et al. 2011). By extrapolating this regression line, the authors estimated that remnant level Mean *C* values (Mean *C* > 3.5) could potentially be reached in 70-90 years post-disturbance.

Further data on this topic are needed to resolve the generality of the conflicting trends observed in these two studies. Specifically, how typical are asymptotic versus linearly increasing Floristic Quality trends? How typical are different ceilings and times to reach peak scores across sites? Furthermore, there is no understanding of how FQA scores are expected to behave beyond the early- to mid-successional time frames that they have been studied so far (e.g., Appendix- Figure A.4). In general, further studies on successional-temporal trends in FQA are needed that: 1) report ambient changes to scores across later successional stages (long-term Floristic Quality data), 2) assess multiple habitat types (i.e., other than wetland restorations), 3) report scores from different regions, and 4) compare scores from sites with different landscape settings and surrounding land uses, to assess the role of regional species pools, habitat buffers, etc.

To summarize, habitat maturity or successional status does not necessarily equate with Floristic Quality. Researchers using FQA scores to study particular treatments, must consider the potentially complex background trends that FQA scores undergo if their study compares sites of different ages (e.g., restorations), or if the study is conducted over more than one year. This is especially true in dynamic early successional habitats. For example, because Spyreas et al. (2010) did not account for successional age differences among their study treatment plots, the negative effects on Floristic Quality from an exotic species invasion that they reported were probably *underestimated* (see also, Tulbure et al. 2007).

Chapter 1: Does Floristic Quality Differ Among Habitat-Types?

Abstract

Comparing the conservation value of habitats across a landscape in a repeatable, quantitative, fashion has proven to be difficult— especially when the areas compared are of different vegetation- or habitat-types (e.g., a forest versus a grassland). Floristic Quality Assessment (FQA) is an increasingly popular way to compare the conservation value of areas in this way. FQA metrics register the effects of anthropogenic disturbance on an area's plant species composition and diversity to estimate its conservation value. Their popularity stems in part from their ease of use and flexibility, as site values are readily compared across wide regions and habitat-types. Some now question the assumption that FQA values are equivalent in their meaning across habitat-types, as other ecological metric values, such as species richness, differ across habitat-types for reasons that are unrelated to a site's disturbance level or conservation value (e.g., topography, productivity, soils). This study compared Floristic Quality values among habitat-types across Illinois (USA). It used vegetation data from 346 randomly selected and 53 specifically chosen high-quality reference sites, to assess a range of conditions. Randomly selected wetlands and forests differed in their FQA values on average, due to the relative abundance of highly degraded wetlands across the state. Among reference quality habitat-types, floodplain forests had lower average Floristic Quality than sedge meadows, marshes, and prairies. Reference marshes had lower FQA values than sedge meadows and prairies when compared with the richness weighted Floristic Quality Index (FQI). Between forest types, floodplain forests had lower average Floristic Quality values than upland forest (both reference and randomly selected sites). The differences demonstrated here may present limitations to how FQA values should be compared among habitat-types. Specifically, where value differences are attributable to a habitat's inherent ecological or biophysical characteristics (e.g., hydrology, soils), habitat categorizations may need to be accounted for (because most FQA use assumes that anthropogenic degradation solely dictates values). However, because only

certain specific habitat-types and metrics were shown to inherently differ (i.e., reference marsh and floodplain forest), such instances are limited.

Keywords

Floristic quality, Conservation value, Species conservatism, FQA, Wetland, Forest, Prairie, Grassland, Habitat-type, Vegetation-type, Plant community, Anthropogenic disturbance, Habitat degradation, Floral assemblage

Introduction

At large spatial scales, the different vegetation-types that characterize ecosystems or biomes have been compared for their relative conservation value (Sampson and Knopf 1994, Hoekstra et al. 2005). But, at smaller scales, other than a few specific habitat-types having been identified as being rare (Izco 1998), or unique (e.g., serpentine communities in, Kruckeberg 1984), comparative assessments of areas across a given landscape that are of different habitat-types have not been conducted. This is largely because the means to quantify the conservation value of different communities across landscapes are poorly developed (Smith and Theberge 1986, Izco 1998, Chapman et al. 2009). This is despite the exhaustive history of defining and delineating plant community-types at this scale (Whittaker 1978, Ricklefs 2008, Grondin et al. 2014). Comparative assessments of sites of different habitat-types are further complicated because the differences in plant species composition and diversity that determine their conservation value may arise from either natural or anthropogenic processes (van Breugel et al. 2013, Garcıa-Madrid et al. 2014).

Over the past 30-years, Floristic Quality Assessment (FQA) has become an increasingly popular and influential way to compare the conservation value of areas based on their plant assemblages (Mack 2007, Matthews et al. 2009b; Appendix). A site's Floristic Quality level is estimated by two metrics that are calculated from its plant species list. The first metric is the Mean Coefficient of Conservatism (Mean *C*) of its species, where species Coefficients of Conservatism (*C*-values) are *a priori* assigned values that range from 0-10 (*C*-values are described further in the Methods). The other metric incorporates both site species richness and site Mean *C* to create the Floristic Quality Index (FQI). Sites with higher FQA values (FQI or Mean *C*) are of greater conservation value because they are less biologically degraded from their pristine, historical, condition having incurred less anthropogenic disturbance and retaining a fuller complement of species that are sensitive to anthropogenic disturbance (i.e., Conservative species with *C*-values between 7-10). Once calculated, any site's FQA value can then be compared with any other site's value to determine its relative degradation level or conservation value within a given FQA region. FQA regions are the scales at which species *C*-values are assigned— most often a state or province. Site FQA values can only be compared within their FQA regions. Furthermore, sites of any community- or habitat-type are assumed to be comparable within an FQA region (e.g., grassland versus forest) (Wilhelm and Ladd 1988, Swink and Wilhelm 1994b, Andreas and Lichvar 1995, Nichols et al. 2006, McIndoe et al. 2008, Mortellaro et al. 2012, Bried et al. 2013).

Indeed a large part of the popularity of FQA metrics stems from their ease of use, and from their flexibility due to these liberal assumptions about value comparability (i.e., an FQA value means the same thing regardless of its habitat-type) (Allain et al. 2006, Medley and Scozzafava 2009). The ability to compare diverse sites across broad regions using a single metric has motivated FQA's use in scientific inquiries. For example, studies have determined the effects of ecological management, exotic species invasion, and landscape context on native Floristic Quality and conservation values across sites of varied habitat-types (e.g., Lopez and Fennessy 2002, Cohen et al. 2004, Balcombe et al. 2005, Matthews et al. 2005, Bowles and Jones 2006, Jog et al. 2006, Mack 2006, Miller and Wardrop 2006, Miller et al. 2006, Spyreas and Matthews 2006, Mack 2007, Johnston et al. 2008, Mack et al. 2008, Mack 2009, Spyreas et al. 2010). FQA values are perhaps most influential due to their *direct* use in making habitat conservation decisions, as in evaluating natural areas for purchase or preservation (e.g., The Nature Conservancy) (Rocchio 2007). Similarly, some federal, state, county, and municipal agencies legally mandate using exact

FQA values to evaluate wetlands within their jurisdictions. A common example is to use an FQI value of 20 as a threshold for triggering regulatory action by state or federal agencies for wetland habitats protected under section 404 of the Clean Water Act (Herman et al. 1997, Streever 1999, Herman et al. 2001, Matthews and Endress 2008, Matthews et al. 2009b).

Whether they are for regulatory, research, or site conservation value assessment, such uses of FQA metrics implicitly assume that the inherent ecological, chemical, or physical factors (soil, hydrology, aspect, disturbance regime, etc.) that *naturally* characterize sites of different habitat-types, do not dictate their FQA values. Instead, FQA values are seen as reflections of site conservation values that are driven by their relative anthropogenic disturbance levels. However, other ecological metrics are not independent of habitat-types. Species richness for example (i.e., site-level alpha diversity, or species density), varies among different habitat-types for reasons that have nothing to do with how disturbed a site is (e.g., productivity, topography, soil heterogeneity, vegetation structure). Some authors have therefore questioned whether FQA metric values are comparable across habitat-types; either because Mean *C* may inherently differ among habitat types (Andreas et al. 2004), or the richness component to FQI could (Bernthal 2003, Matthews 2003, Bowles and Jones 2006).

Studies have not yet determined if FQA values differ across a given region's habitats. A few studies have suggested possible differences in Mean *C* values between relatively unique or uncommon habitat-types, such as bogs or fens (Andreas et al. 2004, Rentch and Anderson 2006, Rocchio 2007, Bried and Edinger 2009). A far more extensive study summarized historic vegetation data across Wisconsin (USA) and found significant differences in Mean *C* values among several habitat-types (Rooney and Rogers 2002). However, because certain regions of the state were comparatively over-sampled in this dataset (Curtis 1959), the variation among habitat-types could not be separated from regional variation. For example, differences could have reflected the latitudinal gradient in anthropogenic disturbance and land use intensity that exists across Wisconsin (Nichols 1999) (see also Chapter 2). Other studies have focused on comparing *FQI* values across habitat-types, and their conclusions have been similarly limited (e.g., Johnston et al. 2008). Furthermore, studies that only report FQI values are not able to discern whether habitat differences are attributable to FQI's richness or Mean *C* component. Finally, several studies have found habitat-type classifications to be unimportant when explaining variation in FQA values, although these have only compared wetland types (Ervin et al. 2006, Miller et al. 2006, Johnston et al. 2009, Johnston et al. 2010).

To summarize, among-habitat FQA value results have been ambiguous. The search for patterns has been hampered by dataset limitations such as small sample sizes, regional sampling bias, few habitat-types being compared, and/or a tendency to only report values for one FQA metric. Further tests are needed to determine if there are large inter-habitat FQA value differences that may need to be accounted for when comparing values among habitat types. To such an end, statewide data with an unbiased site selection criteria (e.g., probabilistic systematic, stratified) are needed to avoid under/over representing regions in a state. Characterizing sites across an entire state is especially crucial because each species *C*-value is scored based on its statewide tendencies (see Methods). Thus, sampling must represent species' entire ranges across states to accurately represent FQA values.

This study seeks to assess overall differences in Floristic Quality among Illinois habitats in two ways. First, it considers randomly selected sites in order to reveal differences in FQA values that are representative of habitats across the state. Second, it considers less-disturbed sites of high Floristic Quality to reveal the upper limit to values for habitat-types. These two datasets are used because differences in FQA values could be attributable to a habitat's history of anthropogenic degradation (assessed by randomly selected sites), or to the habitat's inherent ecological or biophysical characteristics (assessed by high-quality sites). I ask the following specific questions regarding the FQA values of Illinois' major habitat-types:

1. Do site FQA values from randomly selected wetlands differ from those of forests across Illinois? (prairies excluded due to their rarity)

2. Do select, least-disturbed, high quality, *reference* site FQA values differ among Illinois' prairies, wetlands, or forests?

3. If site-level FQI values differ, are differences attributable to lower Mean *C* or lower richness?

Methods

Illinois (USA) is a state with a considerable diversity of native habitats, which makes it a useful test ground for FQA. This variety is in part due to its size, but also due its geographic situation. It lies within North America's eastern forest-prairie transition zone that incorporates floral elements from eastern forests and western prairies, while its considerable latitudinal gradient encompasses boreal and southern coastal plain taxa (Braun 1950, Schwegman et al. 1973). At the same time, Illinois habitat-types have been differentially destroyed or degraded. For example, native grassland (prairie) has largely been eliminated, but considerable native forest remains (Illinois Department of Natural Resources 1994).

Two sets of data were used in this study. The first was randomly selected emergent wetlands and forests throughout Illinois that were sampled as part of Illinois' Critical Trends Assessment Program (CTAP). Forests were visited from south to north across Illinois, from 15 May through 30 June, and wetlands were sampled between 1 July and 31 July. Forest sampling was confined to homogeneous areas of forest with respect to aspect, hydrology, topography, and forest type, when possible. Sampling was generally done at least 50 meters from the forest edge. Sample species lists in forests were created by recording all vascular plants observed in twenty 0.25-m^2 quadrats (0.5-m X 0.5-m) spread across two randomly selected 50-m transects at each site. These ground layer quadrats were nested within larger tree (10-m X 50-m) and shrub (2-m X 50-m) sample plots, from which additional woody plants were recorded. In addition, species within a plot (10-m X 50-m plot) on a third randomly selected transect were also recorded. In wetlands, vascular plants were recorded from a 40m X 50m plot placed on the edge of the wetland and oriented to span its hydrological gradient from the upland inward. A total of 157 forests and 189 emergent wetlands from the years 1997–2012 were sampled, with the vast majority sampled from 2007-2012. See Carroll et al. (2002) for detailed site selection and sample protocols.

The second dataset was comprised of minimally disturbed wetland, prairie, and forest reference sites. The goal was to sample the highest-quality representative sites from the dominant Illinois habitat-types. The following habitats were selected: Dry-mesic/mesic upland forest, wet-mesic/mesic floodplain (bottomland) forest, sedge meadow, marsh, dry-mesic/mesic prairie (Table 1.1). Reference quality forests were widespread enough that their samples could span the state, but this was not possible for high-quality prairie and wetlands, which were clustered in the northern third of the state (Figure 1.1). The extent to which reference sites were undisturbed and undegraded is based on their characterization for the Illinois Natural Areas Inventory conducted in the 1970's (White 1978). While the aim was to sample the highest existing FQA values for these habitat-types, it was unavoidable that some reference sites have incurred some degradation from their historic condition. For the purposes of this study, however, the important assumption was that there was no difference in the relative disturbance levels among habitat-types. A total of 53 references site were sampled from 2000-2009. Other than their site selection criteria, random and references sites were similarly sampled.

Site FQA values are based on Coefficients of Conservatism (*C*-values) ranging from 0 to 10 that have been assigned to each native plant species (and some sub-species) in Illinois. Highly Conservative plants (7-10) are only found in minimally degraded natural areas, whereas species that persist in, or readily invade degraded areas are given lower numbers. Non-natives receive $C = 0$. A species C-value is determined by considering its simultaneous occurrences and behavior everywhere across a state. Some taxa may show different Conservatism behavior in different regions of a state, but C-values reflect a species' average statewide behavior. Species C-values were obtained from Taft et al. (1997), with minimal modifications (e.g., *Phragmites australis* was considered non-native, Spyreas, Ellis et al. 2004). Some taxa were counted as native when sample sites were within their native range and exotic when the sample site was elsewhere in the state (e.g., northern versus southern sites for *Taxodium distichium, Pinus echinata, Robinia pseudoacacia, Catalpa speciosa, Viburnum recongnitum*). In some cases, where a specimen could not be identified to species, the lowest C-value in its genus that it resembled was used as a conservative estimate (e.g., *Stachys* sp. $C = 5$). Or, in a few cases, where two species could not be distinguished and their values were close one another their average value was used (e.g., *Impatiens capensis* $C = 2$ and *I. pallida* C = 4; *Ulmus rubra* = 3 and *U. Americana* = 5). FQA metrics were calculated as follows. The Mean Coefficient of Conservatism ($\overline{C}$ or Mean C) for all species detected was calculated as ($\overline{C} = \sum C/S$), and the Floristic Quality Index was calculated as (FQI=$\overline{C}$ $*(\sqrt{S})$), where S is total species richness, and C is each species' Coefficient of Conservatism value. We included non-native species for both S and C in calculations (see, Spyreas et al. 2012).

Analysis

Because sample area (plot-size) differed for forests versus wetlands/prairies, their richness and FQI values cannot be compared due to the effect of sample area on richness. However, Mean C value comparisons are robust to moderate differences in site sample intensity (Appendix), and they were compared across all habitat-types. One comparison did not meet ANOVA assumptions of normality or equality of variance (Shapiro-Wilk, $p < 0.05$; random wetlands versus forests). However, because such tests become prohibitively sensitive with large sample sizes (i.e., N= 345), and because ANOVA assumptions can be modestly violated and still produce robust results (Gotelli and Ellison 2012), and because examination of the distribution of the test residuals showed a good model fit (Figure 1.9), parametric tests were assumed robust. Nonetheless, a Kruskal-Wallis One Way ANOVA on ranks was also conducted, and since it also resulted in a highly significant difference between group means ($p < 0.001$), differences were interpreted confidently. All other comparisons met test assumptions. Significant *post-hoc* pairwise comparisons were tested using the Holm-Sidak method with an alpha level of $p < 0.05$. ANOVA tests were conducted using SigmaPlot for Windows12.0, (Systat Software, Inc., Chicago, IL).

Results

Site Mean *C* values were significantly different between randomly selected forests and wetlands across Illinois (Figure 1.2; ANOVA: $F_{1,345} = 107.4$. $p < 0.001$). Mean *C* comparisons among reference habitat types showed significant differences between floodplain forest and both marsh and sedge meadow (Figure 1.3; ANOVA: $F_{4,50} = 4.48$, $p < 0.004$). Reference marsh had significantly lower FQI and richness values than sedge meadow and prairie (Figure 1.4; ANOVA: FQI- $F_{2,27} = 5.68$, $p < 0.009$; richness- $F_{2,27} = 10.3$, $p < 0.001$). Site FQA values and richness were lower in floodplains in randomly selected forests (Figure 1.5; ANOVA: Mean *C*- $F_{1,152} = 8.63$, $P < 0.004$; FQI- $F_{1,152} = 19.8$, $p < 0.001$; richness- $F_{1,152} = 18.1$, $p < 0.001$), while FQI and richness were lower in uplands in reference forests (Figure 1.5; ANOVA: Mean *C*- $F_{1,23} = 1.99$, $p < 0.171$; FQI- $F_{1,23} = 5.88$, $p < 0.02$; richness- $F_{1,23} = 34.7$, $p < 0.001$).

Site FQA values have been shown to vary by latitude across Illinois (Chapter 2). And, while randomly selected wetland and forest site comparisons are not affected by latitude because of their spatially balanced site selection, other categorical comparisons could be (i.e., reference sites, upland versus floodplain forest for random sites). In order to ensure that means differences were not attributable to latitude, ANCOVA using latitude as the covariate was conducted (Proc GLM, SAS 9.2, SAS Institute Inc., Cary, NC). Comparisons of Mean *C* across all reference habitat-types (*as per*, Figure 1.3), showed that habitat category differences remained after accounting for latitude. (ANCOVA: Mean *C*- $r^2 = 0.53$; latitude $p < 0.0003$, habitat $p < 0.025$, lat*hab $p < 0.011$). Comparisons of FQI and richness between wetland and prairie reference habitats (*as per* Figure 1.4) showed that their habitat-type differences remained after accounting for latitude (ANCOVA: FQI- $r^2 = 0.70$; latitude $p < 0.0001$, habitat $p < 0.0001$, lat* hab $p < 0.59$; richness- $r^2 = 0.56$; latitude $p < 0.018$, habitat $p < 0.0003$, lat*hab $p < 0.72$). Analysis comparing randomly selected forests (*as per* Figure 1.5) showed that habitat differences remained after accounting for latitude (ANCOVA: Mean *C*- $r^2 = 0.42$; latitude $p < 0.0001$, upl.flood $p < 0.0001$, lat*upl.flood $p < 0.14$; FQI- $r^2 = 0.44$; latitude $p < 0.0001$, upl.flood $p < 0.0001$, lat*upl.flood $p < 0.74$; richness- $r^2 = 0.24$; latitude $p < 0.0001$, upl.flood $p < 0.0001$, lat*upl.flood $p < 0.18$), while comparisons of reference forests showed that only richness significantly differed between habitat-types after accounting for latitude (ANCOVA: Mean *C*- $r^2 = 0.13$, latitude $p < 0.74$, upl.flood $p < 0.09$, lat*upl.flood $p < 0.88$; FQI- $r^2 = 0.24$; latitude $p < 0.08$, upl.flood $p < 0.10$, lat*upl.flood $p < 0.50$; richness- $r^2 = 0.64$; latitude $p < 0.02$, upl.flood $p < 0.0001$, lat*upl.flood $p < 0.15$).

Discussion

This study's main findings were that randomly selected forests had higher Mean *C* values than randomly selected wetlands overall (Figure 1.2). At the same time, reference forests (especially floodplain) had lower Mean *C* values than wetlands and prairies (Figure 1.3). Reference marshes had lower FQI values than reference prairie and sedge meadows, due to their lower richness (Figure 1.4). Across forests, upland forests consistently had higher Mean *C* values than floodplain forests (Figure 1.5; especially after latitude was accounted for). Upland forest FQI and richness was also higher than floodplain forest in random forests, but in reference forest uplands FQI values were lower along with, and due to, lower richness.

The differences between average forest and wetland Mean *C* values found in this study are not attributable to these habitats' inherent ecological characteristics because *reference* forest Mean *C* values were not higher than those of wetlands (Figure 1.3). In other words, forests do not naturally support higher Floristic Quality than wetlands (i.e., based on values in their pristine, high-quality state). Indeed, wetlands may even be capable of achieving higher Mean *C* values than forests (Figure 1.6 & 1.10). Instead, it was a lack of highly disturbed, low Mean *C* forests, relative to wetlands that drove statewide averages (Figures 1.6 – 1.8). In their survey of 231 habitat patches in southern Illinois, Matthews et al. (2005) found similar results. Their sampled forests had higher mean FQA values than wetlands, presumably because of the wetlands' greater disturbance— being measurably smaller and more isolated than forests. Although, Rooney and Rogers (2002) did not make comparisons between general wetland and forest categories, their results from specific community-types across Wisconsin did not suggest consistent differences between them. One point that is worth noting for all of these studies is that they have not included very young forests, or forests with recent, massive disturbances (e.g., clear-cut). This is assumedly because at that point such habitats are typically not considered forests. Thus, old-fields, successional habitats, or young, sapling-dominated stands (i.e., without a developed tree canopy) have not been represented in among habitat-type comparisons. This is compared to wetlands, which can be very young and disturbed, while still readily being recognizable as wetlands. This could result in a skewed sampling of disturbed forests versus disturbed wetlands overall, and depending on whether Floristic Quality accrues at that same rate in forests during succession as wetlands (Spyreas et al. 2012), this could lead to a bias in FQA by habitat comparisons. Future studies of this type should consider how to integrate early-successional habitats (i.e., nominally young forests). Nonetheless, results from the present study may be seen as providing the most extensive and accurate depiction of ambient Floristic Quality levels of a state's wetlands and forests to date.

For the second research question, FQA value differences among least-disturbed, high-quality habitat types were compared. For Mean *C*, high wetland and prairie values (Figure 1.3), suggest that they may be able to attain higher Floristic Quality levels than forests. Perhaps because their ecological and biophysical characteristics favor species with higher *C*-values (discussed below). However, the present study cannot completely rule out historic anthropogenic disturbance legacies as determining differences among reference forest and wetland/prairie habitats. For example, not all of the forest reference sites were "primary", "old-growth" forests (sensu, Spyreas and Matthews 2006), as many of them were historically logged (Wallace 2001). This is compared to wetland and prairie reference sites, which were primary (e.g., without historical tillage). Therefore, the reference forests sampled could have been inadequate representatives of the highest quality forest possible, and their lesser Mean *C* could be a spurious result (Figures 1.6 & 1.7). In the only other study focusing on higher quality habitats, Rooney and Rogers (2002) found Mean *C* differences among specific habitat-types, but their results were not comparable among the broad habitat categories of wetland, grassland, or forests. Additionally, their habitat-type results could be separated from unequal sampling in certain regions, which means that their results could reflect regional differences in human disturbance intensity and land use across the state. In contrast, regional effects on habitat differences in the present study were either controlled for by a spatially balanced statewide sampling design (random), or by using ANCOVA on reference sites to account for region.

Regarding this study's third research question, site FQA value differences came from the average Floristic Quality of habitats (Mean *C*), species richness, and/or the Floristic Quality Index (FQI), depending on the habitat-type. Previous studies have speculated that the floras of marshes specifically might support lower *C*-value species on average, causing their Mean *C* values to be relatively deficient compared to other habitats (Andreas et al. 2004, Rocchio 2007). Others have speculated that marshes could have lower richness, and therefore, lower FQI values on average (e.g., Bernthal 2003). The present study found that marshes had lower FQI and richness levels, but equivalent Mean *C* values compared to sedge meadows and prairies (Figure 1.4). Alternatively, it was another habitat-type, floodplain forests, which had consistently lower Mean *C* values compared to other habitat-types (Figure 1.3), including upland forests (Figure 1.5). Thus, potential differences in FQA values between habitats may come from either their richness or Mean *C* components.

These results highlight that there may be different reasons for Floristic Quality differences among habitats. They may be related to human disturbance levels (i.e., wetland versus forest overall), or they may arise from inherent habitat properties (i.e., floodplain forest Mean *C*). The mechanism(s) behind differences that may be inherent to a habitat's ecological characteristics have not been studied. Harsh abiotic conditions are previously suggested as predisposing floras in some habitats to having higher Mean *C* values (Andreas et al. 2004, Rentch and Anderson 2006, Bried et al. 2013). This is certainly plausible as more physiologically stressful environments may predispose assemblages to a greater proportion of "stress tolerant plants" (Grime 1979, Keddy and MacLellan 1990, Wisheu and Keddy 1992), which as a group tend to have higher species *C*-values (Taft et al. 1997). With respect to results from the present study, one might then expect the hydrology of marsh habitats to cause greater water inundation and hypoxia stress to their plants and to produce higher Mean *C* values. Similarly, greater hydrologic stress in floodplain versus upland forest might be expected to elevate their average Mean *C* value. This was not observed in either case. The frequency, intensity, or type of disturbances that naturally characterize specific habitat types (e.g., fire, wind-throw, small mammals, soil disturbance, grazing, browsing, flooding) could also shift species compositions towards naturally higher proportions of weedy, low *C*-value species (Collins and Pickett 1987, Gilliam et al. 1995, McLachlan and Bazely 2001, Meiners et al. 2002)— resulting in lower site Mean *C* values. This could be an explanation for floodplain forest results, with their regular flooding acting as a disturbance, rather than an abiotic physiological stress. Finally, another possible mechanism could be site productivity. Greater productivity might allow for greater assemblage representation by matrix "competitor" species (Keddy and MacLellan 1990, Wisheu and Keddy 1992), which tend to fall within the moderately-high range of *C*-values (Taft et al. 1997). Although it is not obvious which habitat type's productivity would be expected to generate the highest Floristic Quality values for in Illinois, as productivity may be measured in different ways.

While Mean *C* differences among habitat-types reflect the *ratio* of conservative to non-conservative species at a site, differences in *FQI* values can result from inherent differences in species richness, which can be attributable to an even wider possible variety of ecological characteristics of habitats. For example, Francis et al. (2000) found that areas within a forest had different FQI, but not Mean *C* values, implicating species richness as determinant. These areas differed in their topography, aspect, moisture regimes, and vegetation structure, meaning that any one of these variable(s) could have been responsible for species richness driven FQI differences. Other studies have suggested that because FQI and diversity differ among wetland types, Mean *C* is the preferred metric for comparing FQA values among habitat-types (Matthews 2003, Cohen et al. 2004). Finally, non-native species invasion can affect both FQA values and species richness (Spyreas et al. 2010). And, because habitat-types vary in their level of non-native species abundance (Spyreas et al. 2004), either due to their relative susceptibility to invasion (Martin et al. 2009) or their disturbance levels, this could also contribute to variation in Floristic Quality observed among habitat-types.

Summary, Implications, and Future study

The ease with which their values are compared across different habitat-types is one of several reasons FQA metrics are unique among existing measures of site conservation value (Medley and Scozzafava 2009, Bried et al. 2013). This study's results show that site-level FQA values vary among certain habitat-types. In certain instances, some of these differences could confound how FQA values are used and they may need to be accounted for. Whether and where FQA values are comparable across habitat-types will depend on the metric used (FQI or Mean *C*), and the habitat-types being compared (e.g., floodplain forest). Users will also need to be conscious of how fine of habitat category distinctions are being made. Anecdotal recommendations are that comparisons should be made among "similar plant community types or similar ecological system types" or "sites within the same ecosystem classification" (Rocchio 2007, Medley and Scozzafava 2009, respectively). Whereas the current study raises the somewhat counterintuitive prospect that finer community-type distinctions are more likely to inherently vary in values (e.g., floodplain versus upland forest), compared to broader habitat categories from obviously different vegetation or physiognomic classes (i.e., forest versus grassland). The question of how broad or fine of habitat categorizations are important for FQA will require further clarification.

Several other factors are also important when comparing habitats. Users should be aware when sites are being compared across broad regions in a state (especially across latitudes), as region can account for more variation in site values than habitat-type (Chapter 2). Users should also consider whether low- or high-quality sites are being compared, as site disturbances will account for far more variation in FQA values than other factors (Appendix). Thus, comparisons of high-quality sites would be more sensitive to confounding factors that are unrelated to disturbance. Finally, whether the habitats being compared are common, rare, or unusual may be important, as the Floristic Quality of rare or unusual habitats are less well characterized.

Further studies need to determine whether results from the present study are mirrored in other states. Studies should also investigate the causes of variation. For example, it has been suggested that habitat based differences could reflect bias in how species the *C*-values of certain habitat's floras have been assigned (Bowles and Jones 2006). Alternatively, they could occur because certain habitat-types have tended to incur greater levels of anthropogenic disturbance historically. Or, they could occur because some habitat type have certain inherent ecological properties that lead them to harbor higher *C*-value species, or that allow them to recover more quickly from anthropogenic disturbances. The extent to which future research can discern the ecological and biophysical, (e.g., natural disturbance regimes, resource availability, abiotic conditions), human disturbance related, or species *C*-value based characteristics that are ultimately responsible for Floristic

Quality differences among habitat types, the more FQA users will be able to adapt metrics to account for them, and the more accurate they will be at characterizing site conservation values.

Acknowledgements

I am grateful to CTAP botanists past and present as well as T.J. Benson, David N. Zaya, and Tara Beveroth. Thanks also to Loy R. Phillippe, John Ebinger, Allen Plocher, Paul Marcum, Jeffrey W. Matthews, Dan Busemeyer, Connie Carroll-Cunningham, and Michael Wallace for help sampling reference sites.

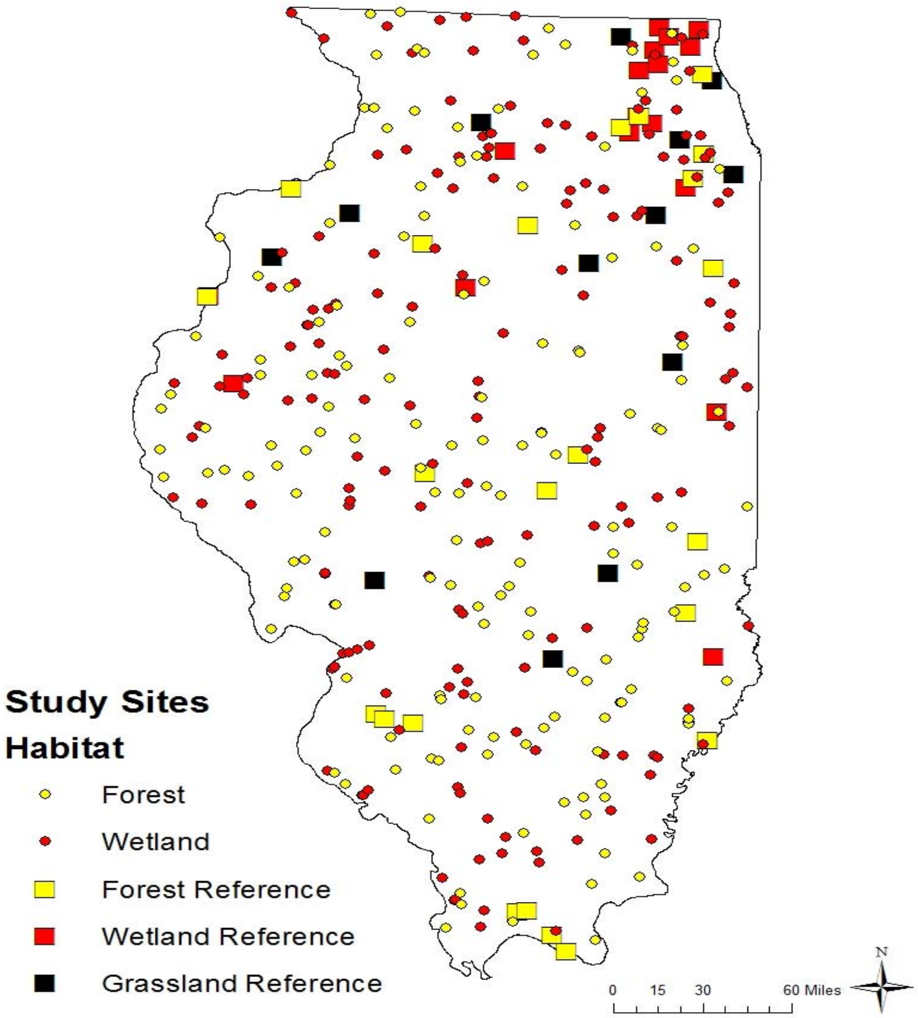


Figure 1.1. Sample site locations. Sites are color coded by habitat-type and symbol size to distinguish reference (squares) versus randomly selected sites (circles). The site selection procedure for random forests and wetlands was spatially balanced across the state, whereas only forest reference sites achieved statewide coverage.

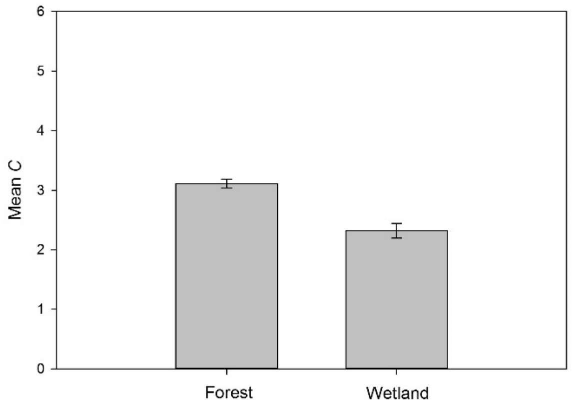


Figure 1.2. Site Mean *C* comparisons between randomly selected forests and wetlands across Illinois (95% confidence interval bars shown). A significant difference was found (ANOVA: $p < 0.001$).

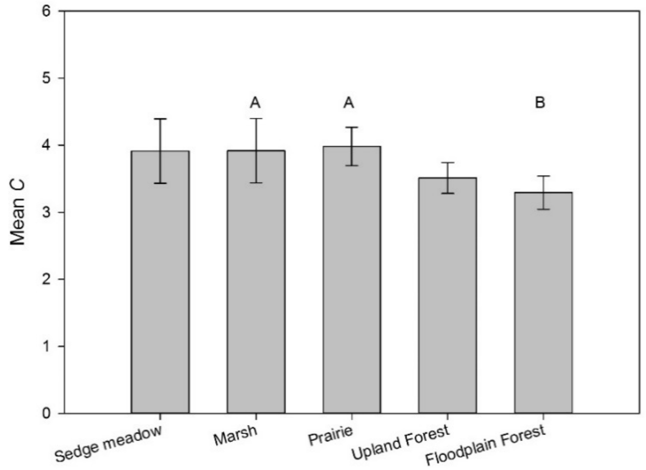


Figure 1.3. Reference site Mean *C* comparisons among Illinois habitat types (95% confidence interval bars shown). A significant overall difference was found (ANOVA: $p < 0.001$), with significant post-hoc pairwise comparisons indicated by different letters (Holm-Sidak method: $p > 0.05$). Habitats without letters are not significantly different from any other habitat.

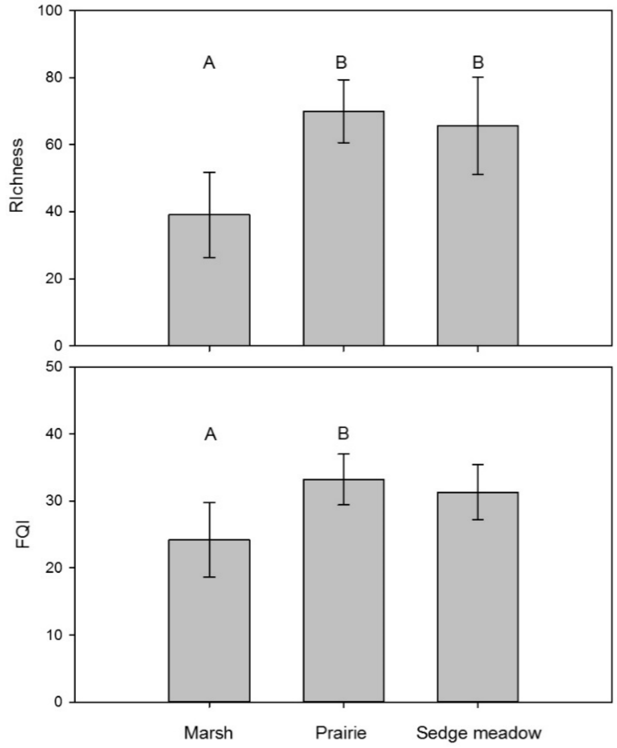


Figure 1.4. Reference site FQI and richness comparisons among Illinois wetland and prairie habitats (95% confidence interval bars shown). Significant overall differences were found (ANOVA: $p < 0.001$), with significant post-hoc pairwise comparisons indicated by different letters (Holm-Sidak method: $p > 0.05$). Habitats without a letters are not significantly different from any other habitat.

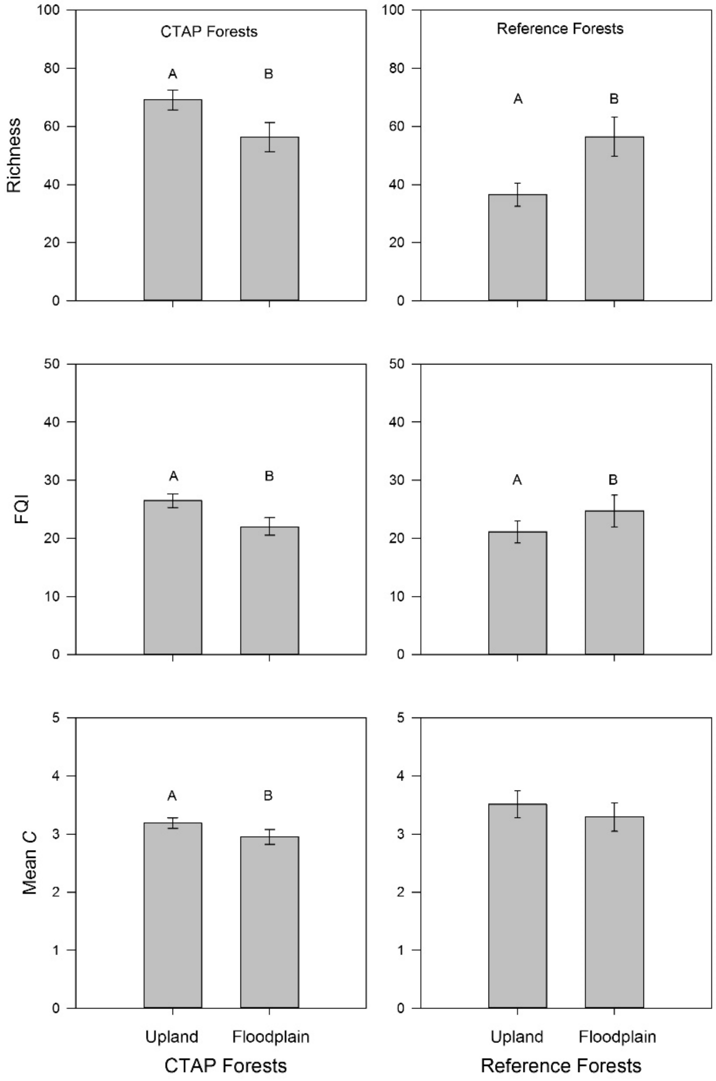


Figure 1.5. Site FQA and richness comparisons between upland and floodplain forest across Illinois using both reference and random data (95% confidence interval bars shown). Significant comparisons indicated by different letters (ANOVA: CTAP Mean *C*- $p < 0.004$; Reference FQI- $p < 0.02$; all other comparisons, $p < 0.001$).

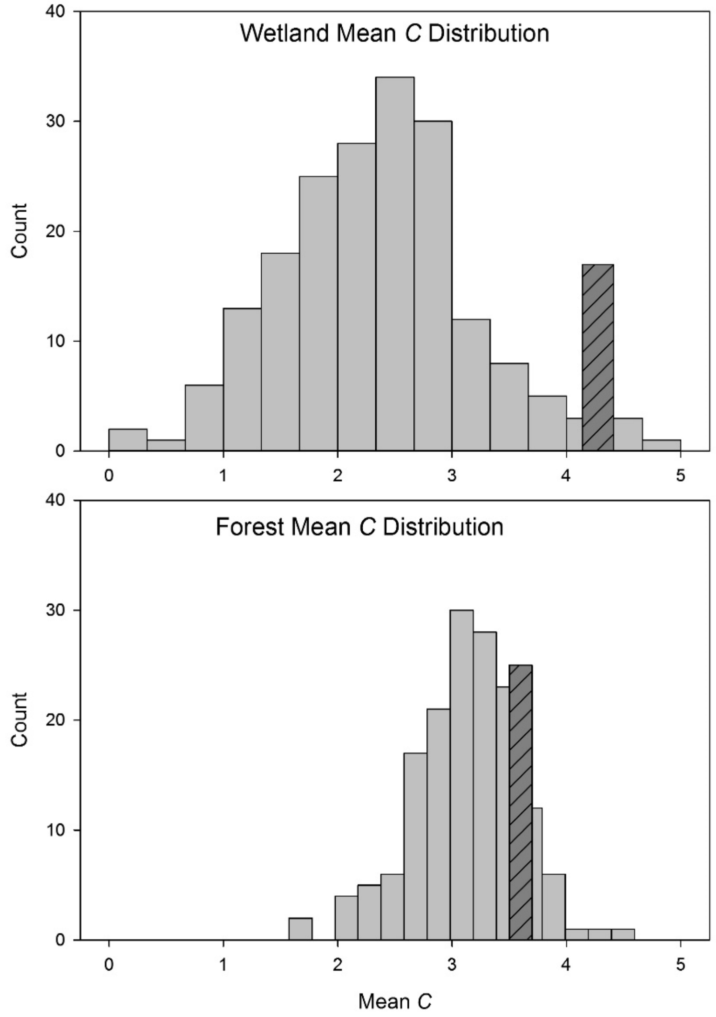


Figure 1.6. Histograms showing the distribution of Mean *C* values from randomly selected sites (light grey columns). Reference sites (dark grey striped column) are put into a single column bin at their mean value because their range was small and not of interest here.

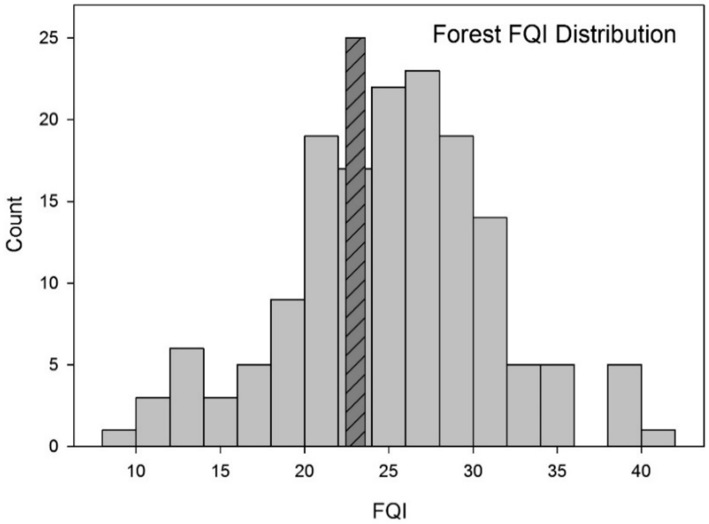


Figure 1.7. Histograms showing the distribution of FQI values from randomly selected forest sites (light grey columns). Reference sites (dark grey striped column) are put into a single column bin at their mean value because their range was small and not of interest here.

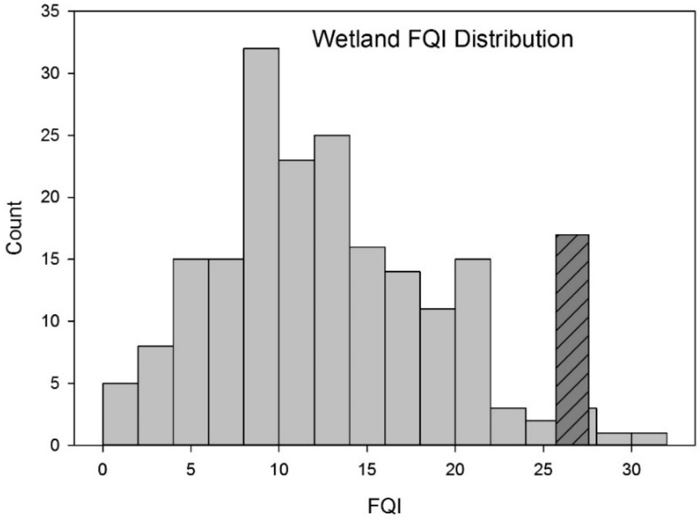


Figure 1.8. Histograms showing the distribution of FQI values from randomly selected wetland sites (light grey columns). Reference sites (dark grey striped column) are put into a single column bin at their mean value because their range was small and not of interest here.

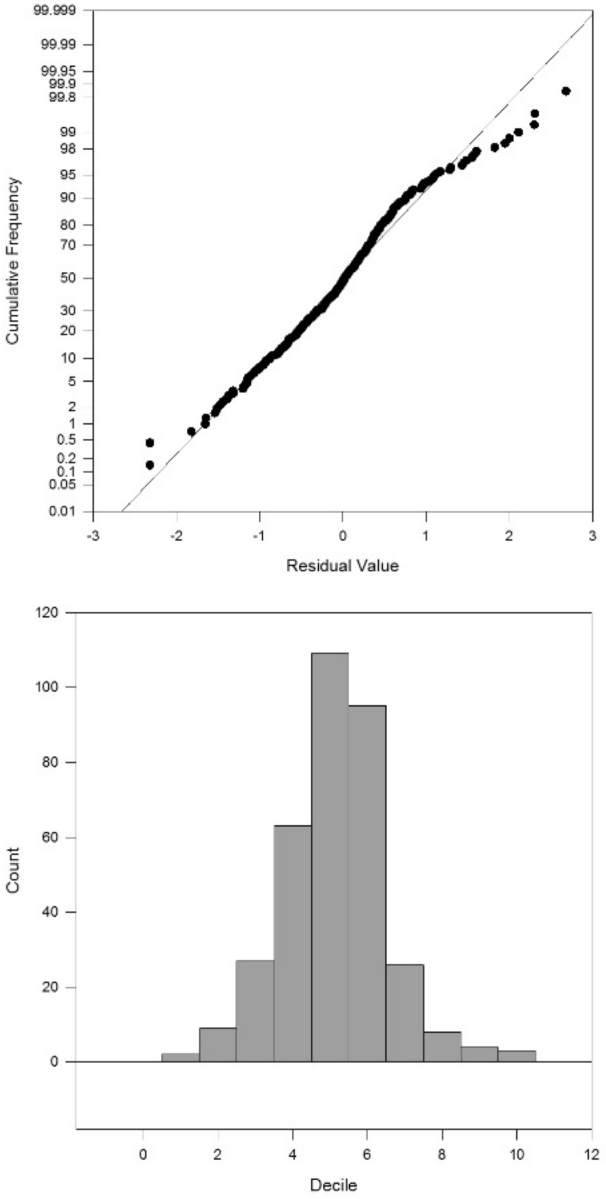


Figure 1.9. Histogram of residuals (bottom) and normal probability plot of residuals (top) from ANOVA comparison among site Mean *C* values from random wetlands and forests (see Figure 1.2).

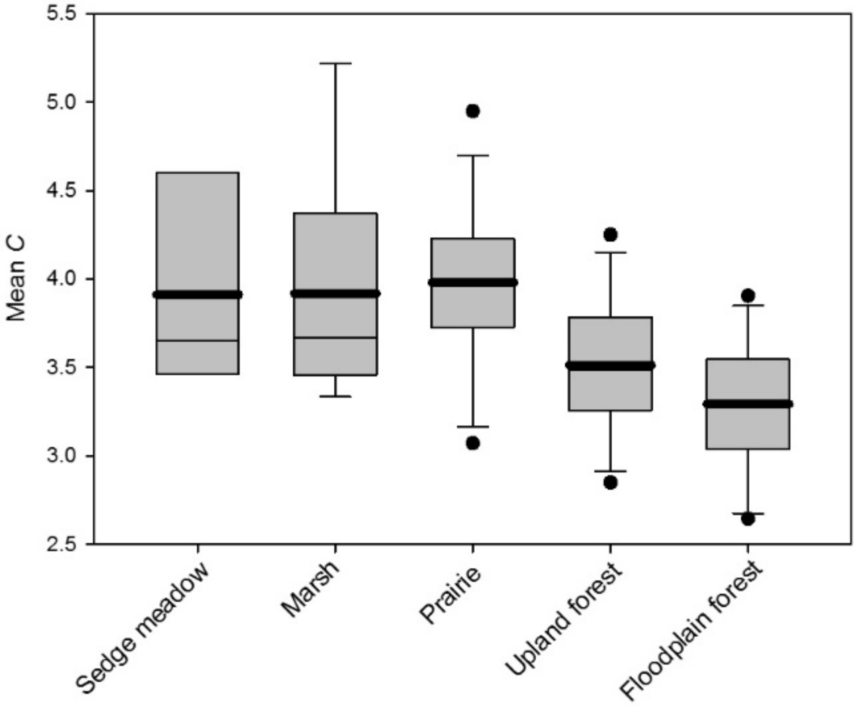


Figure 1.10. Mean *C* value box plots for reference site habitat types. The boundary of the box indicates the 25th percentile and 75th percentiles. Whiskers (error bars) above and below the box indicate the 90th and 10th percentiles, which sedge meadow lacked because at least nine points are required to calculate them. The thin line within the box marks the median and the thick line the mean. Dots indicate outlier values.

Table 1.1**.** Data summary. Number of sites per habitat type. Italicized names describe highly disturbed agricultural or cultural lands within the randomly selected dataset.

Habitat	**Site Selection**	**Specific Type Sampled**
Forest (n=176)	Random (n=157)	Most commonly Dry-mesic upland forest (typically oak-hickory), and Wet-mesic floodplain forest (typically ash-maple-elm)
	High-quality reference (n=23)	Dry-mesic upland forest (n=11) Wet-mesic floodplain forest (n=12)
Wetland (n=206)	Random (n=189)	Most commonly marsh, *wet old field, wet pasture*, sedge meadow, and emergent pond
	High-quality reference (n=17)	Marsh (n=9), sedge meadow (n=8)
Prairie (n=13)	High-quality reference (n=13)	Dry-mesic and mesic prairie (n=13)

Chapter 2: Regional Patterns in Floristic Quality

Abstract

Quantifying the conservation value of habitat patches in such a way that their values may be compared across regions has proven be an intractable goal. Floristic Quality Assessment (FQA) is an increasingly popular and influential way to measure the conservation value of areas based on their floras. A site's FQA value estimates how much it has been degraded by anthropogenic disturbances, as reflected in its plant species composition and diversity. The popularity of FQA's two metrics (Mean *C* and FQI) stem in part from their ease of use and flexibility, as any site value can be compared with any other across a region. This assumption of regional equivalence and comparability is increasingly questioned, however, as other ecological metrics, such as species richness, vary regionally due to factors unrelated to a sites conservation value (e.g., climate, topography). It is not known if Floristic Quality values vary across regions. I compared site Mean *C*, FQI, and species richness values across Illinois' forests and wetlands. Mean *C* increased with latitude overall, and this relationship was even stronger for forests. Forest species richness weakly increased to the south and to the west. Weak quadratic relationships between latitude and FQA were found for wetlands. Floristic Quality differences were detected across statewide political boundaries and ecoregions, although these were ultimately attributed to latitude. In summary, surprisingly weak richness patterns at large scales highlight its responsiveness to local factors. Stronger and more consistent responses of FQA to latitude suggest that regional ecological influences *may* confound site FQA values comparisons across large spatial-scales, but only for certain habitat-types and certain metrics. Determining if regional variation represents an error or bias in FQA values will ultimately depend on discriminating between causes of geographic variation, namely: 1) human disturbance legacies, 2) inherent ecological characteristics of habitats, or 3) unintended regional bias in assigned species values (*C*-values). Nonetheless, FQA users should remember that because site-level human disturbance levels will explain far more

variation in values than region, metrics will be suitable for their intended uses in most cases.

Keywords

Anthropogenic disturbance, Conservation value, Ecoregion, Floral assemblage, Floristic Quality, Forest, FQA, Habitat degradation, Latitudinal diversity gradient (LDG), Species Conservatism, Wetland

Introduction

Conservation "hotspots" are often mapped and compared at global scales (Myers et al. 2000, Brooks et al. 2006). Such maps are often underpinned by ecological factors such as the global latitudinal diversity gradient (LDG) (Visser et al. 2014), and by anthropogenic land use patterns. However, comparisons of the conservation value of habitat patches across landscapes— the scales at which most conservation and restoration efforts happen— are not effectively made, because the means to do so are not well-developed (Stein 2002, Ejrnæs et al. 2008). Existing vegetation-based assessments of habitat conservation value at these scales have used endemic species, species diversity, phylogenetic diversity, rare species, indicator species, or indices that combine several metrics to rank areas (Mack et al. 2008, Kiera et al. 2009). But, such measures have not been readily adopted by conservation practitioners, either because they do not have a broad enough ecological scope (e.g., focus on specialized taxa or specific site properties), or because they are too difficult to generate, interpret, or compare across areas (Niemi and McDonald 2004, Pearman et al. 2006). Thus, generally applicable measures for assessing the conservation value of natural area vegetation, which are simple and flexible enough to be readily adopted by conservation practitioners, are wanting (Niemi and McDonald 2004, Taft et al. 2006).

It is in this context that Floristic Quality Assessment (FQA) has increased in use and influential over the past 30 years (especially in North America) (Appendix). A site's plant species list is used to calculate its Floristic Quality value with two primary metrics.

The first is the mean of *a priori* conservation values assigned to each species. Species conservation values are termed Coefficients of Conservatism, or *C*-values (further described in Methods), and their average for a site is the Mean *C*. The second metric incorporates site species richness and Mean *C* into the Floristic Quality Index (FQI). Floristic Quality metrics (Mean *C* and FQI) are said to measure an area's biological Integrity, conservation value, or habitat degradation level via the amount of human induced changes to that area's plant assemblage (Taft et al. 1997, Spyreas and Matthews 2006, Mack 2007). FQA metrics are popular for a number of reasons: they perform well, being characterized as "remarkably robust" for their stated objectives (Miller et al. 2006); they capture information that is unique among existing conservation measures (Appendix); they only require a plant species list for their calculation; their values can be compared over time (Spyreas et al. 2012); and their values can easily be compared between all types of sites within a given region (FQA regions are most commonly states/provinces, Appendix).

The ability to compare varied sites across large regions with a single metric has spawned many different uses. For example, they are used in research on site restoration success and failure across vast areas (e.g., Mushet et al. 2002, Balcombe et al. 2005, Matthews et al. 2009a), and to study the causes and patterns in statewide habitat degradation (Nichols 2001, Milbauer and Leach 2007, Kraszewski and Waller 2008, Rooney and Leach 2010, Smart et al. 2011). FQA is also influential because of its direct use as a habitat evaluation criterion by various agencies. For example, the Nature Conservancy uses FQA to "prioritize high quality natural areas to include in their natural areas networks" (Rocchio 2007). Some federal, state, county, and municipal agencies have legally mandated that FQA metrics be used to evaluate sites across their jurisdictions. For example, a fixed benchmark FQI value threshold of 20 has been used by state and federal agencies as the determinant criteria for mitigating wetland destruction, as regulated under the Clean Water Act (Herman et al. 1997, Streever 1999, Herman et al. 2001, Matthews and Endress 2008, Matthews et al. 2009b). Such use assumes that a value of 20 is an equivalent and absolute value, such that any area's FQI value can be held up against this benchmark and its meaning will be equivalent. For example, any site scoring below 20 is considered easily replaceable and might receive less protection (Wilhelm 1992, Matthews et al. 2005).

Use of FQA across large regions has therefore been encouraged as objective and repeatable; a quantitative criterion that has encouraged “uniformity in natural area evaluation, enabling planners, land custodians, ecologists and other practitioners to make standardized comparisons among various open land areas” (Wilhelm and Ladd 1988). Region-wide metric uses assume that site values only vary due to anthropogenic factors (e.g., livestock grazing, logging, exotic invasive species, ecological management regimes) (Spyreas and Matthews 2006, Milbauer and Leach 2007, Spyreas et al. 2010). Thus, Floristic Quality is otherwise assumed independent of location, even across large U.S. states and Canadian provinces (e.g., in Florida: Cohen et al. 2004, Ohio: Mack 2006, the Dakotas: Hargiss et al. 2007, Colorado: Rocchio 2007, Alberta: Raab and Bayley 2012).

Despite the assumption of regional neutrality, other conservation metrics and ecological indicators are inherently varied across regional ecological gradients (e.g., Pearman et al. 2006, Muratet et al. 2008). For example, plant species richness varies inversely with latitude at global scales, and it varies regionally in association with many other non-anthropogenic factors (e.g., productivity, climate, disturbance regime, glacial history, elevation, edaphic conditions, biogeography) (Mutke and Barthlott 2005, Sarr et al. 2005). As richness is a constitutive component of FQI, FQI is also expected to respond to spatial variation in these factors. On the other hand, Mean *C* is not tied to richness and it might be assumed to be independent of spatially varying ecological factors. The level to which FQA values vary due to regional ecological gradients that are unrelated to human disturbance, would need to be accounted for in their region-wide use (Bernthal 2003). If not, this would confound the use of FQA in making conservation decisions. For example, land acquisition and protection in regions where ecological characteristics naturally generate lower FQI values could be disfavored.

Studies have not yet determined if FQA values vary within a region. The largest existing study found latitudinal differences in FQI, Mean *C*, and species richness in ponds and lakes across Wisconsin (USA) (Nichols 1999). However, Nichols (1999) also pointed out that observed patterns could have simply been reflecting the different habitat-types sampled in different regions (oligotrophic versus eutrophic ponds; Chapter 1). Conversely, neither Cohen et al. (2004) nor Miller et al. (2006) found latitudinal differences in wetland FQA values. Finally, Johnston et al. (2010) found strong latitudinal variation in values from wetlands bordering the Great Lakes. Unfortunately, because they compared *C*-values from different states, their results must be viewed cautiously (i.e., because each state's floras have been assigned different *C*-values for a given species; Appendix). Additionally, because they only reported FQI values, it was not possible to determine if their observed patterns were due to differences in richness, Mean *C*, or both.

A considerable amount of data is required to make conclusions about statewide Floristic Quality patterns and regional trends in FQA metrics, and this has hindered their understanding thus far. If regional variation in FQA is to be understood, analyses *must* be conducted at statewide scales, because each species *C*-value is scored with respect to its *simultaneous behavior across the entire state* (see Methods). Therefore, because values in one area of a state might be counterbalanced by another part, assessments of regional FQA patterns must have complete statewide coverage. The following additional criteria for data exist if regional FQA patterns are to be understood: 1) habitat-type should be controlled for; 2) an unbiased site selection criterion with probabilistic sample design should be used (randomized, stratified-systematic, etc.) (Genet and Olsen 2006); 3) site sampling should be consistent (sample intensity, area, or season can affect FQI values, Appendix); 4) studies should examine more than wetlands or aquatic habitat-types.

This study seeks to determine if there are regional patterns in site-level Floristic Quality values or species richness across Illinois (USA). To assess general patterns, this study will quantify variation with respect to latitude and longitude. Illinois is an excellent test case because of its considerable latitudinal and longitudinal range, along which species richness is projected to increase from north to south, and from west to east (Withers et al. 1998, Qian et al. 2007, Qian and Ricklefs 2007). Next, because other ecological metrics are known to vary according to the ecological factors that typify certain specific regions (e.g., areas with specific biogeographical history or topography), FQA values among areas with shared ecological characteristics (ecoregions) will be compared. Finally, because conservation value is often assessed by agencies with specific jurisdictions, patterns across political boundaries will be evaluated. The specific tests and predictions in the study are as follows:

1) Studies that span North America predict that plant species richness will increase from north to south, and west to east, across Illinois. Alternatively, the assumptions underlying the use of FQA are that their values do not vary across the state. This study seeks to determine if site FQA values (Mean *C* or FQI) and species richness values in Illinois' wetlands and forests vary with latitude or longitude.

   It is expected that species richness and FQI will increase from north to south due to the effect of richness, while Mean *C* will not vary regionally. Similarly, for longitude, richness and FQI are expected to increase towards the east, while Mean *C* is not expected to vary.

2) Ecological characteristics at smaller spatial scales (e.g., ecoregions) often surpass factors operating at continental scales in explaining floristic and vegetation patterns. This study seeks to determine if average Floristic Quality levels in wetlands and forests vary across smaller ecological gradients across Illinois represented by ecoregions.

   It is expected that Mean *C* values will not differ across ecoregions.

3) Finally, because the same conservation value criteria are often used by agencies across different jurisdictions, this study asks if average site FQA values in wetlands and forests vary across Illinois political boundaries (U.S. Army Corps of Engineers districts).

   It is expected that FQI will vary across these boundaries (due to species richness' influence), but Mean *C* will not.

Methods

Sites were sampled as part of Illinois' Critical Trends Assessment Program (CTAP), which samples randomly selected emergent wetlands and forests throughout Illinois (Molano-Flores 2002). Grasslands were excluded from analyses because native Illinois grassland (prairie) has largely been eliminated from Illinois (Illinois Department of Natural Resources 1994). Forests were visited from south to north across Illinois, from 15 May through 30 June, and wetlands were sampled between 1 July and 31 July. Forest sampling was confined to homogeneous areas of forest with respect to aspect, hydrology, topography, and forest type when possible. Sampling was generally done at least 50-m from the forest edge. Sample species lists in forests were created by recording all vascular plants observed in twenty 0.25-m^2 quadrats (0.5-m X 0.5-m) spread across two randomly selected 50-m transects at each site. These ground layer quadrats were nested within larger tree (10-m X 50-m) and shrub (2-m X 50-m) sample plots, from which additional woody plants were also recorded. In addition, species within a plot (10-m X 50-m plot) on a third randomly selected transect were recorded. In wetlands, vascular plants were recorded from a 40-m X 50-m plot placed on the edge of the wetland and oriented to span its hydrological gradient from the upland inward. A total of 157 forests and 189 emergent wetlands from the years 1997–2012 were sampled, with the vast majority sampled from 2007-2012. See Carroll et al. (2002) for detailed site selection and sample protocols.

Site FQA values are based on Coefficients of Conservatism (*C*-values) ranging from 0 to 10 that have been assigned to each native plant species (and some sub-species) in Illinois. Highly Conservative plants (7-10) are only found in minimally degraded natural areas, while species that persist in, or readily invade degraded areas, are given lower numbers. Non-natives receive $C = 0$. A species *C*-value is only valid for the state it is assigned in, as it is determined by considering the species occurrences and behavior across that state. Some taxa may show different behaviors in different regions of a state, but the *C*-value reflects a species' average statewide Conservatism. Species *C*-values were obtained from Taft et al. (1997), with minimal modifications (e.g., Phragmites australis was considered non-native: Spyreas et al. 2004). Some taxa were counted as native when sample sites were within their native range and exotic when the sample site was elsewhere in the state (e.g., northern versus southern sites for *Taxodium distichum, Pinus echinata, Robinia pseudoacacia, Catalpa speciosa, Viburnum recongnitum*). In some cases, where a specimen could not be identified to species, the lowest *C*-value in its genus that it resembled was used as a conservative estimate (e.g., *Stachys* sp. $C = 5$). Or, in a few cases, where two species could not be distinguished and their values were close to one another their average value was used (e.g., *Impatiens capensis* $C = 2$ and *I. pallida* $C = 4$; *Ulmus rubra* $C = 3$ and *U. Americana* $C = 5$). FQA metrics were calculated as follows. The Mean Coefficient of Conservatism ($\overline{C}$ or Mean *C*) for all species detected was calculated as ($\overline{C} = \sum C/S$), and the Floristic Quality Index was calculated as ($\mathrm{FQI} = \overline{C} *(\sqrt{S})$), where S is total species richness, and C is each species' Coefficient of Conservatism value. We included non-native species for both S and C in calculations (see, Spyreas et al. 2012).

Analysis

Simple ordinary least squares regressions were performed to determine longitude and latitude relationships. In a few instances, simple regressions against latitude did not meet the assumption of normality (Shapiro-Wilk: $p < 0.05$), and based on the shape of their residuals, they were then fit by a regression model with a quadratic term. FQA values were compared across U.S. Army Corps of Engineers districts (map of district boundaries available- http://www.lrc.usace.army.mil/Portals/36/docs/regulatory/maps/ilmap.pdf), and natural divisions (ecoregions) across Illinois with ANOVA. Army Corps districts where used because different districts may have common FQA thresholds (e.g., FQI < 20), and ecoregions were used to compare Floristic Quality values among areas of relative ecological homogeneity, with similar topography, glacial history, bedrock, soils, native plants, and animals (Schwegman et al. 1973, Bailey 1995, Woods et al. 2006). Two ecoregions were lumped for this analysis because they had too few sample points (Middle Mississippi Border and Western Forest-Prairie). FQI and richness values were not compared among ecoregions because there were unequal numbers of wetland versus forest site samples in ecoregions, which could reflect species richness differences due to sampling. The ecoregion ANOVA did not meet assumptions of normality or equality of variance (Shapiro-Wilk: $p < 0.05$), because regions with smaller numbers of sample points had highly inflated variances (e.g., Shawnee Hills, Coastal Plain). Examination of the distribution of the test residuals, however, showed a good model fit, so parametric tests were assumed to be robust (Figure 2.11). Nonetheless, a Kruskal-Wallis One Way ANOVA on ranks was also conducted, and since it also found a highly significant difference between the ecoregions ($p < 0.001$), significant differences were interpreted. Post-hoc pairwise comparisons were made using the Holm-Sidak method ($p < 0.05$). Finally, in order to determine if ecoregion was important to site values beyond the influence of latitude, ANCOVA was carried out.

Global Polynomial Interpolation in ArcMap 10.1 was used to illustrate localized spatial patterns in FQA metric values. Global Polynomial Interpolation (GPI) fit a smooth surface of gradual changes among data points that was defined by a third-order polynomial function. The “Extract By Mask” function was then used to bound the surface to Illinois. All ANOVAs and regressions were conducted using SigmaPlot for Windows12.0, (Systat Software, Inc., Chicago, IL), and ANCOVAs were conducted using the Proc GLM procedure in SAS 9.2 (SAS Institute Inc., Cary, NC).

Results

Simple ordinary least squares regression showed that site Mean *C* decreased from south to north in all cases (All sites- $p < 0.001$, Forest- $p < 0.001$, Wetland- $p < 0.01$; Table 2.1, Figure 2.6, 2.9, 2.10), as did FQI and richness in forests (FQI- $p < 0.001$, Richness- $p < 0.001$; Table 2.1, Figure 2.10). Wetland FQI and Mean *C* (but not richness) showed a significant quadratic relationship with latitude (Mean *C*- $p < 0.001$, FQI- $p = 0.002$; Figure 2.8 & 2.9), but only Mean *C* was better described by the quadratic versus simple linear relationship ($r^2_{adj} = 0.11$ versus $r^2_{adj} = 0.03$). Forest species richness was the only value significantly related to longitude; values weakly increased from east to west ($p < 0.05$; Table 2.1). ANOVA results indicated that Mean *C* differed among ecoregions (Mean *C*- $F_{12, 333} = 4.16$, $p < 0.001$; Figure 2.4). Mean *C* differences among ecoregions were nullified after the influence of latitude was accounted for (ANCOVA: Mean *C*- $r^2 = 0.19$; latitude $p < 0.0001$, ecoregion $p = 0.14$, lat * ecoregion $p = 0.04$). ANOVA results indicated that both FQA values differed among USACOE districts (FQI- $F_{3,342} = 6.2$, $p < 0.001$; Mean *C*- $F_{3,342} = 7.09$, $p < 0.001$; Figure 2.5).

Discussion

Contrary to expectations, Mean *C* values increased from north to south across all habitats, although a non-monotonic, quadratic relationship described the latitudinal trend in the subset of wetland habitats. FQI results were as predicted, increasing to the south, except for the quadratic relationship in wetlands. Species richness increased to the south as was expected, but only in forests. It also unexpectedly increased to the west in forests, although both richness trends were weak. Both FQA metrics differed among U.S. Army Corps of Engineers district boundaries, and Mean *C* values differed among several ecoregions across the state, however, these categorical differences were better explained by site latitude. As the first study to characterize differences in Floristic Quality values across an entire FQA region, it demonstrated a considerable latitudinal trend, especially in forests. In addition, previously suggested gradients in richness across the region (e.g., Withers et al. 1998), were either unsupported (wetlands), supported (latitude for forests), or contradicted (longitude for forests), but all of these relationships were very weak.

The consistent trends observed for Floristic Quality values were unexpected as FQA metrics are generally considered to be regionally independent within their state of origin. Some authors have expressed concerns that if Floristic Quality values were found to vary regionally, and if this was not accounted for in site comparisons, a confounding bias would be present (Herman et al. 1997, Bernthal 2003). For example, this could mean that FQA values would be reflecting the effect of the ecological characteristics of different regions (e.g., topography, climate), as opposed to their human disturbance levels and conservation value. The potential magnitude of confounding geographical variation in Floristic Quality values is illustrated by the magnitude of the latitude effect. Almost a full point, a 39% increase, in average Mean *C* values can be expected across Illinois wetlands and forests from north to south (2.3 predicted by the regression at the northern state boundary compared to 3.2 in the south; Figure 2.6), with an approximately 8 point difference in FQI values (Figure 2.7).

One of Karr's (1987) primary tenets for biological assessments is "regionalization"; as metric values may have different meanings in different regions, they should be calibrated for regions. Vegetation based Indices of Biotic Integrity (IBI) have long been created for, or relativized by, specific regions within states (Genet and Olsen 2006). Some authors have suggested that if regional variation in site-level FQA values exist, this would necessitate that site comparisons are restricted to specific regions in states (Nichols 1999, Reiss 2006). But, determining the appropriate regions and scales to confine comparisons would require detailed knowledge of Floristic Quality patterns within a state, perhaps even beyond that of this study (e.g., Figures 2.1-2.3). As an alternative, Johnston et al. (2010) suggested adding a simple correction factor based on latitude when comparing FQI values across regions to address a potential latitude bias. In some respects, the current study's results would seem to support this simple recommendation, as ecoregions were not determinant of site Floristic Quality levels above and beyond the linear effect of latitude.

However, resultant maps also illustrate that a simple linear correction could also be a facile response to what are actually rather nuanced patterns (e.g., Figures 2.1-2.3). Furthermore, the variation explained by latitude was not high, meaning that the predictive power of its effect on FQA values is actually quite limited. Finally, a single, linear correction ignores that many other factors, especially anthropogenic ones (e.g. human disturbance, land-use intensity), are also correlated with latitude, and it is not clear what is being corrected for. Ultimately, it should be the *cause* of the latitudinal gradient in Floristic Quality that determines if regional differences are confounding, and if they should be corrected for. Three potential explanations for regional FQA patterns that need to be discriminated are: 1) anthropogenic disturbance legacies and land-use patterns, 2) natural variability associated with inherent ecological gradients, and 3) regional bias in assigned species *C*-values— as when the floras of certain regions have unwittingly been assigned higher *C*-values. Each one of these explanations are discussed below.

Anthropogenic disturbance legacies and land-use patterns

Latitudinal variation in Floristic Quality has been anecdotally attributed to human disturbance trends (Nichols 1999, Reiss 2006). If regional FQA patterns are reflective of anthropogenic disturbance and land-use, values would reflect what they have been assumed to measure— anthropogenic degradation of site Floristic Quality— and regional differences may (continue to) be ignored in FQA. Lower Floristic Quality sites could have incurred more anthropogenic disturbance directly (e.g., livestock use, logging), or their landscape context and the condition of their surrounding areas could cause their lower site values. For example, habitat buffering, or an unavailability of seed and propagule pressure from Conservative species in local species pools might slow recovery from past site disturbances (Lopez and Fennessy 2002, Cohen et al. 2004, Mack 2007, Matthews et al. 2009a).

For Illinois, the southern and the western regions have more forest cover, a more rugged topography, and/or more protected public natural areas (i.e., Shawnee National

Forest) (Illinois Department of Natural Resources 1994). Additionally, the north is more heavily urbanized (i.e., the Chicago Region), and the central part of the state is the most extensively converted to intensive crop production (Illinois Department of Natural Resources 1994). Integrated land-use maps clearly illustrate that urbanization and agriculture are the dominant disturbances and land-sues in the state (Theobald 2013), such that regional variation in these two variables would seem to explain the broad FQA value gradient observed. But, when one considers the relatively high Floristic Quality of wetlands in the highly urbanized northern third of the state (Figure 2.2), it is obvious that coarsely mapped land-use on its own (e.g., 80 m resolution, Theobald 2013), does not adequately predict all the regional FQA patterns demonstrated in this study. Therefore, either localized, site-level human disturbances explain locally idiosyncratic or unexpected results (e.g., FQI and richness in northeastern wetlands), or other regional factors are responsible for localized patterns (e.g., natural ecological gradients, discussed below). To summarize, there is good support that land-use and disturbance, operating at both local and regional scales, drive most of the observed spatial patterning in Floristic Quality, however, other potential explanations of variation require future consideration to assess their relative contributions.

Natural variability and ecological gradients

Regional FQA patterns could also reflect natural ecological gradients that are unrelated to human degradation across sites. This would mean that if the same sites were sampled across Illinois prior to European settlement, the same relative regional Floristic Quality patterns would be observed (although values would be higher). In terms of identifying causal variables, any number of ecological and environmental variables show strong latitudinal variation across Illinois and could potentially explain observed patterns. Global latitudinal diversity gradients (LDG) are explained using evolutionary (i.e., diversification rates, time for speciation) or ecological variables (topographical heterogeneity, temperature, precipitation, net primary productivity) (Visser et al. 2014). At smaller regional scales, other variables can augment these in their importance (e.g., disturbance regimes, edaphic conditions, glacial and biogeographical history, biotic interactions, Mutke and Barthlott 2005, Sarr et al. 2005). While it is one thing to identify the correlative variables that best explain the ecological gradients associated with Floristic Quality patterns, it is another to determine the mechanisms. The proposed *mechanisms* behind LDG's environmental variables include: species' physiological adaptations and traits linked with harsh or stressful environments (e.g., cold or arid climates); higher productivity and biomass leading to larger populations and lower extinction rates; and greater habitat diversity/niche-space (Visser et al. 2014). These mechanisms alter richness, and there would seem to be little reason why a species composition based FQA metric like Mean *C* would share its explanatory mechanism with diversity.

Nonetheless, regionally explanatory mechanisms do seem to overlap. Harsh or stressful abiotic conditions have been suggested as predisposing certain habitats to higher Mean *C* values (Andreas et al. 2004, Rentch and Anderson 2006, Bried et al. 2013). At the community level, physiologically stressful environments may favor "stress-tolerant plants" in assemblages (Grime 1979, Keddy and MacLellan 1990, Wisheu and Keddy 1992), and these types of species may tend towards higher *C*-values (Taft et al. 1997). A latitudinal "stress gradient" across Illinois that would explain the current study's results is not apparent. Cold-stress does not increase to the south, while the aridity gradient from north to south across Illinois is minimal (Kartesz 2014b). Productivity is another relevant variable, as greater productivity could allow for greater representation by matrix "competitor" species in plant communities (Keddy and MacLellan 1990, Wisheu and Keddy 1992), which tend to have relatively high *C*-values (4-6), as opposed to weedier less-conservative taxa (Taft, Wilhelm et al. 1997). And, net primary productivity could approximate this study's regional Floristic Quality pattern (e.g., Kucharik et al. 2000), although again, this gradient does not seem strong enough to explain the latitudinal Floristic Quality gradient. Finally, the frequency, intensity, or type of disturbances that naturally characterize regions (e.g., fire, wind-throw, hydrology) could cause regional patterns by increasing the representation of weedy species in communities (Collins and Pickett 1987, Gilliam et al.

1995, McLachlan and Bazely 2001, Meiners et al. 2002), thereby, lowering Mean *C* values. In any case, it would also be interesting to know if regional ecological gradients (e.g., productivity, soil properties, climate) act on FQA values by favoring higher *C*-value species in different regions, or if they act by allowing a site's species to be more tolerant of, or recover more quickly from, anthropogenic disturbances in different regions. Nonetheless, because regional richness and FQA patterns were qualitatively similar in some cases (e.g., forests and latitude, Table 2.1), this suggests that they could share potentially explanatory ecological variables and mechanisms across regions.

When considering the relative support for ecological gradients at explaining Floristic Quality patterns it is informative to note that the only other studies that have reported FQA values across a region have suggested that their values increase to the north (in Wisconsin, Michigan, and Florida, see, Nichols 1999, Reiss 2006, Johnston et al. 2010), the opposite direction from what was demonstrated across Illinois. Therefore, any explanatory ecological gradient for FQA values is probably not shared across these states (e.g., a negative relationship with growing degree days suggested in, Johnston et al. 2010). Alternatively, while human land-use patterns vary in their latitudinal direction among these states, they are all congruent with observed FQA patterns. This supports human disturbance as the predominant determinant of regional FQA patterns.

Discerning regional bias in assigned species *C*-values

Another possible explanation for observed FQA patterns is *C*-value assignment during the original statewide scoring process (e.g., an unperceived affinity for southern floras). This would mean that southern ranging species in Illinois have been assigned *C*-values beyond what their occurrences in pristine versus disturbed habitats support there, and this would lead to higher site-level FQA values in the south. Determining whether such a bias or error in scores exists would be difficult, as it would not be clear whether to test the overall southern species pool or average site-level assemblages across the southern region. A few studies have attempted to assess individual species *C*-values for their accuracy, either post-score assignment, or during the scoring process itself (Nichols 1999, Mushet et al. 2002, Cohen et al. 2004, Forrest 2010, Landi and Chiarucci 2010, Bried et al. 2013). Matthews et al. (In Review) provide a useful template for how empirical data can be used to compare *C*-values for their relative over- or under-valuation within a flora, as they found evidence for small scoring errors among plant functional groups (e.g., perennials, woody taxa). However, there are no studies assessing systematic error or bias in scores across regions.

Future Study and Conclusions

FQA metrics are unique among existing site conservation value measures in that their values are readily compared across landscapes, because of their perceived regional independence (Miller et al. 2006). Whether this type of use is justified is contingent upon the extent and cause of statewide FQA patterns. At a basic research level, further studies should assess whether the regional patterns observed in this study exist in other states. More intensive study is needed to quantify and discern the causes in Floristic Quality's regional variation— with the goal of understanding contributions from land-use and human disturbance versus inherent ecological factors.

Distinguishing natural versus anthropogenic effects on vegetation is challenging because they often co-vary (Findlay and Houlahan 1997, Nichols 1999, Johnston et al. 2008). Indeed, some variables are even difficult to categorize as being either anthropogenic or natural. For example, non-native species abundances are of interest because they vary across regions (e.g., across Illinois, Spyreas et al. 2004), and because they depress native Floristic Quality and diversity levels (Spyreas et al. 2010). But, non-native plant abundances are a function of both anthropogenic and natural ecological factors (Martin et al. 2009), which obfuscates their meaning with respect to Floristic Quality. A promising research direction would be to analyze statewide FQA data from less-disturbed, remnant habitats. Such sites would be relatively free from human land-uses and disturbances, so they could reveal regional FQA patterning associated with inherent ecological gradients.

Finally, although the importance of a site's habitat-type in explaining its Floristic Quality level is debated (Chapter 1; Johnston et al. 2009), because forests and wetlands exhibited different regional patterns in this study, future inquiry must account for potential interactions between latitude and habitat-type when studying regional patterns.

The regional FQA trends revealed in this study could have considerable implications when site FQA values are compared across latitudes. However, it is important to remember that latitudinal patterns are only explaining a small amount of the variation found in site values (Table 2.1). Alternatively, where site human disturbance levels have been quantified, the amount of variation they explain is far greater (e.g., between 60-90%, Appendix). This means that while the potential for regional nuisance variation is a concern for FQA (Bried et al. 2013), the effects from human disturbance on site Floristic Quality values *far exceeds* effects from other potential sources of variation. Thus, I make the following specific conclusions for using FQA. First, the evidence does not suggest that regional patterns in Floristic Quality are strong enough, or are of a source that would confound its common usage. Second, in instances where values are compared across large latitudinal gradients, users should be aware that *some* of the variation in FQA values *could* be attributable to factors other than anthropogenic degradation, and this may need to be accounted for. However, efforts to account or correct for variation with a simple linear factor are not recommended because of the complexity of the spatial patterns observed in this study.

Acknowledgments

CTAP botanists past and present, Thomas Jefferson Benson, Scott "Pasta" Chiavacci, Dennis Skultety, David N. Zaya

Table 2.1- Simple regression results for statewide FQA predicted by latitude and longitude.

		Latitude					Longitude			
	N	Intercept	Regression coefficient	r^2_{adj}	*F*		Intercept	Regression coefficient	r^2_{adj}	*F*
All Sites										
Mean *C*	346	9.51	-0.17	**0.09**	**34.7*****		5.34	0.03	0.00	0.41
Forest										
Mean *C*	157	10.6	-0.19	**0.29**	**65.9*****		2.69	-0.005	0.00	0.91
FQI	157	115	-2.24	**0.25**	**51.5*****		-43.8	-0.77	0.01	2.14
Richness	157	225	-4.01	**0.08**	**14.6*****		-252	-3.56	**0.02**	**4.91***
Wetland										
Mean *C*	189	6.76	-0.11	**0.03**	**7.04****		10.0	0.09	0.00	1.71
FQI	189	23.4	-0.27	0.00	0.83		81.8	0.78	0.01	2.77
Richness	189	7.30	0.58	0.00	0.43		134	1.16	0.00	0.70

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

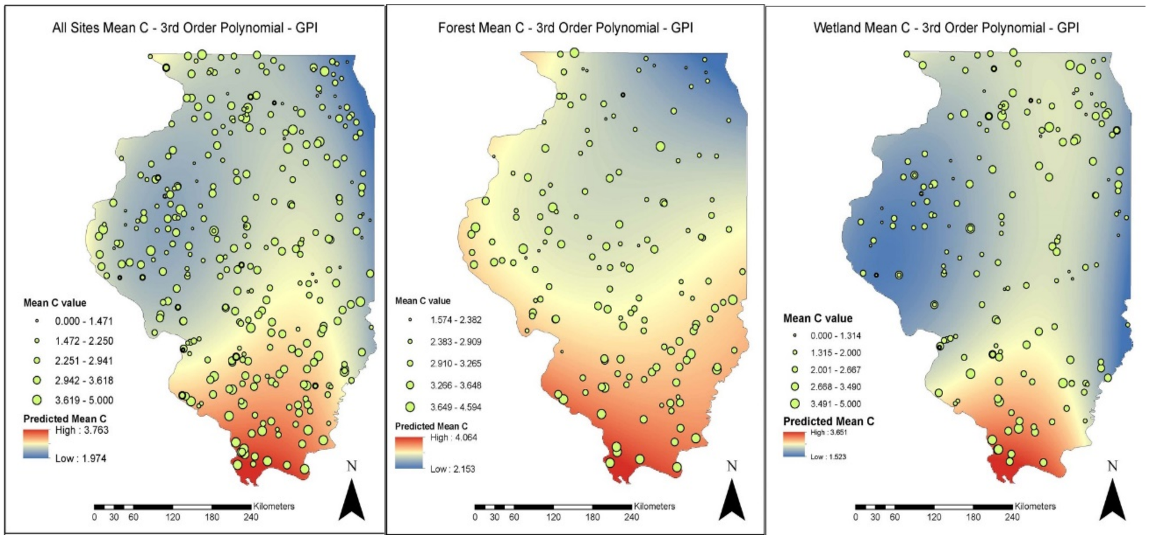


Figure 2.1- Spatial interpolation of site-level Mean *C* values in emergent wetlands, forests, and both combined across Illinois. Symbols depict actual sample points, sized by their relative values.

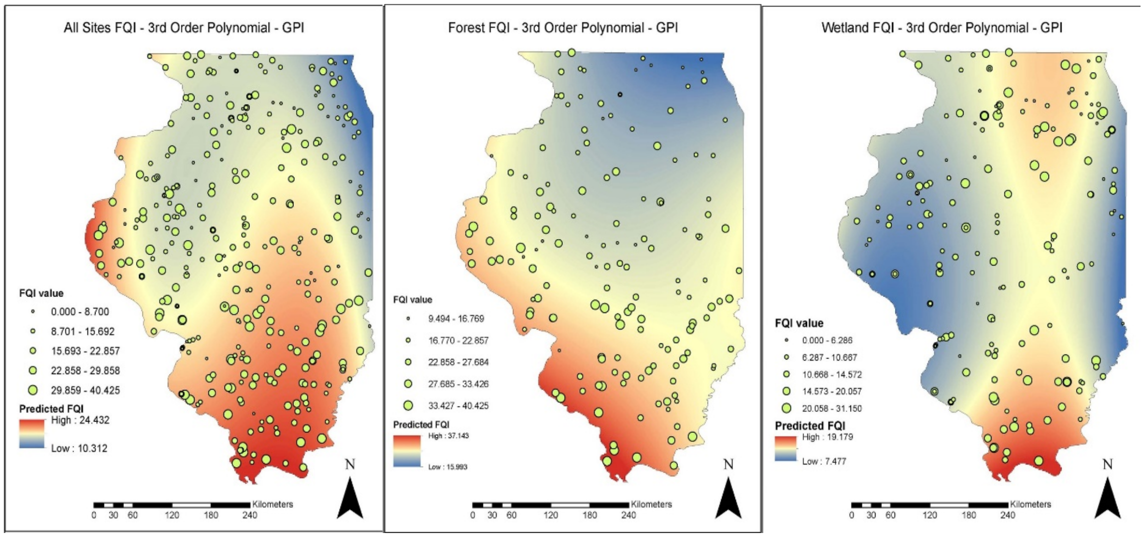


Figure 2.2- Spatial interpolation of site-level FQI values in emergent wetlands, forests, and both combined across Illinois. Symbols depict actual sample points, sized by their relative values.

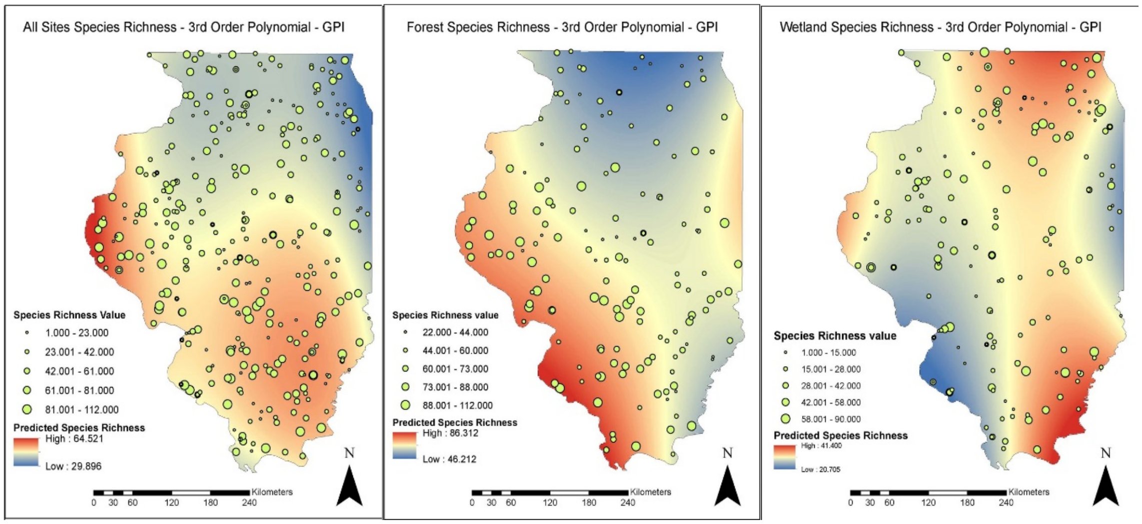


Figure 2.3- Spatial interpolation of site-level species richness in emergent wetlands, forests, and both combined across Illinois. Symbols depict actual sample points, sized by their relative values.

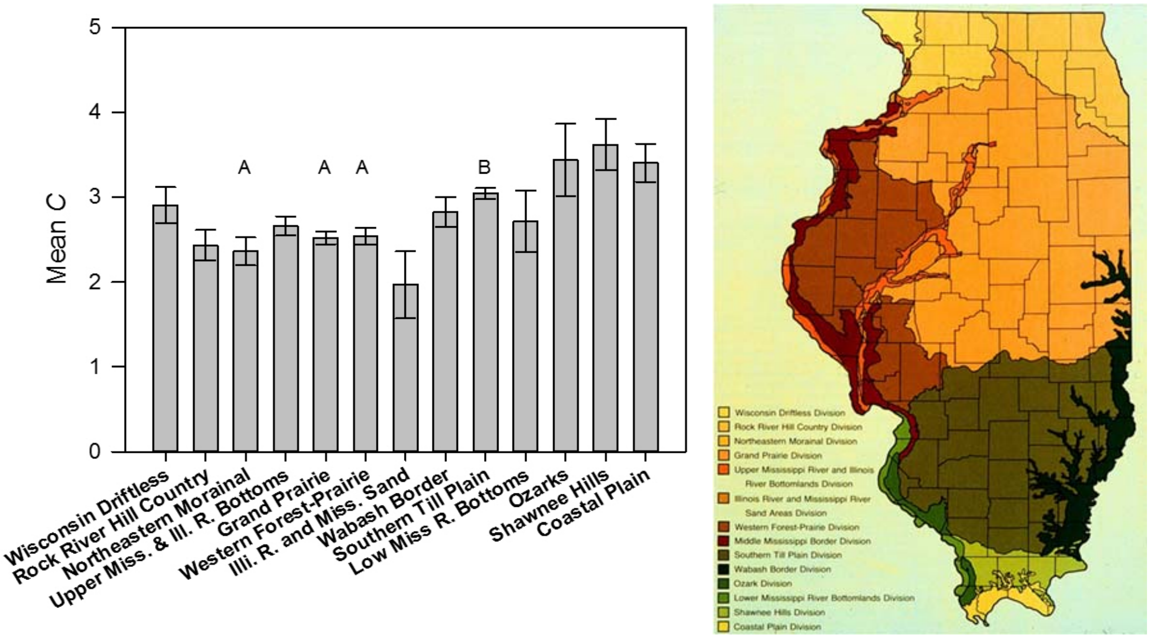


Figure 2.4- Mean *C* comparisons among Illinois ecoregions (natural divisions). Standard error (+/- 1 SE) bars are shown. A significant overall difference was found (ANOVA: $p < 0.001$), with significant post-hoc pairwise comparisons indicated by different letters (Holm-Sidak method: $p > 0.05$). Ecoregions without letters are not significantly different from any other ecoregion.

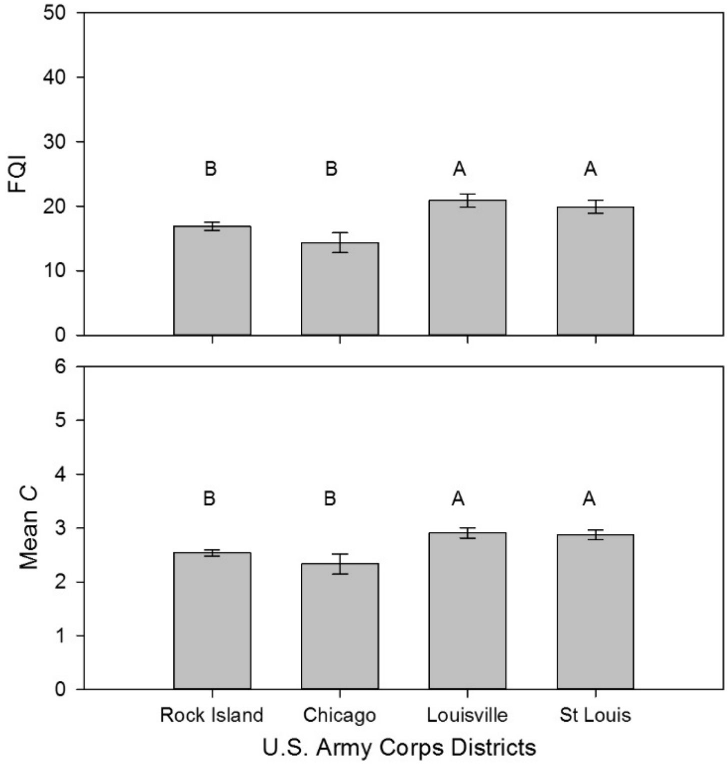


Figure 2.5 – Mean FQA value comparisons among U.S. Army Corps of Engineers jurisdictions in Illinois (for district boundaries see http://www.lrc.usace.army.mil/Portals/36/docs/regulatory/maps/ilmap.pdf). Standard error (+/- 1 SE) bars are shown. A significant overall difference was found (ANOVA: $p < 0.001$), with significant post-hoc pairwise comparisons indicated by different letters (Holm-Sidak method: $p > 0.05$).

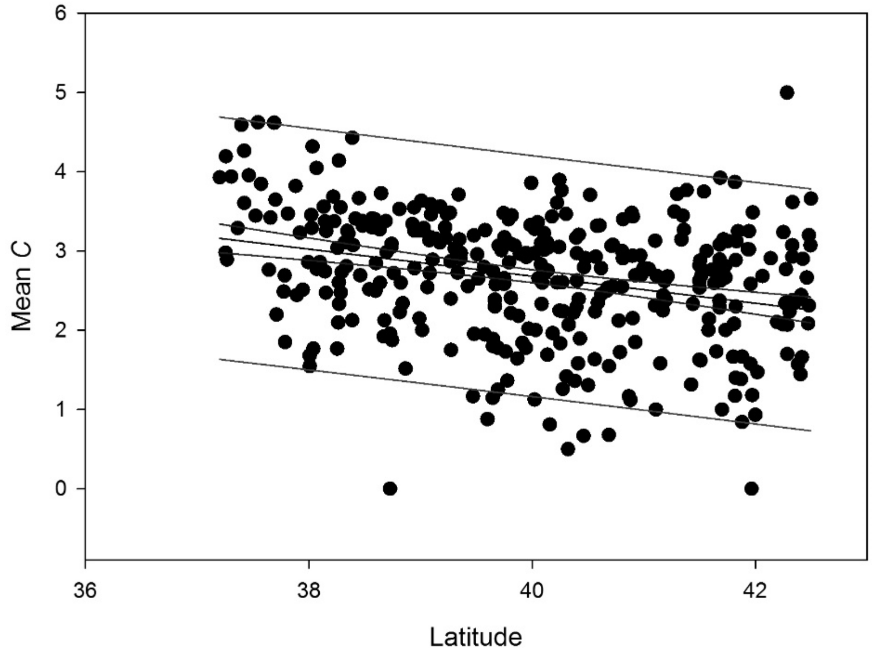


Figure 2.6. Scatter plot and regression line of the relationship between site Mean *C* and latitude (°N) for all sites combined. Regression line (centerline), is bounded by 95% confidence interval, and the outermost 95% prediction interval. See Table 2.1 for regression statistics.

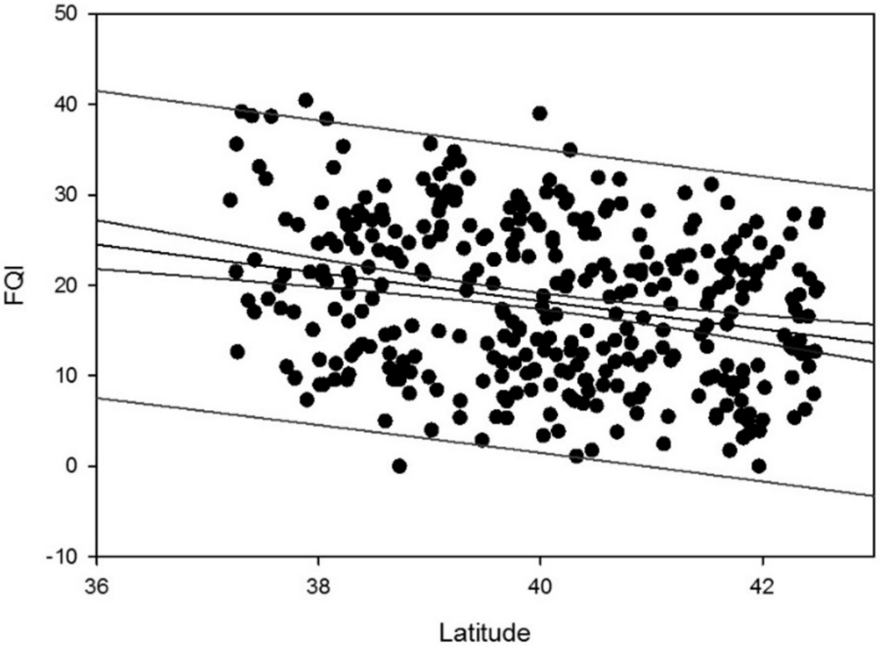
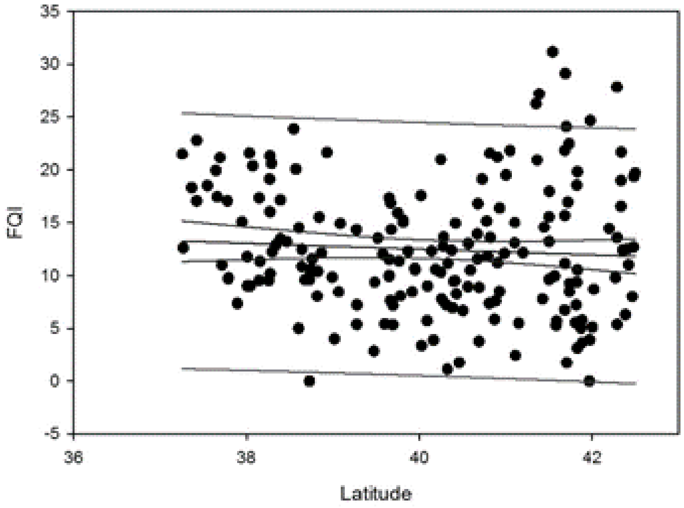


Figure 2.7. Scatter plot and regression line of the relationship between site FQI and latitude (°N) for all sites combined. Regression line (center line), is bounded by 95% confidence interval, and the outermost 95% prediction interval. Regression statistics are not shown as this graph is intended to illustrate the distribution of points and is not suitable for statistical tests due to sample differences between wetlands and forests and their effects on richness and FQI site values.



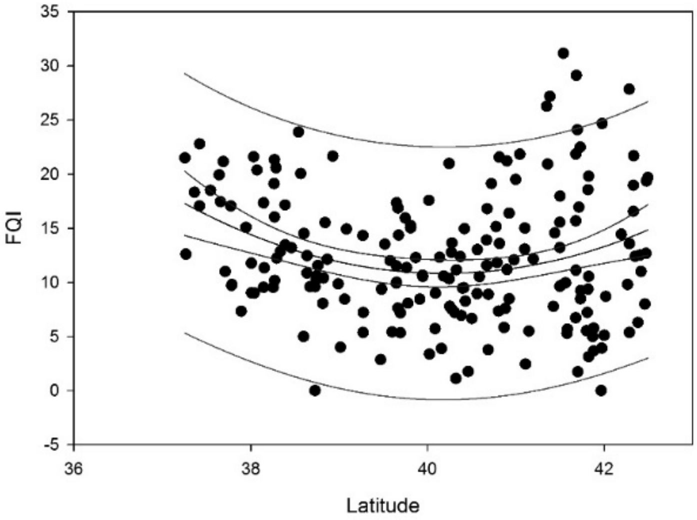
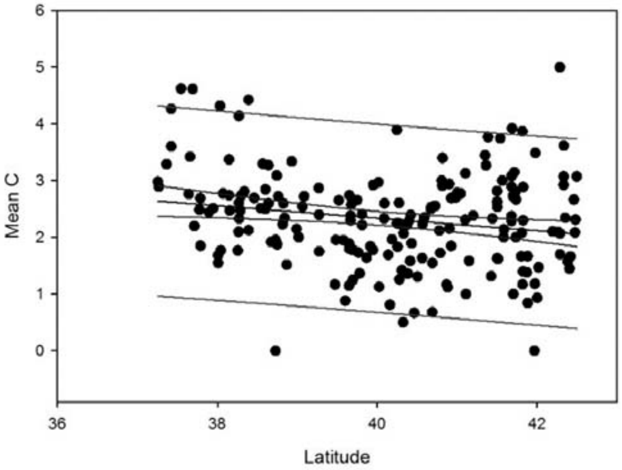


Figure 2.8. Scatter plot of simple (top) and quadratic regression (bottom) between site FQI and latitude (°N) for emergent wetlands. Regression line (center line), is bounded by 95% confidence interval, and the outermost 95% prediction interval. Simple regression was not significant ($p = 0.36$, $r^2_{adj} = 0.00$; see Table 2.1). The quadratic relationship was highly significant, ($F_{3,186} = 6.48$ and $p = 0.002$), with a best fitting quadratic polynomial equation: $Y = 1225 + -60.5x + 0.75x^2$; $r^2_{adj} = 0.06$.



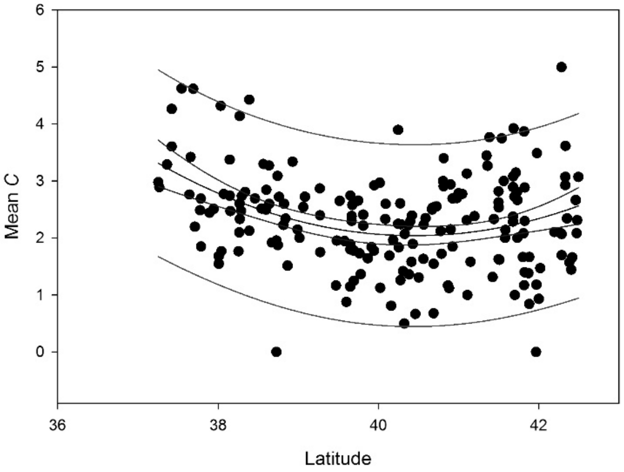


Figure 2.9. Scatter plot of simple (top) and quadratic regression (bottom) between site Mean *C* and latitude (°N) for emergent wetlands. Regression line (center line), is bounded by 95% confidence interval, and the outermost 95% prediction interval. Simple regression was significant ($p = 0.03$, $r^2_{adj} = 0.03$; see Table 2.1). The quadratic relationship was highly significant, ($F_{3,186} = 12.7$ and $p < 0.001$), with a best fitting quadratic polynomial equation: $Y = 205.8 + -10.1(x) + 0.12(x^2)$; $r^2_{adj} = 0.11$.

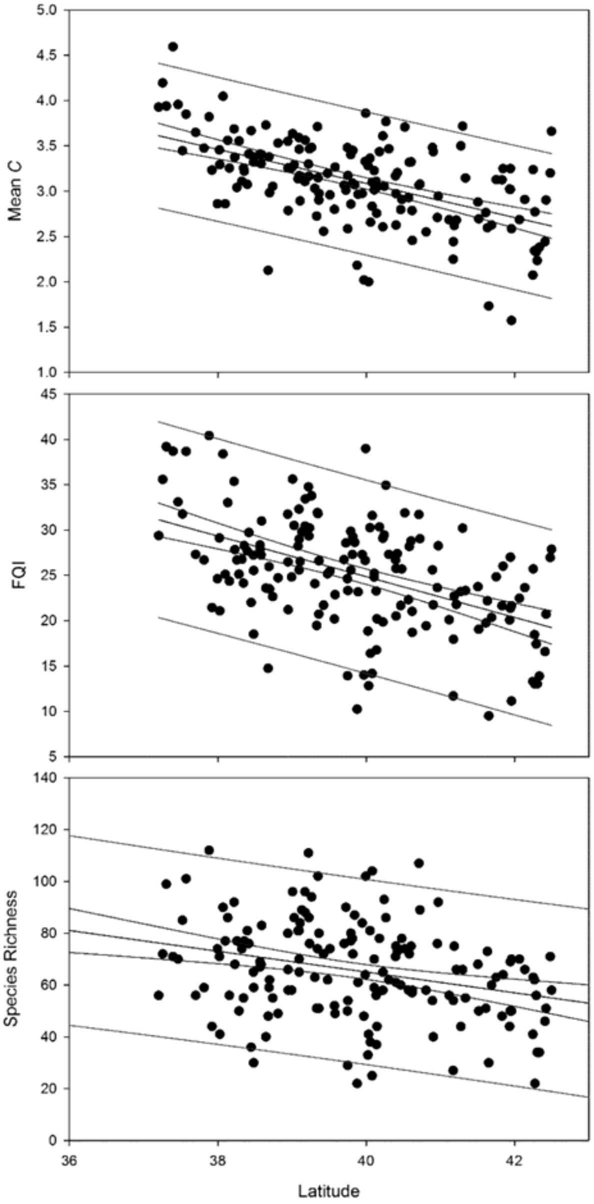


Figure 2.10. Scatter plot and regression line of the relationship between site richness, FQI, and Mean *C* and latitude (°N) in forests. Regression line (center line), is bounded by 95% confidence interval, and the outermost 95% prediction interval. See Table 2.1 for regression statistics.

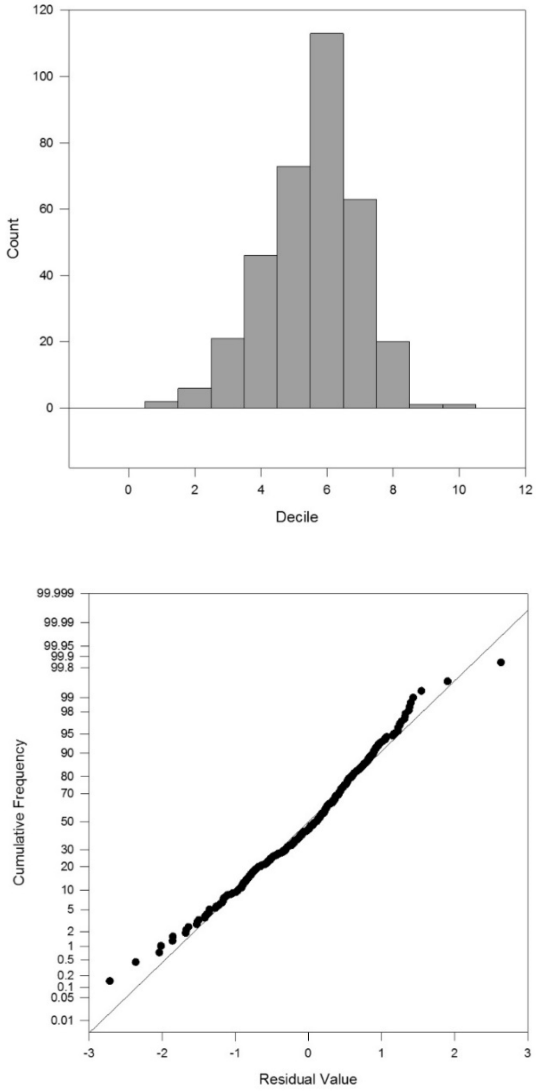


Figure 2.11. Histogram of residuals (top) and normal probability plot of residuals (bottom) from ANOVA comparison among site Mean *C* values from ecoregions (natural divisions) (see Figure 2.4).

Chapter 3: Successional trends in Floristic Quality[1]

Summary

1. Simple, conservation-relevant, plant community measures are sought by resource managers. In this context, the use of Floristic Quality Assessment (FQA) has increased exponentially over the past 30 years. FQA measures a habitat's floristic quality and conservation value by summarizing the relative anthropogenic disturbance tolerances of its plant species (i.e. their conservatism). However, despite their widespread use in research, restoration, and conservation work, the behaviour of FQA values in communities during succession is not understood.

2. We analyzed FQA values in 10 old-fields over 50 years of unaltered succession. We determined whether floristic quality followed a predictable increasing successional trend, assessing four specific predictions: 1) FQA values will follow an asymptotically increasing, rather than peaked or linearly increasing trajectory, 2) field initiation treatments (abandoned as hayfield or cropfield) will not lead to long-term differences in FQA values, 3) trajectories will be consistent regardless of the particular species composition of fields, 4) trajectories will be robust to common variations in FQA metric formulations (non-native species, varied spatial scale).

3. In all cases, a negative exponential rise to an asymptote best described FQA value trajectories over time. Field abandonment treatments did not affect FQA value trajectories. Furthermore, trends were consistent among fields despite differences in species

---

[1] This chapter appeared in its entirety in the Journal of Applied Ecology and it is formatted accordingly. The citation is: Spyreas, G., Meiners, S.J., Matthews, J.W., Molano-Flores, B. (2012) *Successional trends in Floristic Quality* (49) 339-348. This article is reprinted with the permission of the publisher and is available from http://onlinelibrary.wiley.com/, and using DOI: 10.1111/j.1365-2664.2011.02100.x

composition among fields. Overall, the results suggest a predictable, deterministic path for FQA values over early- to mid-successional timeframes studied.

4. *Synthesis and applications*. Understanding the temporal behaviour(s) of floristic quality is necessary for setting realistic restoration goals, evaluating habitat recovery, and for adapting management to achieve high conservation value natural areas. By illustrating the temporal consistency of floristic quality metrics during succession this paper demonstrates the robustness of FQA for such uses. The FQA value trajectory described here also establishes a background trend model for expected values in recovering habitats, which will allow for the assessment of an individual habitat's progression relative to the background trend. Such comparisons *en masse* will highlight the constraints of greatest importance to community-level floristic quality restoration. For example, FQA values in this study were ultimately limited by conservative understory plant re-establishment from adjacent old-growth forest. As this is not unlike species recovery patterns observed in other habitats, it suggests that restoration practitioners would do well to focus on conservative species.

Key Words

Anthropogenic disturbance, Conservation value, Deterministic vs. stochastic succession, Floristic Quality Assessment (FQA), Floristic Quality Index (FQI), Invasion impacts, Mean *C*, Remnant flora, Restoration monitoring, Successional trajectory

Introduction

Successional trends in plant communities and habitat restorations are commonly tracked, studied, and compared using simple measures such as diversity, structure, or biomass. However, these fail to capture the properties most immediately relevant for conservation– species identity and community composition (Filippi-Codaccioni et al. 2010). Thus, the means to compare plant assemblages with regards to their levels of endemism, rarity, regional uniqueness, taxonomic distinctness, and specialization are needed (Izco 1998, Ricotta 2004, Devictor et al. 2008, Chapman et al. 2009). However, quantifying such properties in ways that allow for easy comparisons among sites and over time has proven difficult.

It is in this context that the use of Floristic Quality Assessment (FQA) has increased exponentially over the past 30 years (e.g., North America; LaPaix et al. 2009, Europe; Bonanno and Giudice 2010). FQA utilizes "conservatism scores" assigned *a priori* to each plant species in region. A species score is based on its sensitivity to anthropogenic disturbance and its likelihood of being found in high quality remnant natural areas (Taft et al. 1997). Simple univariate summaries can then be used to characterize an area's floristic quality. Thus, an areas' "floristic quality" refers to the degree to which its plant assemblage is intact, resembling that of a remnant, native habitat. This is dependent on how much anthropogenic degradation the area has accrued and how many of its sensitive conservative species remain. Floristic quality metrics have been shown to effectively measure anthropogenic disturbance and site conservation value (Cohen et al. 2004, Miller and Wardrop 2006, Mack 2007, Mack et al. 2008). As only a plant species list is required, the ease of use and novel ecological information captured by FQA has spurred its increasing use in choosing natural areas for acquisition or legal protection. Land managers and researchers commonly also use FQA to determine the effectiveness of management techniques over time (e.g., Brudvig et al. 2007, Foster et al. 2007). In the United States, legal mandates for habitat monitoring and assessment often require FQA based criteria (Matthews and Endress 2008, U.S.E.P.A. 2010). Finally, FQA is increasingly used in basic ecological and conservation research (e.g., Panzer and Schwartz 1998, Spyreas and Matthews 2006, McNicoll and Augspurger 2010).

A key assumption to using FQA is that changes in metric values at a site are orderly and predictable over time. Insufficient understanding of the temporal dynamics of conservation metrics can lead to their misuse (Niemi and McDonald 2004). For example, high plant species richness is often considered indicative of less-disturbed, high conservation value habitats, but this generalization is unwarranted given the inherently non-monotonic trend in richness over time (Fleishman et al. 2006). Likewise, high floristic quality values are commonly equated with “mature”, “late”, “advanced”, “climax”, or “stable” successional states (e.g., Swink and Wilhelm 1994a, Middleton and Bever 2012), implying that FQA values increase in accordance with successional advance over time. This is not an unreasonable assumption given that rare, specialist, or disturbance-sensitive species are often prevalent in or restricted to the oldest or least disturbed habitats (Peterken and Game 1984, Honnay et al. 1998, Kindscher and Tieszen 1998, Honnay et al. 1999). However, if floristic quality values do not follow simple, predictable increases during succession as is assumed, their interpretation and use may be confounded.

Studies examining temporal changes in floristic quality values after anthropogenic disturbances have not shown consistent results. Time since logging disturbance in mature forests has been shown to correlate with higher floristic quality values (Francis et al. 2000, Wallace 2001). Chronosequence comparisons typically find older restorations to have higher floristic quality values (Mushet et al. 2002, Balcombe et al. 2005), while studies tracking individual sites often show unexpected deviations from monotonic increases over time (Spieles et al. 2006, McIndoe et al. 2008, Matthews et al. 2009b, Middleton et al. 2010). Decreasing floristic quality values in these instances have been concomitant with observations of non-native species invasion, suggesting that non-native species may dictate floristic quality values. However, these studies have only observed early successional (<20 years) restorations and the long-term relationship between invasion, succession, and floristic quality is unexamined. Community invasions which *persist over time* could suspend succession by native plants (Flory and Clay 2010), and/or lead to novel anthropogenic communities (Hobbs et al. 2006), thereby dampening native floristic quality values. Were non-native invasions to prove persistent, their negative effects on native floristic quality would be substantial and widespread (Spyreas et al. 2010). Alternately, invasion effects on floristic quality may be fleeting and largely limited to early successional stages if non-native species do not persist.

Even without non-native invasions or other obvious catalysts, developing plant communities can take unpredictable paths towards unexpected states (Hobbs and Norton 1996). Paths towards alternate community types may lead otherwise similar sites to become dissimilar in species compositions over time. However, if floristic quality metrics only measure accrued anthropogenic degradation and the time since disturbances, then the stochastic successional processes that produce differing species compositions should not lead to differences in floristic quality values. Furthermore, the trajectory of floristic quality values over time should not vary among sites that differ in species composition, if the sites have shared anthropogenic disturbance legacies. The temporal predictability of FQA values have not been studied in this way because the restoration sites compared thus far have differed in their anthropogenic disturbances.

We analyzed the temporal dynamics of floristic quality values in 10 old-fields over 50-years of unmanipulated succession after abandonment. If floristic quality is inexorably linked to time since anthropogenic disturbance and advancing successional state as is assumed, then floristic quality values in these fields will follow a predictable, increasing trajectory during succession. We address four specific predictions:

***1)** We predict that an asymptotically increasing trajectory will be a better descriptor of temporal trends in floristic quality values than either a linear or a peaked model.* Previous studies have shown that FQA values in the initial years of wetland restoration commonly exhibit an asymptotically increasing trajectory (Matthews et al. 2009b). Alternatively, a peaked trajectory to floristic quality values could arise if FQA values follow species richness over early- to mid-successional timeframes (Anderson 2007), or if fields become increasingly invaded by non-native species (Matthews et al. 2009b). A linearly increasing trajectory could reflect a strong link between floristic quality values and advancing successional states; where fields would consistently accumulate floristic quality as succession proceeded, and they would not slow or reach an asymptote in values, until rates of species turnover slowed and/or when fields reached successional equilibrium (i.e., as old-growth forest in the present case).

2) *We predict that field condition at abandonment (row crop vs. hayfield) will not have long-term effects on floristic quality values.* Thus, even if there are initial differences in floristic quality values associated with abandonment treatments, values will quickly converge on a common trajectory as time since disturbance (i.e., age) becomes the primary floristic quality determinant.

3) *We predict that floristic quality values will exhibit a consistently predictable trajectory regardless of differences in the particular species composition of individual fields*. Variation or divergence in FQA values corresponding with variation or divergence in field species compositions would suggest strong controls on FQA values beyond the time since site disturbance (e.g., stochastic successional phenomena) that could limit their utility.

4) *While several variants in metric formulations have been proposed for FQA, we predict that the asymptotic FQA trend model will be robust to differences in metric calculations, including those that vary in their spatial sampling scale and those that exclude non-native species*.

Materials and methods

The study used data from the Buell-Small Succession Study (BSS) fields, located within the piedmont region of New Jersey, USA (40° 30’ N, 74° 34’ W; http://www.ecostudies.org/bss). The BSS fields were farmed from 1701 to 1958-66, at which time they were abandoned from agriculture and allowed to re-vegetate without management or manipulation. Fields were abandoned as pairs in alternate years from 1958 to 1966. At abandonment this parcel was not seen as having been “farmed out”, although the site’s soils are characterized as naturally droughty and not very fertile. Since abandonment, the vegetation has been monitored with 48 permanently marked 0.5 × 2.0 m plots within each of 10 fields, from which percentage cover of all species present in plots has been annually or bi-annually recorded in mid to late July. Plots are arranged in a regular pattern that varies slightly with the shape of the field. Most fields abut a nearby old-growth forest preserve. Data collection occurred every year since release, until 1979, when sampling switched to alternate years. The fields also differed in their season of abandonment (autumn or spring), final crop (hayfield or row crops) and soil treatment (ploughed or intact vegetation). ‘Season of abandonment’ and ‘soil treatment’ have been found less important than ‘final crop’ in their effect on succession in the fields (Meiners et al. 2002). Therefore, only the ‘final crop’ treatment was considered in our study.

Floristic quality metrics are composed of Coefficients of Conservatism (*C*) previously assigned to New Jersey’s flora (Allen et al.). Where species sampled in BSS plots were not found in this database, *C* scores were taken from the nearest available state or as the average of the two nearest (e.g., West Virginia, Pennsylvania). Scores range from zero (tolerant of anthropogenic disturbance, no fidelity to remnant habitats) to ten (conservative species, intolerant of human stressors, exclusive to remnant habitats) (Taft et al. 1997). All non-native species are assigned zeroes.

Analysis

To determine which trajectory would best describe trends in floristic quality values over time, we used nonlinear least squares regression, using a Gauss-Newton algorithm in SYSTAT 11 to describe floristic quality values over time using three models for comparison (Engelman 2005). These models were chosen based on previously demonstrated success at characterizing successional dynamics (Zedler and Callaway 1999, Morgan and Short 2002, Gutrich and Hitzhusen 2004, Anderson 2007, Matthews et al. 2009b). The first model assumed that the value of an FQA metric (*Y*) increased linearly over time (*t*):

$$Y(t) = Y_0 + bt \qquad \text{eqn 1}$$

The second model assumed the value of a metric (*Y*) increased to an asymptote, a trend that is well described by the negative exponential function:

$$Y(t) = Y_0 + a(1 - e^{-bt}) \qquad \text{eqn 2}$$

where *t* is site age in years, *a* represents the asymptotic maximum, *b* is a slope parameter, and Y_0 is a y-intercept. Alternatively, values could initially increase to a peak and then decline. Such a trajectory is well described by a double exponential function:

$$Y(t) = Y_0 + a(e^{-ct} - e^{-bt}) \qquad \text{eqn 3}$$

Note that equation (8) reduces to equation (7) if the additional slope parameter *c* equals zero (i.e., there is no decline from the peak). Support for competing regression models was compared using Akaike's information criterion, corrected for small sample sizes (AIC*c*). We ran analyses using the age of the fields or the year of the sample (i.e., X-axis as field age or calendar year), but these produced similar results so we present data from field ages. The number of fields with data available for analysis varied at any given age (see vertical bars Figure 3.1) for two reasons: first, some fields did not have data for the oldest age classes because fields were abandoned in different years (final field ages ranged from 42-50), and second, because of the bi-annual sampling cycle in last half of the study.

We compared effects of field abandonment treatments on floristic quality using ANOVA to compare values in the first year and at the final age that had data for all ten fields. Comparisons at the final age used either 43 or 44 year old fields because of the bi-annual sample scheme. Both treatments had equal representation by 43 and 44 year old fields, the last age that all fields had reached.

Sorensen's distance values were used to represent differences in species composition among fields, as a means of addressing our prediction that floristic quality values will exhibit a consistently predictable trajectory regardless of differences in their species composition. Specifically, we used field level species presence-absences to calculate all pair-wise Sorensen's distances among fields for a given year. During transition ages when some fields were being sampled every year and others were already on alternate year sampling schedules, we used all the composition data available, but kept the sample size constant to calculate standard error using the same number of independent comparisons per year (45). Similarly, values only extend to an age of 46 to maintain full sample size for comparisons.

Because various formulations have been proposed for calculating floristic quality metrics (Ervin et al. 2006, Miller and Wardrop 2006), we examined the robustness of floristic quality-time models under different scales of species aggregation, and where non-native species were included or excluded from calculations. The first metric compared was Mean C ($\overline{C}$):

$$\overline{C} = \sum C /(S) \qquad \text{eqn 4}$$

where C is the Coefficient of Conservatism values of plant species, and S is the number plant species. Native Mean ($\overline{C}_n$) only considered native species:

$$\overline{C}_n = \sum C_n /(N) \qquad \text{eqn 5}$$

where C_n is the Coefficient of Conservatism values of native plant species, and N is the number of native plant species. The Floristic Quality Index (FQI), Floristic Quality Assessment Index (FQAI), and Native Floristic Quality Index (FQI_n), were calculated as follows:

$$\text{FQI} = \overline{C} * (\sqrt{S}) \qquad \text{eqn 6}$$

$$\text{FQAI} = \overline{C} * (\sqrt{N}) \qquad \text{eqn 7}$$

$$\mathrm{FQI_n} = \overline{C}_n * (\sqrt{N}) \qquad \text{eqn 8}$$

With respect to scale, Mean *C* values were calculated in the following ways. First, ‘site’ level values for a given age were calculated from the species list generated from all species encountered in all 480 plots. Second, ‘field’ level values were calculated from the accumulated species in the 48 plots in a given field. Finally, ‘average of plots’ values were calculated as the value within plots, averaged across all 48 plots in a field. This third value using plot level averaging has the effect of emphasizing frequently occurring species. Its calculation was intended to examine suggestions that using plot level averages may give a more realistic assessment of the floristic quality of a field by dampening contributions from outlier, rare, or ephemeral species (McIndoe et al. 2008). This has the same effect as weighting values by their frequency in a community, which has also been suggested for FQA’s use (e.g., Francis et al. 2000, Cohen et al. 2004). FQI values could not be compared in instances where sample effort and richness-area effects would bias comparisons (e.g., across years at the site level).

Results

There was no eventual decline in floristic quality (Figures 3.1 & 3.7), effectively reducing Equation 8 (peaked function) to Equation 7 (asymptotic) (Table 3.1). Based on this information and AICc, we selected the asymptotic as the more parsimonious model (Table 3.1). Visual examination of the asymptotic function suggests that Mean *C* and FQI values are near their maxima 50 years after field abandonment (Figure 3.7). The asymptotic trend was consistent whether or not the metrics included non-native species in their calculation (Table 3.1), although values without non-native species were higher (Figure 3.2). An asymptotic curve was also the best predictor of Mean *C* across the different scales that species were sampled/aggregated (Figure 3.3). Overall species richness in fields declined after a maximum value approximately 35 years after field abandonment (Figure 3.4). Although non-native species richness declined following abandonment (Figure 3.4), its trajectory did not vary inversely with FQA values. Non-native species dominance (percentage cover) over time did appear to vary inversely with FQI values. However, the trajectory of non-native dominance did not mirror that of Mean *C* values over the last ~25 years of the study period.

Abandonment conditions had neither initial nor long-term effects on floristic quality (Mean *C*: Age 1, *t*=48; d.f.=8; *P*=0.65; Age 43-44, *t*=0.33; d.f.=8; *P*=0.75; FQI, Age 1, *t*=0.49; d.f.=8; *P*=0.64; Age 43-44, *t*=0.89; d.f =8; *P*=0.4). The only apparent difference in the trajectory of values between abandonment treatments was a more rapid initial rise in row crop fields, approximately between ages 4-8, after which treatment values quickly converged and showed similar trajectories (Figure 3.5; FQI displayed a qualitatively similar pattern and is not shown). Fields varied over time in the number of species they shared, although they generally converged upon an intermediate level of dissimilarity in species composition (Figure 3.6). Variation in shared species among fields over time contrasts with variation in FQA values among fields (Figure 3.1), which were rather consistent except for a spike in variation at the end of the study period, which was an artefact of the reduction in sample size.

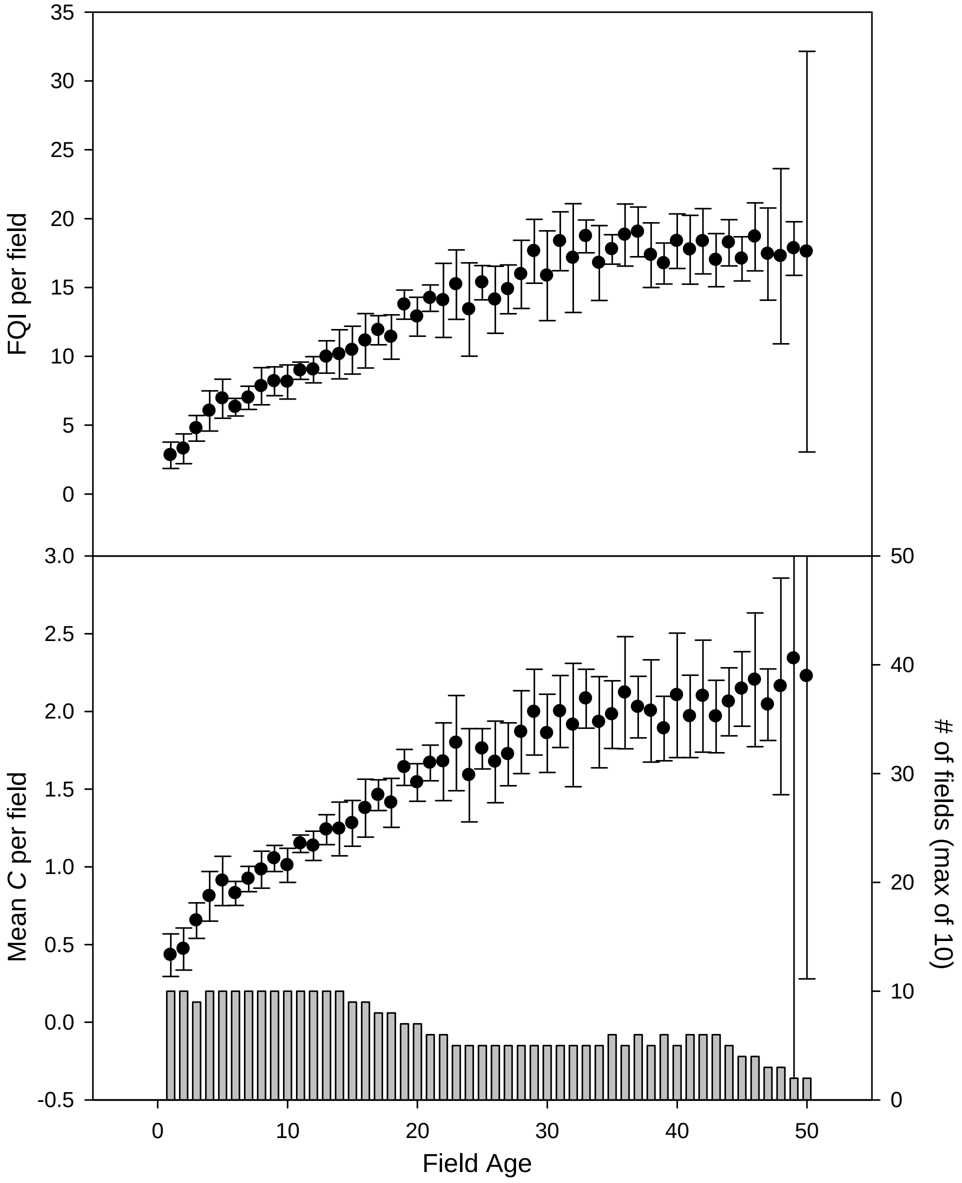


Figure 3.1. Trends in floristic quality measures in BSS fields over time (± 95% C.I.). Sample size for any given age in both graphs is indicated with vertical bars in the lower panel (i.e., right vertical axis), for this and all figures following. Non-native species are included in metric calculations.

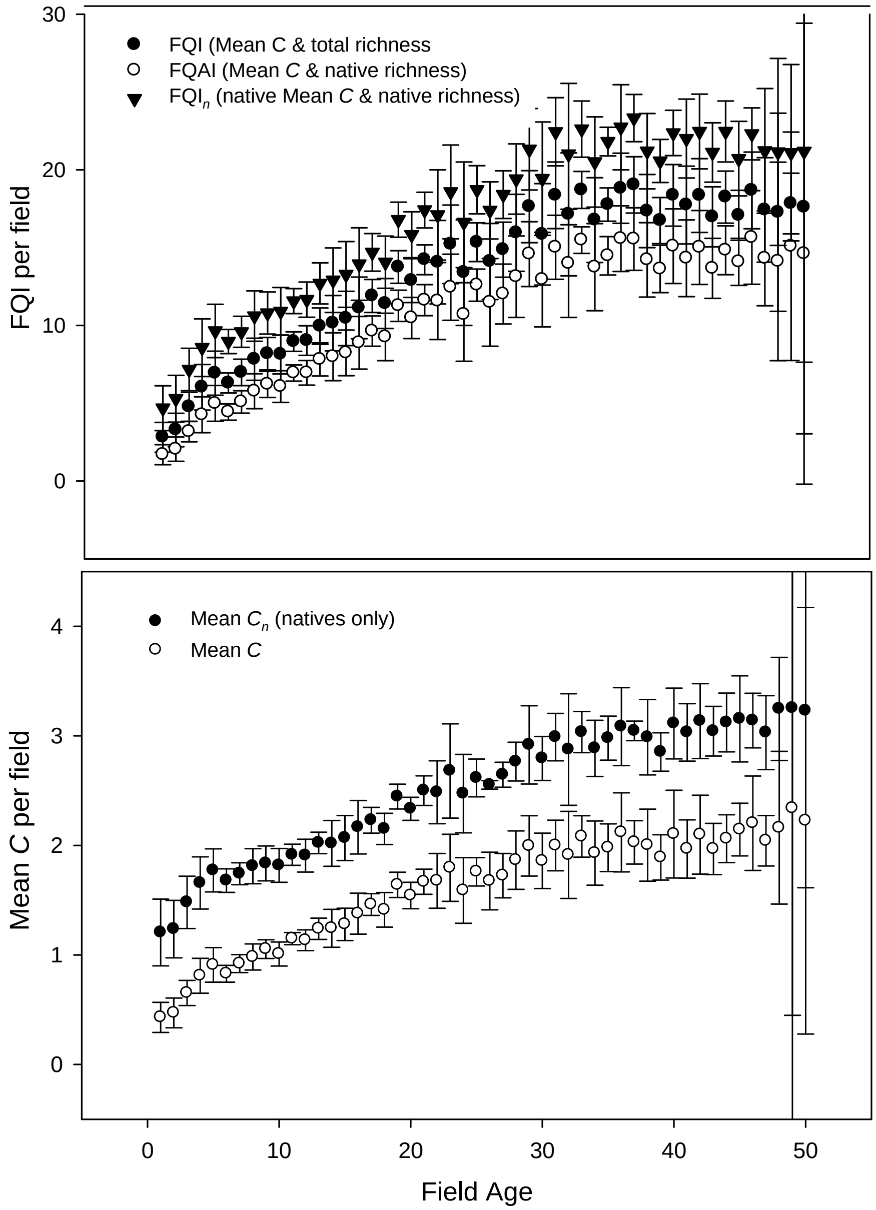


Figure 3.2. Trends in FQA measures in BSS fields over time using calculations that include or exclude non-native species (± 95% C.I.).

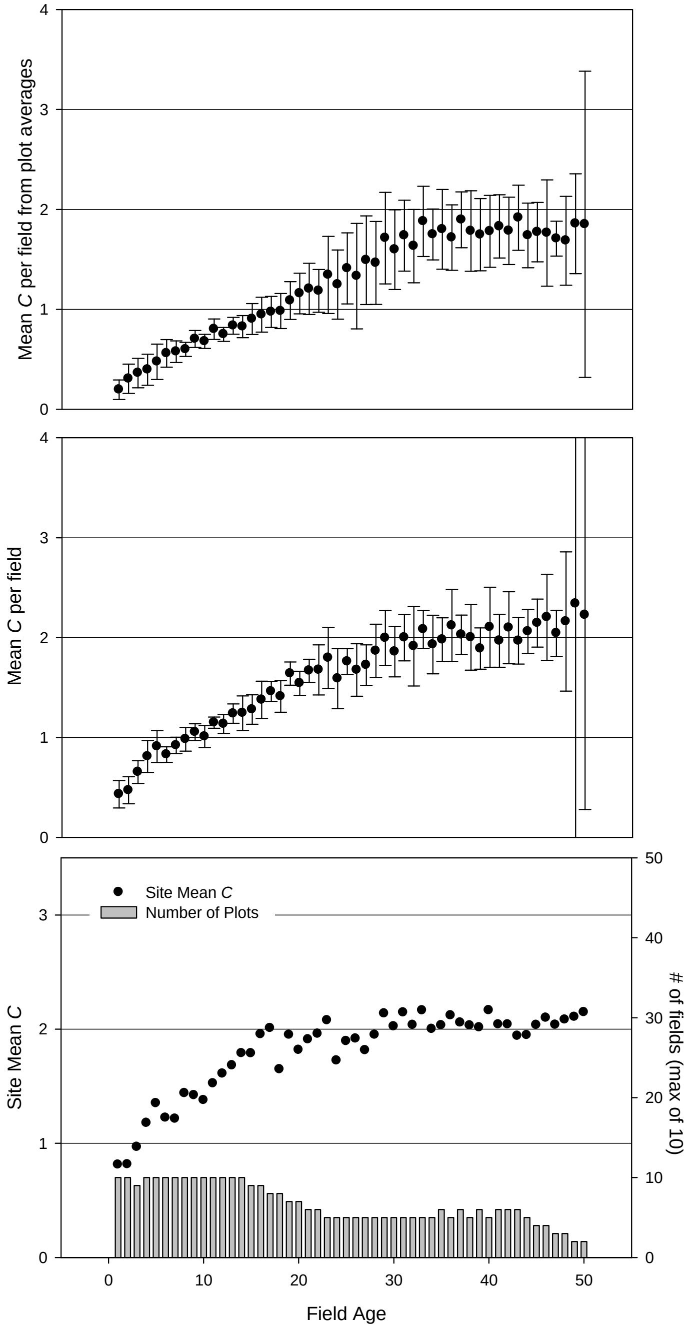


Figure 3.3. Mean C calculated across sampling spatial scales (the site level flora, field level floras, and average of plots per field) (± 95% C.I.). Non-native species are included in metric calculations.

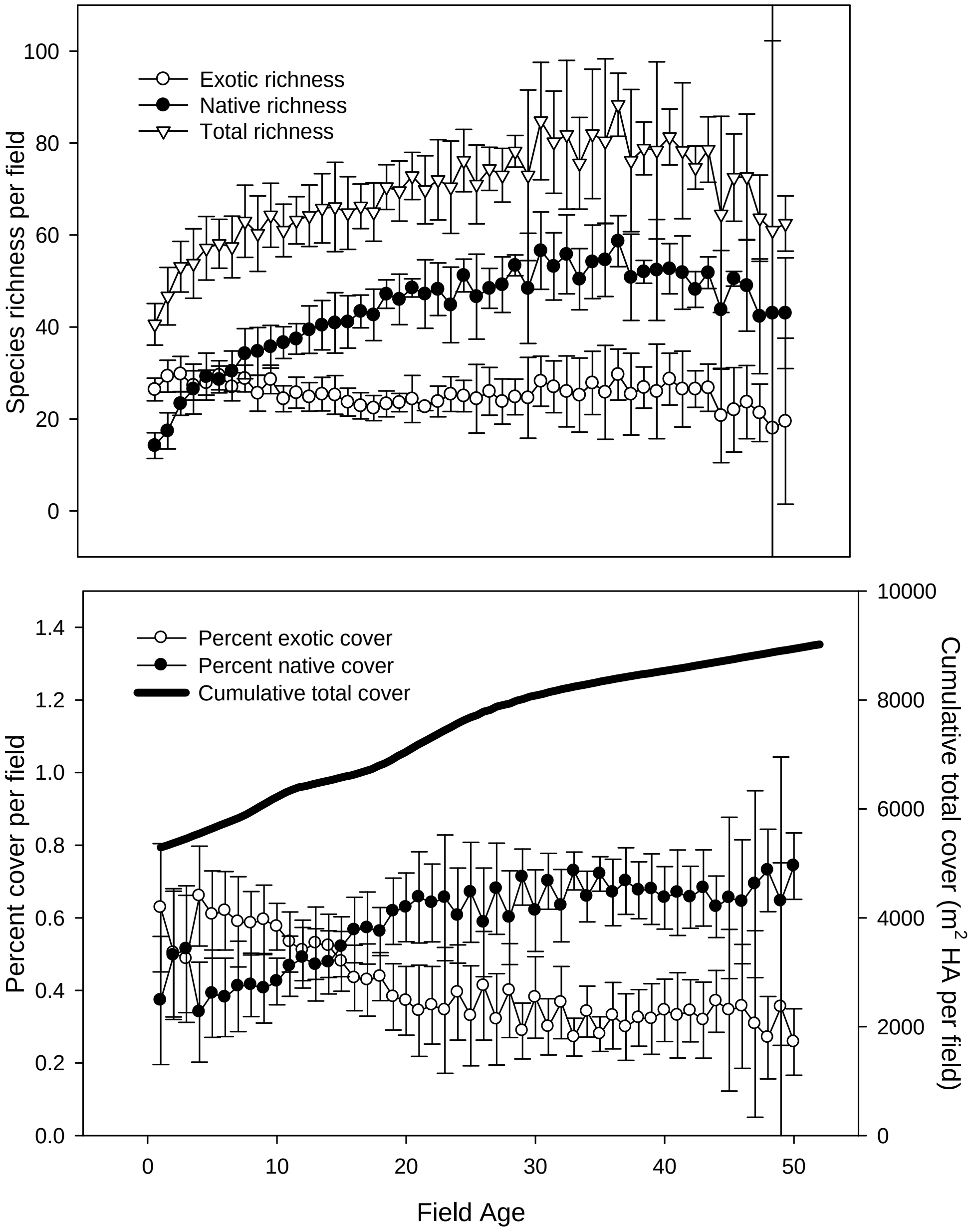


Figure 3.4. Trends in non-native and native species in BSS fields over time (± 95% C.I.). Percentages are relative contributions to total cumulative cover.

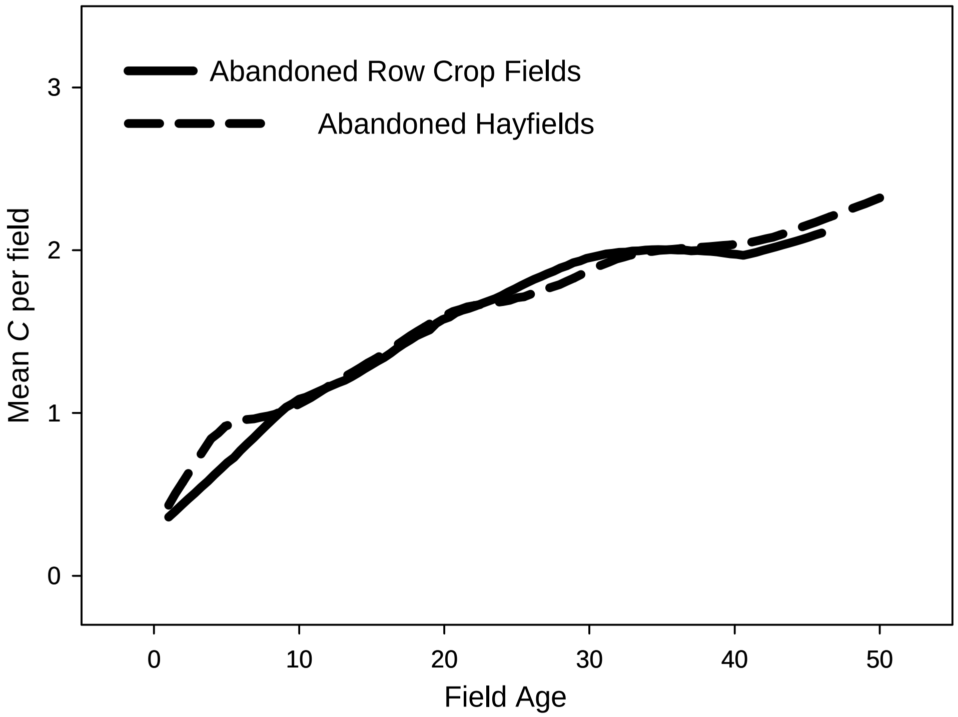


Figure 3.5. Trends in Mean C values for field abandonment treatments. Non-native species are included in metric calculations. FQI trends were qualitatively similar and are not shown.

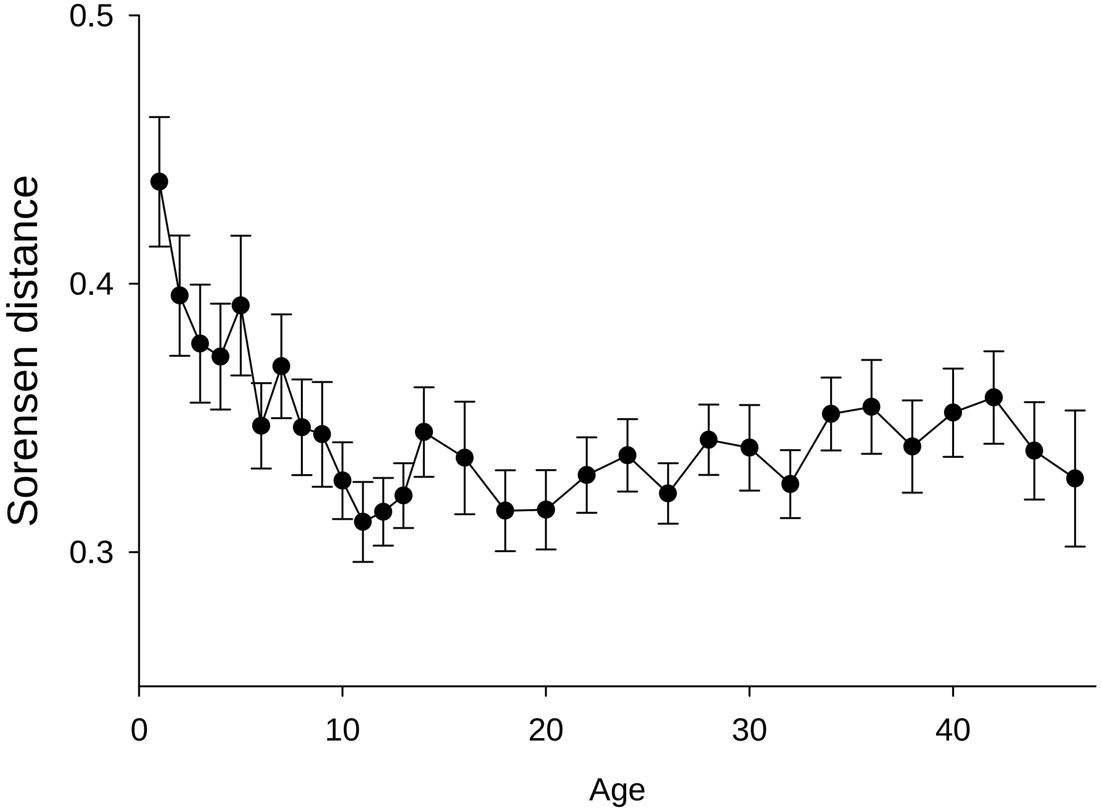


Figure 3.6. Dissimilarity in species composition among fields based on Sørensen's distance. Data plotted are average compositional distances among all fields at the same age (± 95% C.I.). Analyses switch to alternate years past age 15 reflecting the change in sampling periodicity.

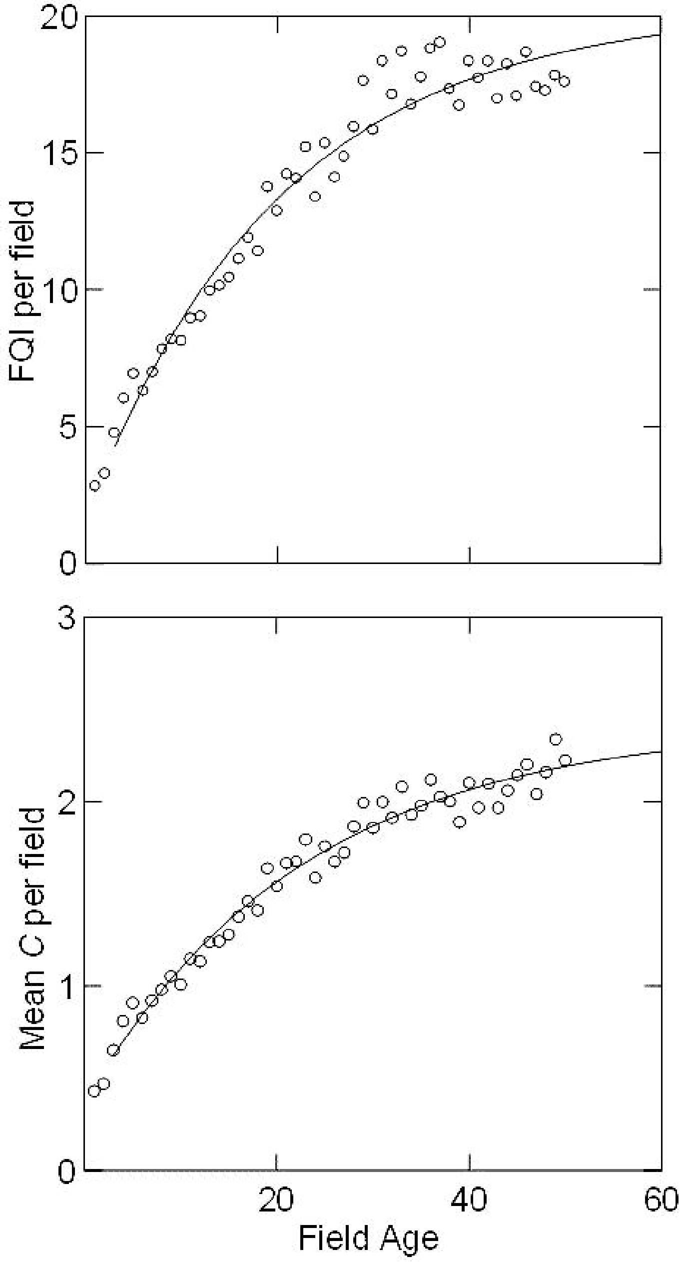


Figure 3.7. Negative exponential model fit line for FQA measures in BSS fields over time. Circles represent average of field scores as in Figure 3.1. Model fit lines are extrapolated beyond data to suggest an asymptote. Non-native species are included in metric calculations.

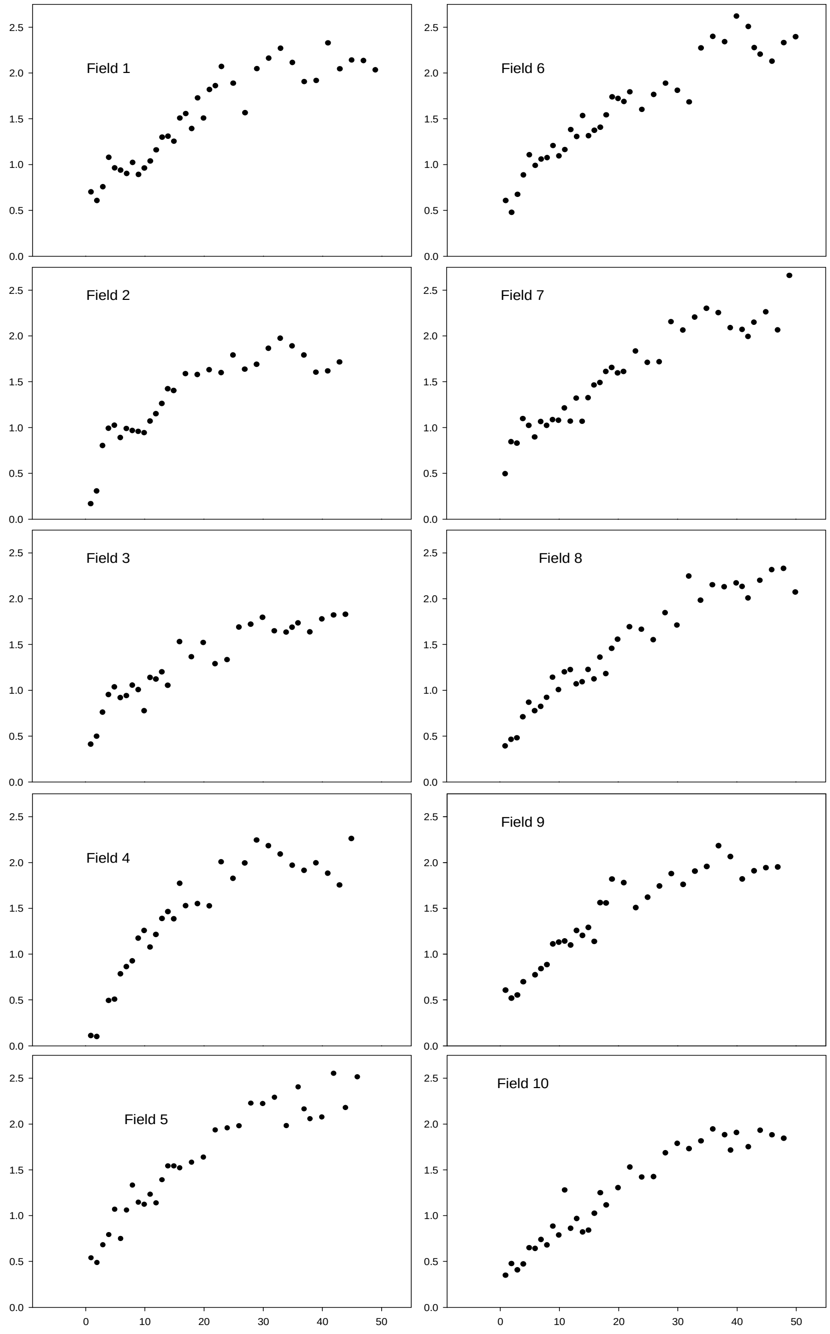


Figure 3.8. Trends in Mean C values over time in individual BSS fields. Non-native species are included in metric calculations. Y-axes are field Mean *C* value and X-axes are field age in years.

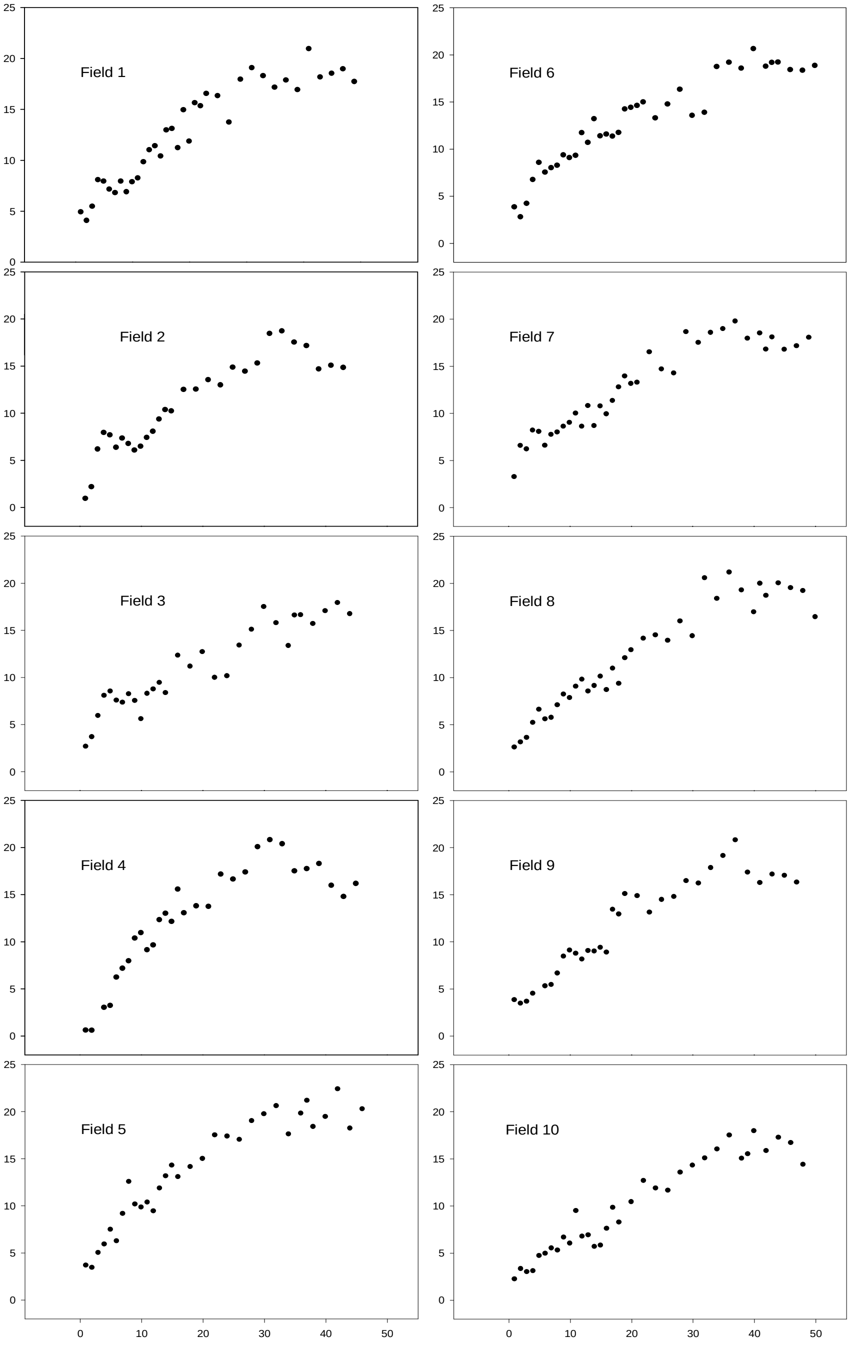


Figure 3.9. Trends in FQI values over time in individual BSS fields. Non-native species are included in metric calculations. Y-axes are field FQI value and X-axes are field age in years.

Table 3.1. Model comparisons and estimated parameters (Y0, a, b, and c) for floristic quality measures. Sample size is 50 in all cases. Results were qualitatively similar for Mean C values calculated at different scales and are not presented.

Model	**K**	**AIC$_c$**	**ΔAIC**	**likelihood**	**weight**	Y_0	R^2	***a***	***b***	***c***
Mean C										
Negative exponential	3	-103.9	0	1	0.77	0.375	0.975	2.048	-0.044	-
Peaked exponential	4	-101.5	2.367	0.306	0.23	0.375	0.975	2.048	-0.044	0
Linear	2	-76.89	26.96	0.000	0.00	0.767	0.905	-	0.033	-
Mean C$_n$										
Negative exponential	3	-12.79	0	1	0.77	1.192	0.976	2.588	-0.032	-
Peaked exponential	4	-10.42	2.367	0.306	0.23	1.192	0.976	2.588	-0.032	0
Linear	2	6.285	19.07	0.000	0.00	1.512	0.935	-	0.039	-
FQI										
Negative exponential	3	87.24	0	1	0.77	1.725	0.962	18.57	-0.049	-
Peaked exponential	4	89.61	2.367	0.306	0.23	1.725	0.962	18.57	-0.049	0
Linear	2	112.5	25.27	0.000	0.00	5.902	0.865	-	0.300	-
FQAI										
Negative exponential	3	81.13	0	1	0.77	0.536	0.963	16.24	-0.050	-
Peaked exponential	4	83.49	2.367	0.306	0.23	0.536	0.963	16.24	-0.050	0
Linear	2	106.9	25.77	0.000	0.00	4.242	0.864	-	0.262	-
FQI$_n$										
Negative exponential	3	94.98	0	1	0.77	3.748	0.957	20.97	-0.047	-
Peaked exponential	4	97.34	2.367	0.306	0.23	3.748	0.957	20.97	-0.047	0
Linear	2	117.4	22.45	0.000	0.00	8.226	0.866	-	0.338	-

Discussion

The best model for floristic quality values over the first 50 years of succession was a negative exponential increase to an asymptote. This trajectory was consistent whether or not non-native species were included in calculations and it was robust to scales of vegetation sampling. Initial field condition had some early effects on floristic quality values, but trajectories quickly converged among fields and values did not vary between treatments over the long-term. The consistency of FQA value trends despite large temporal variation in species dissimilarity among fields suggests that values are dictated by deterministic successional processes over early to mid-successional stages.

The trajectory of floristic quality values over time

A few studies have reported community-level floristic quality values over time. Matthews et al (2009b) tracked 29 wetland restorations in Illinois for 5 to 14 years after their creation. Although floristic quality values were far more variable among sites and over time compared to those in our study, the majority of their sites were also best described by an asymptotic trajectory model. A similarly-shaped logarithmic trajectory best described FQI values in eight Ohio wetland restorations (Gutrich and Hitzhusen 2004), which on average reached an asymptote 8 years after their creation. Finally, values from an Indiana grassland restoration generally increased over 13 years (McIndoe et al. 2008), although the shape of the trajectory was too erratic to be defined.

While asymptotic trends are most often supported, there appear to be stark differences among studies and systems in the length of time until values plateau. Peaks within 5-10 years typify wetland restorations, whereas at least three decades were necessary in our study's upland fields. Comparatively rapid peaks to floristic quality in wetland restorations could have several causes. First, conservative species are planted in most of these restorations. This is compared to BSS fields which underwent natural colonization and showed gradually increasing trends. Second, relatively low dispersal limitation and high productivity in wetlands allows for rapid establishment by highly competitive taxa whose dominance then resists new colonizations (Chen et al. 2010). Finally, emergent wetlands could have earlier peaks because their terminal state as a herbaceous community lacks the woody and shade tolerant forest taxa accompanying the ongoing physiognomic change of BSS fields to forests.

Succession and floristic quality

The floristic quality trajectories of BSS fields were notable for their consistent shape (Figures 3.8 & 3.9) and variation over time (Figure 3.1). Additionally, there were no patterns in floristic quality values related to year of abandonment or spatial position at the site (data not shown). Therefore, while minor differences in slopes or asymptote values were apparent, no field FQA values took idiosyncratic or divergent paths suggesting that they were dictated by historical contingency or spatial stochasticity (Vaughn and Young 2010). Similar successional trends to floristic quality values may not seem surprising for fields sharing the same species pool and abandoned under similar abiotic conditions (soils, etc.), as this would likely lead to similar species assemblages in fields. However, species dissimilarity among BSS fields was actually quite variable over time, while FQA trends remained consistent. Thus, different species in different fields were producing the same floristic quality trends across the site. This is particularly surprising for a metric like FQI, the components of which, species richness and composition, are frequently erratic and unpredictable during succession (Matthews 1979, Christensen and Peet 1984). Furthermore, initial field conditions (hayfield vs. bare ground) are known to have differentially affected fields in other aspects for 30 years or more after abandonment (e.g., relative representation by annuals and forage grasses, native vs. exotic richness, Meiners et al. 2002), but floristic quality values between treatments followed nearly identical trend lines throughout. In total, these results suggest that floristic quality was dictated by deterministic processes over time, and that FQA measures behave predictably in unmanipulated habitats over early- and mid-successional timeframes.

This finding is also supported by comparing patterns of richness and floristic quality in plots versus fields. While floristic quality values had similarly increasing trajectories when calculated per plot, per field, or at the site level, species richness behaved differently at different scales. Richness (total and native) per field exhibited distinctly unimodal trends, whereas species richness per individual plot has remained very consistent in BSS plots over time (Meiners et al. 2002). Therefore, species of greater conservatism replaced less conservative species in plots, without a net change in species density per plot. However, the same increasing floristic quality trends were generated by different increasingly conservative species in different fields.

On the other hand, species life form was clearly related to successional trends in floristic quality values, especially for dominant plants. For example, the first group to dominate was comprised of weedy ephemeral taxa with low *C* values (e.g., *Ambrosia artemisiifolia* L. *C*=0, *Erigeron annuus* (L.) Pers. *C*=0), whose populations collapsed within 10 years (Meiners et al. 2008). The second group to ascend was comprised of slightly more conservative perennial herbaceous taxa (e.g., *Aster pilosus* Willd. *C*=1, *Solidago juncea* Aiton, *S. Canadensis* L., *S. gigantea* Aiton, *S. rugosa* Mill. *C*=2, *Apocynum cannabinum* L. *C*=2). The third group was made up of the trees, shrubs, and woody vines that dominated during later years of the study (e.g., *Acer rubrum* L. *C*=3, *Rubus allegheniensis* Porter *C*=3, *Cornus florida* L. *C*=5, *Vitis* spp. *C*= 4). They first increased floristic quality values as they came to dominate communities, and then they maintained values at their asymptotic levels as old-field herbs declined. However, despite the seeming coupling of life form with species conservatism levels during succession, life form and conservatism are not synonymous. Both highly conservative and non-conservative species are well represented among all life history, functional group, and species trait categories in regional floras. Further study of the yet untested relationship between life form and species conservatism certainly seems warranted.

A fourth group of species influencing temporal patterns in floristic quality values were non-native species, which generally decreased over time in BSS fields relative to natives. Non-natives *directly* decrease floristic quality values when included in metric calculations (Equations 1, 3, 4; Figure 3.2). However, because there were no differences in the shapes of trajectories for metrics that included or excluded non-natives, non-native presence or richness alone did not determine floristic quality value trajectories. Non-native species effects on floristic quality values can also occur as an *indirect* function of invader dominance by displacing native species with higher *C* values or by decreasing opportunities for them to establish. Even though several of the most invasive plants in North America (e.g., Rosa multiflora, Microstegium vimineum, Lonicera japonica, Alliaria petiolata, Lonicera maackii; Meiners et al. 2001, Gibson et al. 2002, Spyreas et al. 2004), are common in BSS fields, decreasing overall non-native dominance may have explained the asymptotic trajectory shape in these fields, rather than the peak-and-decline trajectory sometimes observed for FQA values over time. While some studies have not found a relationship between floristic quality and non-native species (Tulbure et al. 2007, Brewer 2010), our study does not dispute the majority of evidence which suggests considerable depressive effects from strong invasions (e.g., Spyreas et al. 2010). As non-native species and their impacts have been suggested as being comparatively minimal in mature forests (Von Holle et al. 2003, Meiners et al. 2008, Martin et al. 2009), it will be highly informative to follow continued maturation of BSS vegetation with respect to non-native invasions and their effects. Furthermore, because understories contain a disproportionate amount of the plant diversity in these forests, future study should consider invasion in different strata and their effects on floristic quality in different strata.

Even though BSS fields had become young forests by the end of the study, and despite their adjoining old-growth forest seed source, their understories show a glaring absence of conservative shade-tolerant native forest herbs. Conservative forest herbs were sporadically detected in plots throughout the study period (e.g., *Actea pachypoda* Elliott *C*=5*, Athyrium felix-feminina* (L.) Roth *C*=7*, Circaea lutetiana* L. *C*=6*, Monotropa* *uniflora* Small *C*=8*, Phryma leptostachya* L. *C*=8*, Podophyllum peltatum* L. *C*=6), but these were singular occurrences that did not persist. The potential for future sustained colonization by these taxa could initiate a second period of increasing floristic quality values in BSS fields. However, the notoriously slow migration and establishment by such species into mature forests, suggests that this will not occur for hundreds of years, even with adjacent propagule sources (Matlack 1994, Brunet and von Oheimb 1998, Singleton et al. 2001, Spyreas and Matthews 2006). Re-colonization rates by *conservative* species in other habitat types have not been directly studied, but long-term comparisons of site histories suggest that if passive recovery by remnant taxa occurs in non-forest habitats, it will be measured over centuries as well (Gibson and Brown 1991, Kirkman et al. 2004, Ejrnæs et al. 2008). For example, conservative species are notably absent from grassland restorations even with propagule sources that are directly adjacent (Kindscher and Tieszen 1998, Foster et al. 2007).

Implications for the use of FQA

It could be argued that the increases in floristic quality values demonstrated here provide evidence that “hands-off” approaches to restoration are likely to be successful given enough time; however, we reject this interpretation. Restorations are prone to failure from non-native species invasions (Matthews et al. 2009b). Furthermore, the maximum values in BSS fields (Mean $C \approx 2.25$, FQI ≈ 17), were still well below values in remnant habitats with intact floras (e.g., Mean C ≈ 5-6, FQI ≈ 45-55, Swink and Wilhelm 1994a), as the highly conservative species characterizing remnant habitats did not establish. Barring a few exceptional cases (e.g., in North America, Sperry 1994, Gardner 1995), even the oldest restoration projects show considerable deficiencies in their floristic quality. Therefore, restoration efforts would do well to focus on conservative species. In instances where restorations have achieved FQA value parity with remnants, they have received massive planting and management efforts over dozens of years (e.g., repeated over-seeding, hand planting of plugs, careful introduction of missing conservative species, meticulous monitoring, regular prescribed fire, invasive species control).

Three conclusions can be drawn from these results with respect to assumptions underlying FQA's use. First, by illustrating the consistency of floristic quality metrics during succession, we demonstrate the robustness of FQA for use across temporal gradients. Second, because these fields reached an asymptote in their FQA values even though they continue to undergo rapid successional turnover (data not shown), temporal changes in FQA values cannot be considered synonymous with succession or with the successional states of communities. Finally, while the relationship between floristic quality and time since anthropogenic disturbance may be consistent and predictable, it is not simple (i.e., it is non-linear). Therefore, FQA users must carefully consider background successional trends in floristic quality when using FQA metrics across temporal gradients or for habitats of different ages. For example, Tulbure et al. (2007) concluded that an increase by an invasive species did not decrease a community's floristic quality over time. However, the lack of an invasion effect may have been obscured by background increases in floristic quality that were likely occurring across the site, which was undergoing rapid succession after a recent disturbance. Similarly, controlling for ambient successional changes in floristic quality values in a study of deer browsing effects on the floras of young grassland restorations may have allowed for treatment differences to have been better discerned (Anderson et al. 2007).

While the asymptotic trajectory model we have described will require further testing for its general applicability in other habitat types, successional stages, regions, and landscape settings, we suggest it for use as a baseline expectation for predicting floristic quality values over early- to mid- successional timeframes. Deviations from this expected baseline trajectory could highlight relative successes or failures in recovery progress or management practices at sites. Comparative study of site trajectories and their deviations from the expected baseline *en masse* would reveal patterns in the relative importance of specific ecological constraints to the recovery of community level floristic quality.

Acknowledgements

This research was supported by NSF grant DEB-0424605 to Scott Meiners. Greg Spyreas is supported by IDNR funding of the Critical Trends Assessment Program (CTAP)

Conclusion

Floristic Quality Assessment is a popular and influence ecological measure. The properties of several of its metrics have not been established, nor have several of its assumption for use been tested. The studies in this thesis add to the understanding of FQA metrics in several ways. First, there was a finding of variation in Floristic Quality levels among habitat types. Second, there was a finding of some regional variation in FQA metric values. Finally, there was a finding that site-level FQA values showed consistent asymptotically increasing trends over mid-successional timeframes. With regards to use of FQA, these results suggest that some variation in values could be attributable to the natural ecological characteristics of regions or habitat-types (i.e., not all variation in values attributable to anthropogenic degradation/disturbance), and depending on the use, this non-anthropogenic variation may need to be accounted for. Specifically, site comparisons across very large latitudinal gradients, or of areas from certain specific habitat-types (e.g., upland versus floodplain forest), or ones where only very high quality, pristine habitats are being compared, are instances where FQA score comparisons should proceed cautiously. However, because variation in Floristic Quality across habitat-types and regions was either found to be very small, or it was probably a reflection of human degradation levels across sites, most general FQA use is supported as is. With respect to temporal patterns in scores, the consistency of FQA values over time show that when sites of different ages are compared with one another, an asymptotic trend should be expected. Future study on these topics will continue to refine the understanding of FQA metrics.

Appendix: A review of the study, concept, and application of Floristic Quality Assessment

Measures of Habitat Conservation Value

This papers goal is to examine the conception, theory, criticism, study, methods, use, misuse, and potential future directions in Floristic Quality Assessment. But, it is first placed alongside existing measures of habitat conservation value in a way that highlights their limitations, as opposed to the relative strengths of Floristic Quality Assessment that have precipitated its popularity. Conservation practitioners and ecologists are tasked with evaluating the naturalness, Biological Integrity, and conservation value of natural areas in a way that is as objective as possible. Vegetation-based measures are frequently employed for such purposes, for ecological studies, habitat assessments, and monitoring. But, the history of such measures has been one of limited effectiveness and utility (Smith and Theberge 1986, Mace 2005). This is because in order to be widely useful, such measures must not only be straightforward in their implementation, but they must be flexible and repeatable. Furthermore, they must be easily standardized, compared, and interpreted across different habitats and over space and time. Traditionally measures like plant biomass, productivity, diversity, or structure have been used to in this capacity because they fit most of these criteria (i.e., easily quantified, standardized). However, because they do not directly consider species identities or community composition, such measures do not assess conservation-relevant properties directly. This measurement problem has long been recognized; Diamond (1976) stated that “species must be weighted, not just counted,” because the critical issue for conservation interests is often not an area’s total number of species, but the area’s “species doomed to extinction.” Consequently, plant composition-based habitat measures have emerged that achieve greater conservation relevance. These often focus on recognizable ecological properties. Examples include measures of species endemism, rarity, regional uniqueness, taxonomic distinctness, or specialization within a community’s flora (Izco 1998, Ricotta 2004, Devictor et al. 2008, Chapman et al. 2009, Filippi-Codaccioni et al. 2010). Further, such measures are often able to be summarized into univariate values. Unfortunately, such measures often require sophisticated calculations with multiple steps in data manipulation, which limits their ready-adoption by conservation practitioners. Furthermore, many habitats of high natural quality or conservation value do not necessarily have an abundance of rare, endemic, or specialized plants.

An alternative to quantifying specific ecological characteristics is to focus on the overall completeness or intactness of a habitat's plant assemblage. In other words, how many of the plants that one would expect to find in a given plant community are actually there, and conversely how many can be assumed to have been eliminated by habitat degradation? Completeness or intactness can be simply measured as a percentage of similarity to a habitat's historic composition (e.g., conditions in North America before large-scale European settlement), or to a contemporary reference habitat (e.g., Scholes and Biggs 2005, Hawkins 2006). Measuring this single property– assemblage intactness– then encompasses, or serves as a proxy for many ecological characteristics and other conservation-based measures. This is because intact assemblages will contain rare, endemic, and specialized taxa, and it will also contain appropriate functional group representation and diversity. Furthermore, such an indicator is attractive because it uses an unambiguous benchmark (reference communities), it is conceptually simple, and it avoids focusing on what may be subjective or synthetic ecological properties.

Completeness measures are, however, typically limited in their readymade convenience and widespread utility because their data requirements are intensive. Making direct historic comparisons with natural areas' biotas works well where historic records are available. For example, fish assemblages at watershed scales have worked well (e.g., the "Fish Faunal Intactness" indicator, National Research Council 2000). Some historic plant records are available (e.g., some older North American records, Robinson et al. 1994, Drayton and Primack 1996, Bowers and Boutin 2008). But, they cannot be widely used for plants, because with the exception of trees, knowledge of historic floral compositions is limited by a paucity of phytosociological data on pre-settlement plant communities (e.g., in North America see, Gleason 1909, Noss 1985, Smith and Theberge 1986, Council 2000, Norris and Farrar 2001, Povilitis 2002, in Europe see, Ejrnæs et al. 2008).

Where historic records are inadequate, existing natural areas may be used as reference comparisons. However, matching a reference to its target has been a notoriously capricious process, and accurately determining the appropriateness of a reference dataset for comparison is an involved process that is often ignored (Hawkins 2006, Nielsen et al. 2007). Reference datasets are often mismatched in spatial scale, habitat type, or geographic location with the target. Alternatively, comparisons against contemporary communities may be useful. Large contemporary reference datasets that contain a spectrum of both degraded and high-quality communities are especially useful because they can allow sites to be placed along a spectrum (Matthews and Spyreas 2010). Such datasets are rarely available. Therefore, using completeness measures to assess natural areas is too laborious to be commonly practical. From an ideological perspective, using completeness-intactness measures may also over-emphasize matching idealized historic assemblages as the end goal (e.g., as an absolute restoration target)– which may deny the pervasiveness, inevitability, and potential value of novel modern species associations and combinations (Hobbs et al. 2006, Hobbs et al. 2009). Finally, because such reference comparisons weigh all species equally, they also ignore that certain species (e.g., weedy or common species) are of less conservation interest than others when matching or comparing targets to reference communities.

Introduction to Floristic Quality Assessment

Floristic Quality Assessment (FQA) has come to prominence as a conservation measure because it is more easily used than other measures described, but it also provides considerable information about the ecological characteristics and conservation value of habitats. FQA was originally developed in the late 1970's in the Chicago Region (U.S.A.), to identify protection-worthy lands with a simple, repeatable, quantitative method based on site floras (Wilhelm 1977, Swink and Wilhelm 1979). First developed for prairies and other open lands, FQA is now widely used to study, monitor, rank, and guide management of natural areas of all habitat types (see *History* below). FQA is based on each plant species' "Conservatism" value. Species values are uniquely assigned for each specific region. In the USA, species values are most often assigned to plants within an entire state (for the sake of simplicity, "state" is henceforth used to refer to any FQA region, but for other regional coverages, see Figure A.1). Species Conservatism values range from 0-10, increasing with a plant's exclusivity to remnant, anthropogenically undisturbed habitats in the region. This remnant fidelity is based on species assumed sensitivity to anthropogenic disturbances (Swink and Wilhelm 1994a).

The Conservatism values of a site's plants are used to create univariate summary metrics that quantify the area's Floristic Quality. Floristic Quality scores reflect the extent to which a plant community is composed of Conservative taxa, as opposed to non-Conservative taxa that are not exclusive to remnant, undegraded, natural areas. Sites with high FQA scores are considered biologically undegraded, because they have incurred less anthropogenic disturbance and have retained a greater richness of Conservative plants or a greater ratio of Conservative to non-Conservative species. This assumes communities to have a more intact floral assemblage, with a full complement of Conservative species from their remnant condition. Furthermore, the fewer remnant native habitats that remain in a landscape, the more likely it is that the site's Floristic Quality scores will be useful measures of its *conservation value*. For example, in a region solely composed of undegraded native habitats, even the most pristine site would be rather expendable. But, where few remnant habitats remain a site with a high FQA score will be valuable for its conservation value. This has been in terms of habitat irreplaceability: "Clearly, a tract of land occupied prevailingly by non-Conservative species not only cannot be viewed as natural area, but is quite replaceable. Our interest, then, in natural area identification and assessment should focus on the extent to which constellations of Conservative species are present." (Wilhelm and Ladd 1988)

In some ways, FQA is similar in concept to some widely used metrics in stream biotic indices (e.g., Hilsenhoff's Biotic Index), where a stream's human disturbance legacy is revealed by the presence or absence of sensitive fish or aquatic invertebrates that are intolerant of human stressors and pollution in the assemblage (Hilsenhoff 1987). Although they were created in North America at the same time as FQA, and they probably influenced its development, stream IBI's tend to be used to monitor pollution/degradation to waterways via the biota, while FQA users are more often interested in evaluating the conservation value of the biota itself. Several other related measures and terms share conceptual similarities with species Conservatism (see discussion in *Related concepts*).

The number of studies published studies using FQA has increased exponentially every decade since its creation (Spyreas unpublished), and geographic coverage for floras is quickly expanding within (Figure A.1), and outside of North America (Tu et al. 2009, Bonanno and Giudice 2010, Malik et al. 2012). Despite its growing popularity, FQA has been criticized as being imprecise, inconsistent, biased, subjective, tautological, untested, and unsubstantiated by ecological theory (e.g., Bowles and Jones 2006, Nielsen et al. 2007). Such criticisms have not deterred its use, and FQA will likely continue to gain influence and prominence for several reasons: 1) there remains a need for rapid, widely-applicable, univariate, vegetation metrics of its kind, 2) it requires only a plant species list and a simple calculation for use, 3) it captures information that may be unique among existing ecological metrics (i.e., Conservatism), and 4) it has proven to be a highly robust measure of anthropogenic disturbance and habitat degradation. But, despite more than 30 years of use, there has yet to be a comprehensive review of FQA. This paper examines the conception, theory, criticism, study, use, misuse, and potential future directions in Floristic Quality Assessment.

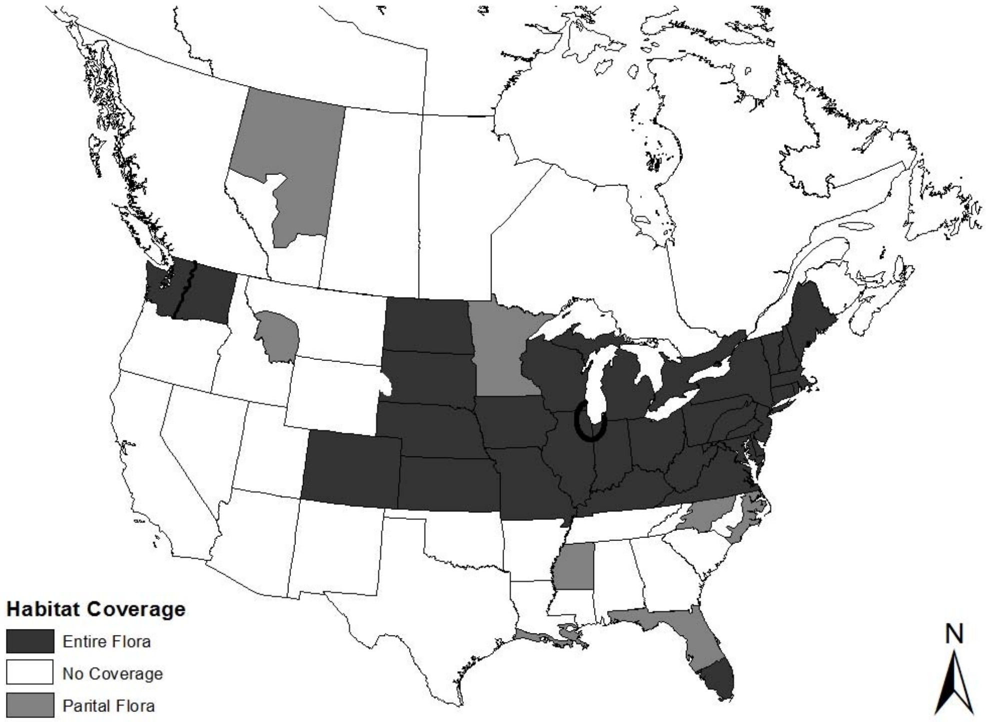


Figure A.1. States and regions where floras have been assigned *C*-values. Partial flora coverages are typically for wetland plants. Several regions species have been scored twice due to overlap in some coverages (e.g., the Chicago Region, parts of New Jersey, Pennsylvania, and the Mid-Atlantic region).

Terminology & Metrics

FQA users must be particularly cautious when navigating the terms and abbreviations associated with the equations, metrics, and concepts underlying it. Considerable confusion has proliferated due to their frequent misuse and improper descriptions. Unfortunately, subsequent users have often deviated from standard FQA texts, and errors and confusion have proliferated (most commonly due to introductions of invalid neologisms or mis-interpreted definitions). To be sure, some of this confusion is attributable to FQA definitions and its lexicon being refined by original authors over time (e.g., Wilhelm 1977, Swink and Wilhelm 1979, Swink and Wilhelm 1994c). However, it's terms, labels, and metrics have been well-established for some time now, and standard FQA texts provide a clear and consistent template to guide use and understanding (Swink and Wilhelm 1994c, Taft et al. 1997). This paper traces these aspects of the history of FQA and highlight preferred language and nomenclature.

Floristic Quality Assessment (FQA)

FQA may most simply be defined as *any* measurement of an area's Floristic Quality that uses plant Conservatism values. The definition of Floristic Quality has evolved over time. Since the first widely available work on the subject (Swink and Wilhelm 1979), its original authors have distilled and refined their initial definition down to a conceptual core; specifically, the idea that anthropogenic disturbance favors less Conservative floras (Swink and Wilhelm 1994a, Taft et al. 1997). And, despite other interpretations by subsequent authors, this is the unifying, canonical concept defining and underlying what FQA measures (see discussion in, *Definitions of Conservatism*). The title "Floristic Quality Assessment" was itself not settled on until 1994. Previous appellations included the "Method for Assessment of Open Land" (Wilhelm 1977, Swink and Wilhelm 1979), "Natural Areas Assessment scheme", and "Natural Area Rating Index (NARI)" . These earlier names were rejected and modified to reflect that the technique had expanded beyond "open lands" to all habitats including forests and aquatic (Wilhelm and Ladd 1988, Bowers and Boutin 2008), and to emphasize that this evaluation method was based on plants (Swink and Wilhelm 1994b). Floristic Quality Assessment is a single method or technique with multiple purposes, uses, metrics, and calculations. Thus, there are not separate Floristic Quality Assessment techniques, as is sometimes stated (e.g., "Several ***FQAI's*** have been developed (Rocchio 2007)"). Users must also be aware that unrelated use of the phrase "floristic quality" occurs in other fields (e.g., Brewer 2010).

In summary, evolving and incorrect use of terminology has created confusion, but inconsistencies should wane over time with careful adherence to standard phrasing and nomenclature established by original authors of FQA (Swink and Wilhelm 1994b, Taft et al. 1997). To further distinguish Floristic Quality Assessment, we suggest that its capitalization will help to highlight that it is a distinct concept with a specific, defined usage, and an established literature, which distinguishes it from other unrelated uses of the phrase.

Conservatism

FQA utilizes Conservatism scores assigned to each plant species *before* its use. Scores are termed Coefficients of Conservatism, or *C*-values. These values have also been abbreviated as CoC, C of C, CS, or CC, but their original, standard abbreviation *C* should be used to avoid confusion. Also, note that its spelling has been incorrect in some FQA studies, and that it is not “Conservativism” or “Conservation”. As previously stated, *C*-values range from 0 to 10; ten indicates a highly Conservative species that is exclusive to remnant habitats without recent or considerable anthropogenic disturbance. Such a distribution largely occurs because such taxa are intolerant of human stressors (see, *Species Conservatism as an environmental indicator*). Rocchio (2007) summarizes the rationale behind scoring:

> “C-values range from 0 to 10 and represent an estimated probability that a plant is likely to occur in a landscape relatively unaltered from pre-European settlement conditions…The C-values essentially represent the collective opinion of local botanical experts regarding a species fidelity to high-quality natural areas, or those areas in which existing plant communities and ecological conditions represent those prior to European settlement.”

Because Conservatism values underlie FQA, accurately defining and understanding Conservatism cannot be overemphasized. Unfortunately, the single most significant and common mistake regarding FQA has been misunderstanding or misrepresenting this concept– what it is exactly that *C*-values measure. In addition to not strictly following Conservatism's original definition, FQA users have also expanded its meaning to include other ecological concepts that have spurious relevance (see discussion below, *Definitions of Conservatism*). To be sure, many mistakenly introduced concepts such as species niche-breadth, specialization, successional status, rarity, invasiveness, etc., are related to species Conservatism, but none of them adequately or precisely encapsulates it, and should generally be avoided. With respect to terminology, users must also be aware that there are unrelated uses of the term "conservatism" in the biological sciences. There are also other ecological measures that are *conceptually* similar to species Conservatism with different names (see *Related concepts*). This overlap means that FQA-based species Conservatism must be clearly identified as a distinct ecological concept with a defined literature to avoid confusion where it is used. This review recommends that FQA users: 1) capitalize its terms (e.g., Conservatism) to distinguish them, 2) and most importantly, users must carefully adhere to the meaning, definition, terminology, and usage rules governing species Conservatism that have been laid out in its standard works, and outlined throughout this review.

Mean C ($\overline{C}$)

Several metrics have been proposed for summarizing *C*-values to measure site Floristic Quality (see, *Other metrics*), but the two original and still most commonly used are FQI and Mean *C*. The Mean *C* (also Mean Coefficient of Conservatism or $\overline{C}$) is calculated as: $\overline{C} = \sum C/S$, where *S* is the number of species at a site and *C* is each species Conservatism value. This terminology has inexplicably transformed over time, where Mean *C* has been variously abbreviated as Mean CC or Mean C of C, but the standard abbreviation is either $\overline{C}$ or Mean *C*. Exotic species may be included or excluded when calculating $\overline{C}$ (discussed in, *Non-native species*). For the sake of clarity, this review suggests that native species calculations be presented as the "native Mean *C*" versus Mean *C* when exotic species are included. It should also be noted that Mean *C* was re-described as a new FQA measure titled, "the modified FQI" (Rooney and Rogers 2002). But, this was an erroneous description, as it has since been pointed out that "the modified floristic quality index of Rooney and Rogers (2002), is simply the mean value of CC (Mean *C*)" (Foster et al. 2007). Therefore, the name "the modified FQI" or "the modified FQAI" (e.g., Reiss 2006) are junior synonyms for Mean *C* that *should be avoided* because they describe a formula that had already been in use for 25 years prior to their publication.

Floristic Quality Index (FQI)

The Floristic Quality Index (*I* or FQI) is the second of the two standard and most commonly used FQA metrics. It is calculated as $I = \overline{C} * (\sqrt{S})$, where *S* is the number of species. Note that **FQA** refers to the concept or process of Floristic Quality Assessment and **FQI** refers to the Floristic Quality Index, one FQA metric. Floristic Quality Assessment Index (FQAI) is also occasionally used to refer to this calculation, but this is an outdated phrase that was rejected in later works (Swink and Wilhelm 1994c), and its use is discouraged in favor of FQI to avoid confusion. As with $\overline{C}$, exotic species may or may not be included when calculating FQI scores. Actually, exotic species may be included in either or both the *S* and $\overline{C}$ parameters of the equation, necessitating that some attention is paid to the methodological specifics when it is used. In practice, these distinctions do not fundamentally alter performance (e.g., Spyreas et al. 2012). Nonetheless users should remember to indicate whether or not exotic species are included, and to label the metrics appropriately (e.g., native FQI). To facilitate ease of computation, the formula for FQI has occasionally been reported as its mathematical equivalent $I = R/(\sqrt{S})$, where R is the sum of C values ($\sum C$) (e.g., Andreas and Lichvar 1995, Lopez and Fennessy 2002, Balcombe et al. 2005). However, because this formulation obscures the fact that FQI is built upon Mean *C*, and because it makes it harder for readers to determine whether exotic species were included, reporting this formulation is discouraged.

Other metrics

Several studies have created alternative or derivative metrics to standard Mean *C* and FQI. Examples of these include the FAQWet index (Ervin et al. 2006), the adjusted FQAI (*I'*) (Miller and Wardrop 2006, Miller et al. 2006), weighted *C* ($\bar{x}Cw$) (Bowles and Jones 2006), adjusted FQAI (Spieles et al. 2006), wCC (Poling et al. 2003), qFQI (Poling et al. 2003), WMF (Anderson et al. 2007), and the modified FQI ($\mathrm{FQI}_{\mathrm{mod}\,t}$) (Cretini et al. 2012). In general, users should be wary of non-standard FQA metrics as their performance may not have been validated. They should also understand that because they typically require additional equation parameters, data manipulation, or data collection, they may be more complicated to use, without significantly improving upon the performance of the standard, highly-vetted metrics (see *Alternate FQA calculations*). Some users have focused on specific subsets of species at sites when calculating scores (see discussion, *Subsets of floras*). For example, plants with *C*-values > 4 (PCC4_10, Bowers and Boutin 2008), or graminoid or forb based scores (grass Mean C, forb Mean C, Smart et al. 2011). These subsets are not considered alongside alternative metrics despite their different labels. Finally, users are also reminded that standard FQA metrics can also be calculated at quadrat or transect scales, rather than for entire site floras. But, again, despite their different labels, these need not be considered alternative metrics because they do not fundamentally change the metric calculations.

In summary, many alternate metrics, terms, and labels exist within the FQA literature. Some of these are erroneous or inaccurate and are to be avoided, while others are only minimally different from standard FQI and Mean *C* metrics, and may have proven useful. Regardless of their relative merits, *any* metrics or calculations that deviate from the standard metrics must be explicitly defined and labeled as such, to avoid adding to the considerable confusion that already exists within FQA terminology, nomenclature, and metrics.

Species Conservatism

History

The term Conservatism was first defined in its modern form in the 1970's by Swink and Wilhelm (1979); when they also assigned scores to the region's plant species and provided a system for their use in FQA. Its definition was refined and clarified in the subsequent edition, which is now widely considered one of two standard FQA works (Swink and Wilhelm 1994a, Taft et al. 1997). While C-values did not exist until the late 1970's, the first known use of the term Conservatism was in the late 1930's (Rocchio 2011). In characterizing Wisconsin prairie plants such as *Stipa spartea*, Gould (1937) described, "*conservative* species that cannot stand extreme environmental disturbances of any kind." Similarly, *Eryngium yuccifoliium* "because of its extreme *conservatism* will probably be among the first of the once abundant prairie plants to become extinct in the county." *Silphium laciniatum* was said to "develop rather *conservatively* as a member of better relic prairie societies." *Conservative* prairie species were described as differing from other prairie species because they could not occupy or spread into roadsides, railroad right-of-ways, pastures, wastelands, or abandoned agricultural lands. These so-called "other prairie species" also differed from "weeds", "introduced weeds", or "sod grasses" that occur where "environmental disturbances" were the greatest. Therefore, although Gould never explicitly defined the term Conservatism, his usage was clearly analogous with its modern interpretation in FQA. Not surprisingly, Gould's *conservative* species (*S spartea, E. yuccifoliium, S. laciniatum*) are highly Conservative species under *C*-value scoring today (9, 8, and 8 respectively in Wisconsin) (Bernthal 2003). A subsequent use of the term extended Gould's *conservatism* by differentiating less-conservative species based on whether they were restricted to black soil versus sand prairies, and by how far they would migrate into non-prairie areas (Thomson 1940). This interpretation was less reflective of modern Conservatism, which makes no community type distinctions. It is not surprising then that Thompson's "most *conservative*" species have relatively modest *C*-values under today's scoring regime (a mean value of 6.2, WI). The final pre-modern use came in Curtis's influential (1959) work describing *conservatism* in prairies. His explanation focused on the climax status of prairie plants. In as far as his description stemmed from species intolerance of massive human disturbance and an inability to quickly re-colonize degraded habitats with dramatically altered environmental characteristics, this description is analogous to the modern Conservatism concept. However, his use of the terms climax and succession to describe *conservative* plants and their habitats implies an adherence to somewhat antiquated theories of succession that are not consistent with contemporary use of FQA (Rocchio 2011).

Related concepts

Users should be aware that there are other unrelated uses of the term conservatism in the biological sciences. The evolution of "conservative" versus "derived" species traits are contrasted (e.g., French and Robson 2012). Qian and Ricklefs (2004) use "ecological conservatism" to describe the degree to which plant genera maintain their environmental associations (climate envelopes, soil types, etc.) over evolutionary time. Similarly, "niche conservatism" is defined as the tendency of species to retain their niche and ancestral ecological characteristics (see also "phylogenetic niche conservatism" (PNC), Wiens and Graham 2005). Due to the tendency of FQA users to equate species Conservatism with species niche breadth, one can see how these phrases have the potential to mistakenly be used interchangeably and therefore confuse FQA users (*Breadth of habitat-types,* discusses why this is problematic below).

Beyond overlapping terminology and phrases, other authors have arrived at *concepts* that are analogous to Conservatism independently. These have most commonly been described for woodland species; for example, "species quality" and "primary woodland species" (Peterken 1974), "ancient forest-woodland species" (Honnay et al. 1998), "vulnerability ranking" (McLachlan and Bazely 2001), "sensitive woodland herbs" (Whitney and Foster 1988), and "old-growth forest specialist species"`(Liira et al. 2007).

But, conceptual analogs have also been conceived for species of minimally disturbed grasslands (Hodgson 1986), or even for non-plant taxa groups ("remnant dependency" in terrestrial insects, Panzer et al. 1995, Twedt 2005). But, none of these authors have specifically defined what these phrases or terms mean, or organized them into concepts that are ready for future or expanded use. Nor have they incorporated them into quantitative habitat/species valuations as has been done for FQA.

The plant scoring system most directly analogous to Conservatism-based FQA is 'hemeroby', which has been used locally in Europe. Like Conservatism, this concept grew out of a desire to quantify naturalness and human impact tolerances, as described in Hill et al.(2002):

> "Hemeroby on the 10-point scale is a measure of human impact varying from 0 (ahemerobic or completely natural) to 9 (polyhemerobic, consisting of pioneer vegetation of railways, rubbish dumps and salted motorways). Most plants in the flora of the Berlin area have been assigned a hemeroby value, with the intention that these values should be used in much the same way as the indicator values of Ellenberg (1979)."

While hemeroby has received some use, it is not popular compared to FQA in North America, perhaps because it may not yet be considered a "a clearly defined concept (Hill et al. 2002)." Furthermore, there are only species values available for the flora around Berlin, Germany, and these values did not work when applied to Britain (Hill et al. 2002). Thus, because the performance of this metric is not known and because it lacks regional scores, its use is limited.

Definitions of Conservatism: avoiding incorrect, inconsistent, and incomplete descriptions

This review has mentioned several problems related to users labeling of species

Conservatism. But, beyond labels, by far the most common mistake of any kind concerning FQA is that users' definitions of Conservatism are frequently imprecise, contradictory, and/or unsubstantiated. Most often this occurs because users have introduced concepts into their definition other than species' remnant dependence and anthropogenic disturbance tolerance. The nonstandard Conservatism definitions that have emerged since the publication of the standard texts have muddled meaning of FQA, and "hindered its widespread utility" (Medley and Scozzafava 2009). Furthermore, additional concepts have not been validated with empirical study in the same way that the concepts of remnant dependence and disturbance tolerance have been shown to dictate Conservatism values (see *Site FQA scores measure human disturbance and biological degradation*). Examples of the mistaken, alternate concepts used to describe Conservatism are detailed below.

Breadth of habitat-types

Habitat-breadth is the most common alternate concept used to define Conservatism. Its use in FQA definitions generally resembles what has previously been described as a species "ecological amplitude", or the number of habitat-types a species will occur in (Packham and Willis 1976). It is also somewhat similar to the "fidelity" or "faithfulness" of plants to a given natural community/association long measured by Phytosociologists (Braun-Blanquet 1932). It has also been defined by FQA users as the range of ecological conditions where a species occurs, which is obviously a parallel concept to the tendency to occur in a specific habitat-type. The use of habitat-breadth also bears some resemblance to other ordinal scale indices that quantify the typical range where a species will tend to occur with respect to a specific environmental variable (e.g., Curtis's Continuum Index, Curtis 1959, Ter Braak and Gremmen 1987, Hydric indicator status, Reed 1988, Ellenberg Values, Diekmann 2003).

Across the FQA literature, this one concept has been referred to with a variety of different terms and phrases including "specialization vs. generalization", "niche breadth", "habitat breadth", "habitat fidelity", "ecological fidelity", and "ecological amplitude".

Under all of these definitions, non-Conservative species lack habitat specificity and occur across a broad suite of habitats (potentially both disturbed and undisturbed), while Conservative taxa are specialized and limited to a few specific *remnant* habitat types. Some studies have gone so far as to consider habitat-breadth as the *only* concept in their Conservatism definition (e.g., Milbauer and Leach 2007, Kraszewski and Waller 2008), thereby, supplanting or ignoring remnant dependence and human disturbance tolerance entirely. With such widespread inclusion into the framework of FQA usage, if this concept is continued to be used, it is imperative that it is verified as an accurate and appropriate Conservatism descriptor, because to this point it has not been.

While the many of the terms listed above are used synonymously within these newer FQA definitions, their meanings outside of FQA can unfortunately be quite different from one another. To begin with, the terms “specialist” and “generalist” have a long debated history in the ecological literature that can be quite unlike their use for FQA. For FQA (e.g., Milbauer and Leach 2007, Kraszewski and Waller 2008), “specialization” is defined narrowly as the breadth of environmental conditions (or habitats) in which a species occurs in. While in the ecological literature, specialization is a much broader term, that is often not synonymous with habitat-breadth (Futuyma and Moreno 1988, Devictor et al. 2010, Clavel et al. 2011). For example, studies can define specialization based on species-specific life-history characteristics, co-evolved mutualisms, food web connections, etc. Therefore, if it is continued to be used for FQA its context must be specified. Another term with a long and unresolved history of debate in the ecological literature is “niche” (Holt 2009). Again FQA users only refer to “niche breadth” in a limited sense— as a habitat specificity— despite the varying and often intractable ways it has been defined and quantified elsewhere (Chase and Leibold 2003, Fridley et al. 2007, De Cáceres et al. 2011, Ricklefs 2012).

To conclude, outside of FQA, specialization and niche are different ecological concepts, whose definitions are complex and disputed across the ecological literature. Not surprisingly, their meaning and use within FQA has often been equivocal or inconsistent across studies. Rooney and Leach (2010) simultaneously use these two different phrases to describe *C*-values, assumedly referring to habitat specificity: "*Niche breadths* of native species were classified according to their coefficient of conservatism (*C*), as habitat *specialists* (*C* = 8–10), moderate *specialists* (*C* = 4–7) and habitat *generalists* (*C* = 0–3)." Therefore, unless there is more consistency and clarification, confusion will likely continue to arise from these concepts use in defining what Conservatism is and what it measures.

The remaining phrases (habitat-breadth, ecological amplitude, habitat fidelity, habitat specificity, etc.), lack the disputed history in the ecological literature. However, the fundamental question of whether they accurately define species Conservatism remains unaddressed. Quantifying each species' habitat-breadth to test for its correlation with *C*-values has not been attempted, undoubtedly because such data would be difficult to obtain for an entire region's flora. However, some example species highlight that habitat-breadth seems ill suited to defining species Conservatism. In Illinois, Conservative species that *regularly* occur in a variety of plant community types, but at the same time are also remnant exclusive, include; *Viola pedata* (C=7), *Amorpha canescens, Ceanothus americanus, Drosera rotundifolia, Gaylussacia baccata*, *Osmunda cinnamomea* (C=8), *Baptisia leucophaea* (C=9), and *Cypripedium candidum, Lilium philadephicum* (C=10) (Masters 1997, Taft et al. 1997).

Taken from the perspective of the habitats rather than species, it might also be argued that anthropogenic habitats themselves may be a rather specific habitat type(s)– offering a narrow ecological niche space or range of environmental conditions rather than a broad one (i.e., general). Many anthropogenic or recently disturbed habitats are quite homogenous and share a set of common abiotic conditions (e.g., high nitrogen availability, high bulk density soil, ample sunlight). Modern agricultural habitats, for example, are subjected to repeated, predictable, disturbances– including tillage, planting, soil fertility treatment, and harvest– that make their resource availability much greater than that of nonarable systems, and that also simplify and homogenize their environmental conditions. This purposefully creates a narrow habitat-breadth (Davis and Landis 2011). From an evolutionary perspective, certain groups of plants are thought to have evolved with, and are adapted to, anthropogenic habitats. For example, in North America certain non-Conservative plants are typically found in association with specific areas in early human settlements (Anderson 1971, Yarnell 1976, Marks 1983). Historically, natural 'disturbance-based niches' would have been less common, and their current ubiquity may obscure the fact that they actually represent a rather specific habitat type. For example, in the Midwestern prairie region the non-Conservative plant *Ambrosia artemisifolia* may have largely been species of bison wallows historically (Taft et al. 1997), whereas it is now widespread because disturbed lands that mimic these formerly localized environmental conditions are now pervasive. Therefore, rather than simply defining them as generalists that are broad in their habitat-breadths, it may be that less Conservative species are just as often restricted to the specific habitats and the conditions created by anthropogenic disturbances (for a discussion of the historic habitat associations of "early-successional" plants see, Marks 1983, Bowers and Boutin 2008).

Since the first iterations of FQA incorporated ecological amplitude or habitat-breadth into their Conservatism definition, it is not surprising that so many users have incorporated this concept (Wilhelm 1977, Swink and Wilhelm 1979). But, eventually the original authors began to voice problems or confusion with scoring species this way:

> "The observed autecological characteristics of some native taxa do not conform well to the previously outlined general concept [habitat fidelity]. For these taxa, while they are usually scarce or rare, we have been unable to discern any predilections for a particular ecological niche or set of niches for which a synecological common denominator can be demonstrated. This could be an artifact of their rarity combined with the mass destruction and desecration of natural lands; the roles of these homeless plants today being those of errant survivors in a world

> that no longer affords them their primal niche. For this reason, a compromise has been forced with respect to their valuation. In most such cases, we have tended to value these species highly, though a few value levels beneath a species of similar rarity but in which there has been observed more fidelity to a recognizable synecological order." (Wilhelm and Ladd 1988)

Not surprisingly habitat-breadth references were notably absent from the third edition of *Plants of the Chicago Region* (Swink and Wilhelm 1994c). Taft et al. (1997) also foresaw problems with using habitat-breadth, and specifically warned against the concept. noting that their scores were not "intending to estimate the degree to which a species is restricted to a certain habitat" Similarly, Masters (1997) advised readers to, "Note that a plant's conservatism is described not by its restrictedness to specific plant communities, but rather to high-quality remnants in a particular region." Therefore, although it may very well be that species *C*-values correlate with the ranges or types of environmental conditions that species occur in, it is not their intent to reflect this.

In summary, habitat-breadth (and its associated terms) should probably be avoided as defining species Conservatism concepts because of the clear examples of species for which it is inaccurately applied, and because of the confusion that these terms' various meanings cause. At a minimum, it is an incomplete Conservatism descriptor that cannot be used on its own. However, justifying its continued FQA use in any form will require: 1) a precise description and argument clearly describing what it means with respect to FQA, 2) avoiding the terms "specialists", "generalists", and "niche" because they have an unresolved, equivocal ecological meaning in the literature, 3) an explicit caveat stating that such a concept is subordinate to the primary concept of remnant dependence in defining Conservatism, 4) an explanation addressing species that do not fit this classification, and 5) quantitative studies to determine how well habitat-breadth correlates with species *C*-values generally.

Species rarity and range

Another concept that is misapplied to definitions of Conservatism is species rarity. For example, Bonanno and Giudice (2010) formally codified this concept into their scoring criteria; “Rare plants considered threatened or endangered were given a coefficient of conservatism ranging from 7 to 10.” However, there are several forms of rarity (Rabinowitz 1981), and the type(s) of rarity referenced when defining Conservatism has not been consistent. For example, *C*-value descriptions have referenced species ranges in their global context: “the degree of human-caused disturbance to an area by accounting for the presence of *cosmopolitan*, native species, as well as non-native taxa” (Fennessy and Roehrs 1997). Local and regional distributions have also been invoked: “Values for *C* are assigned based upon the origin and local or regional distribution of individual species; for example, exotic species and widespread native species receive very low scores, and rare native species receive high scores (Ervin et al. 2006)”. Some studies have not specified the type of rarity or scale, for example, “*C*-values between 2-3…are generally *widespread*” (Rentch and Anderson 2006), and Foster et al (2007), “smaller CC values indicate more *widespread* native species.” These later descriptions probably refer to the distribution within a specific FQA region. Many other studies apply rarity or commonness non-specifically to species *C*-values (Miller et al. 2006, Bowers and Boutin 2008, Marcum et al. 2010).

Several authors have explicitly warned against using rarity (of any type) with respect to Conservatism (Andreas and Lichvar 1995, Herman et al. 2001, Bernthal 2003, Andreas et al. 2004, Rothrock and Homoya 2005, Milburn et al. 2007, Rocchio 2007, MacAllister and Nemeth 2009, Mortellaro et al. 2012). For example, (Andreas et al. 2004) stress that *C*-values in their index “were not assigned based on the rarity of particular species”. The most obvious reason offered for not couching Conservatism in terms of rarity is that rare species are frequently not Conservative. For example, in Wisconsin *Carex pallescens* is a rare “special concern species”, despite having a *C*-value of one because it is found in anthropogenically disturbed areas (Bernthal 2003). *Trifolium stoloniferum* is federally endangered, but it receives a score of five (OH), because it often occurs in modified ecosystems. Similarly, *Populus balsamifera* is endangered and rare in Ohio because it is at the edge of its range; it has a score of three (Andreas et al. 2004). In Illinois, many rare or endangered species are not Conservative (see Figure 3 in, Taft 1994). *Boltonia decurrens* is a globally rare species that is almost totally endemic to a section of the Illinois River Valley, but it is not at all confined to remnant habitats (*C*=4 IL). Alternatively, highly Conservative species can be rather common. For example, *Kalmia polifolia* is described as "a shrub restricted to bogs, which is assigned a 10 [WI], but it is not endangered because bog habitat is still common in northern Wisconsin" (Bernthal 2003). In Ohio, "*Justicia americana* is a common plant…it receives a ranking of 9. Similarly, *Epifagus virginiana* is common in beech-maple forests… it receives a ranking of 10" (Andreas et al. 2004).

At the same time, there is little doubt that most Conservative species are rare or uncommon. This is not because Conservative species necessarily need to be rare, instead it is usually because of the amount of remnant habitat destruction in modern landscapes, or because many highly Conservative species are inherently rare in certain regions (e.g., inhabit rare habitat-types, at the edge of their range). For example, in Illinois, the number of counties native species occur in and their frequency of occurrence in habitat patches across the state are both very highly correlated with species *C*-value (Figure A.2). Because of this correlation, the conceptual distinction between rarity and Conservatism is certainly not intuitive. Indeed, original descriptors of FQA (Swink and Wilhelm 1979) were clearly interested in incorporating rarity into the concept of Floristic Quality, assigning *C*-values from 0 – 20, where *C* values greater than 10 were reserved for rare (15) and endangered (20) species (Swink and Wilhelm 1979, Wilhelm and Ladd 1988, see also, Ladd 1993). However, later revisions removed 15 and 20 scoring species and went so far as to remove all reference to rarity, in order to remove the non-biological value judgment of a species endangerment that could be subject to political whims. More importantly the authors realized that they were not needed and that they caused confusion by including multiple concepts within Conservatism's definition (Swink and Wilhelm 1994a).

To summarize, although it seems contradictory for a measure of conservation value to consciously avoid considering species rarity, users must understand that even the presence of *endangered* non-Conservative plants will add relatively little to site Floristic Quality as defined, and this is why Conservatism was ultimately disassociated from rarity by its original authors. This rarity-Conservatism decoupling actually allows FQA to take on a broader, more versatile utility, as explained by Herman et al. (2001);

> "All too frequently, areas where legally protected species are absent are considered expendable under current formal environmental evaluations. It is precisely because Floristic Quality Assessment is not based on species rarity or legal status that it is a useful tool for assessing the natural quality of an area".

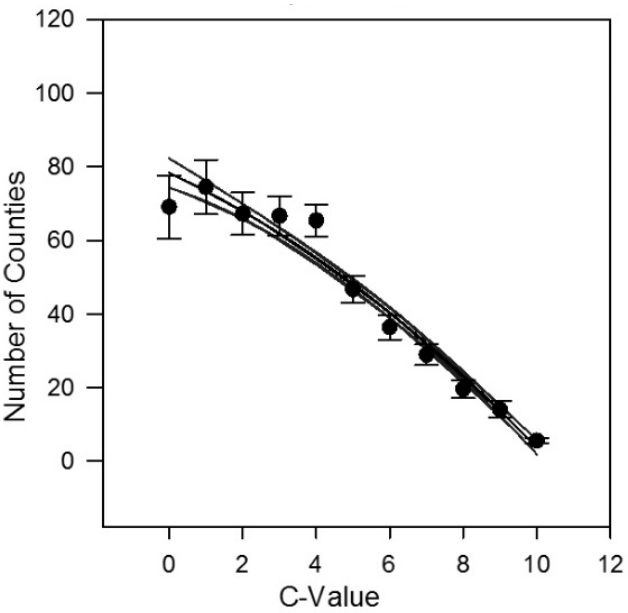


Figure A.2. Relationship between Illinois species' *C*-values (natives only) and the number of counties they occur in ($p < 0.0001$, $r^2_{adj} = 0.50$, $Y = 78.3 + -4.59(x) + -0.29(x^2)$, $N = 1987$), illustrating that species ranges or distributions across that state are highly correlated with their Conservatism. Regression line is bounded by 95% confidence interval and 95% confidence interval bars are shown. Data summarized from ILPIN (Iverson et al. 1997).

Successional status

High Floristic Quality habitats are commonly equated with "mature", "late", "advanced", "climax", or "stable" successional states (Wilhelm and Ladd 1988, Swink and Wilhelm 1994b). Recently species Conservatism has also been defined based on successional status– species tendency to be found in late-successional, stable, or "high-seral" habitats. For example, it is occasionally the sole concept defining Conservatism: "The higher the *C*-value, the later in succession the plant would be expected to appear (Middleton and Bever 2012)." Pluralistic definitions that include succession as one of several defining concepts are more common. For example, Foster et al. (2007) based their definition on successional status, niche breadth, and species range: "CC values range from

0 to 10 with large values indicative of high-quality prairie species of narrow ecological niche and late-successional status. Smaller CC values indicate more widespread native species of earlier successional stages." Curiously, Gutrich and Hitzhusen (Swink and Wilhelm 1979, Gutrich and Hitzhusen 2004) invoked successional status in defining some ranges of *C*-values, but not others: "values between 4–6, species typical of a successional phase of a native community…7-8 taxa typical of stable or 'near climax' conditions" (see also, Bernthal 2003, Bonanno and Giudice 2010).

Regardless of how they are incorporated, there are several problems with including succession-based descriptors. The first is that FQA studies use the classic concept of succession, which represents an understanding of habitat dynamics that is increasingly viewed as specious. Classifying habitats as late- versus early- successional, or as moving inexorably towards a pre-determined climax, is not meaningful or ecologically accurate for most habitats (i.e., Clementsian successional theory). And, because it is difficult to identify the climax, or terminal successional state for most areas, it is difficult to classify what constitutes a late-successional species within this framework. There are Conservative species adapted to habitats that undergo frequent natural disturbance. For example, *Polygala incarnata* (*C*=10 IL) occurs in prairie habitats with regular fire and grazing (Taft et al. 1997). Many early-successional habitats such as cliff faces, beach dunes (Schwegman et al. 1973, Galatowitsch et al. 1999, Gibson et al. 2011)} are characterized by highly Conservative species. This highlights another inconsistency: the distinction between successional status and habitat stability. Conservative species are attributed to "stable" habitats. For example, Reiss (2006) describes highly Conservative species as, "characteristic of stable, reference ecosystems." A habitat's stability is largely determined by its natural disturbance regime, and as previously discussed (*Natural versus human disturbance*), FQA is not dictated by natural disturbance. Therefore, habitat stability (outside of human disturbance) is also a problematic concept to associate with species Conservatism.

To summarize, the repeated FQA-succession association is certainly understandable, as it stems from original FQA authors' observed link between high Floristic Quality communities and successional status (Swink and Wilhelm 1994c, but see, Taft et al. 1997). However, their brief anecdotal observation probably should not merit its continued use as a fundamental concept underlying species Conservatism. Furthermore, recent study has shown that community succession and Floristic Quality are not consistently related (Spyreas et al. 2012). Although there is a desire to link FQA with established succession theory, coupling Floristic Quality with a directional, deterministic, successional framework in communities is not supported. Because of this, arguments for *species* Conservatism values also being dictated by their successional behavior are not reliable either. Barring further study, concepts of habitat and species succession and stability should be avoided when defining FQA and species Conservatism

Plant traits and life history characteristics

Consistent patterns in plant characteristics or life-history traits separating early- versus late-successional plants have long been sought by ecologists (e.g., seed dispersal ability, reproductive capacity, seed germination cues, seed bank persistence, photosynthetic capacity, leaf traits, *r* vs. K evolutionary syndromes). More recently authors have asked whether there are traits that typify species exclusive to high quality remnant habitats (e.g., Meier et al. 1995, Meador and Carlisle 2007, Raab and Bayley 2012). Transient seed bank, germination requirements, age to reproduction, fecundity, growth form, phenology have all been screened in this way in European forests (Hermy et al. 1999). Similarly, species of high quality wetlands have been associated with traits that allow plants to tolerate particular water quality, nutrient conditions, soil pH, flood pulse, and flood requirements (Nelson et al. 1994). Not surprisingly, FQA users have attempted to associate species Conservatism with specific life-history traits and plant characteristics (e.g., Brudvig et al. 2007). The most comprehensive argument for associating plant characteristics with species Conservatism was made by Taft et al. (1997). They suggested Grime's CSR life history categories as a conceptual guide to assigning species *C*-values (Grime 1974, 1979), where "stress tolerators" could describe highly Conservative species, moderately Conservative species are "competitors", and non-Conservative species could be characterized as "ruderals". Other authors have since incorporated this framework into their definitions (e.g., Landi and Chiarucci 2010, Bried et al. 2012). For example, Miller et al. (2006) define low *C*-value taxa as "a class of plants that are comparable to Grime's ruderal and ruderal-competitive species guilds…species that are able to rapidly grow, flower, and set seed and thus tend to withstand frequent and severe disturbance."

While this categorization scheme is appealing because it provides an established theoretical framework for FQA, CSR theory neither adequately nor accurately defines Conservatism in all cases. For example, in the Midwest many types of pristine remnant habitats are not "stressed" habitats in the classical sense– areas where plant productivity is continuously limited by the environment. Highly productive tallgrass prairies have adequate soil nutrients, sunlight, moisture, and moderate temperature regimes. Yet, their undegraded site FQA scores are equivalent with characteristic "stressed" habitats in the region (e.g., xeric sand prairie, rocky exposed bedrock woodlands) and they contain many highly Conservative species that would not be considered stress-tolerant. Taft et al. (1997) recognized this limitation and highlights that species "having *C*-values of 7-10 are less clearly aligned with Grime's model*."* By their own admission, there are several species whose *C*-values do not consistently fit into these different categories, such as Conservative annuals and biennials that are not stress tolerators. Furthermore, the group known as "Competitors" is also troublesome because its species are partially characterized by their dominance or abundance: "Those species with coefficients 4-6 include many **dominant** or matrix species …they have a high consistency of occurrence."(Taft et al. 1997) However, Conservatism values by definition do not reflect abundance (Swink and Wilhelm 1994a). Finally, because Grime's ruderal categorization did not distinguish between natural and human disturbances, more Conservative natural disturbance adapted species could potentially be considered ruderal. Therefore, while CSR strategies may provide an interesting theoretical reference, they do not adequately describe Conservatism and should probably not be invoked when defining it.

In regards to plant traits and species Conservatism generally, there seems to be little doubt that there are some patterns with respect to life history characteristics and Conservatism. For example, only a cursory review of low *C*-value species shows that they are more likely to be weedy annuals than higher value species. Even some more specific physiological characteristics may associate with C-value ranges, for example, in prairie pothole wetlands plants tolerant of brackish water are said to have high *C*-values (Euliss and Mushet 2011). However, it is important to remember that in any region's flora, both highly Conservative and non-Conservative species will be represented among nearly every functional group, life form, and type of life history. This pluralism in trait values suggests that the reasons for Conservatism probably vary by life form, habitat, and species, and that consistent high *C*-values characteristics across the flora does not exist. Therefore, defining species Conservatism by plant traits should be avoided.

Multi-concept definitions

As described above, these auxiliary FQA concepts are inconsistent with standard definitions, and they are often unsupported. Another trend has been to combine several of these concepts in the Conservatism definition. For example, a recent study incorporated six different concepts into its definition (e.g., disturbance tolerance, remnant dependence, CSR life-strategies, species abundance, successional status, niche breadth, Landi and Chiarucci 2010, Raab and Bayley 2012). This has then led to authors to conflate concepts and use them interchangeably throughout the study, which has created even more confusion.

Another inconsistency in definitions occurs when different Conservatism concepts are assigned to different *C*-value classes. For example, Gutrich and Hitzhusen (2004) describe "values of 1 – 3: taxa that are widespread and not indicative of a particular plant community; values of 4 – 6: species typical of a successional phase of a native community;

values of 7 – 8: taxa typical of stable or ‘‘near climax’’ conditions; and values of 9 – 10: taxa exhibiting high degrees of fidelity to a narrow set of ecological parameters.” In this example, species geographic range (rarity), habitat-breadth, successional status, and fidelity to a set of ecological parameters, are all invoked at some Conservatism levels but not others. Other examples of assigning concepts differently across *C*-value ranges abound (Bernthal 2003, Jog et al. 2006, Rentch and Anderson 2006, Bonanno and Giudice 2010). In such cases, the authors do not explain why one ecological concept (e.g., succession) applies to one class of scores (e.g., 4-6), but not others. Surely successional processes or geographical ranges are relevant to all species occurrences. Under such a framework, species in different groups are suggested to be operating under different ecological rules, rather than species having a gradation of responses to ecological forces.

In summary, while a simple, single-concept definition that is based on remnant habitat occurrences and disturbance tolerance is the preferable approach (Taft et al. 1997), it is unfortunately not typical due to more recent inclusion of concepts into definitions. The primary argument against them is their questionable merits and inconsistency. But, it is also true that a remnant-dependence based definition is quite adequate on its own, and it is also the more parsimonious approach when used on its own. Re-establishing this conceptual simplicity will help FQA users and readers avoid the confusion and inconsistency that has increasingly plagued it.

Regionalization

FQA is based on each species’ Conservatism value, which is specifically assigned to each plant in a designated FQA region (typically a state). Values are only relevant to the state they were developed for because they are assigned in relation to other elements in the local flora (Wilhelm and Ladd 1988, Bried et al. 2012). And, as a species’ physiological and ecological variations within its range cause its behavior to vary across its distribution (Rocchio 2007), regional *C*-values also reflect these spatial dynamics. During the species *C*-value assignment process (*Assigning Plant C-values*), botanists are asked to think about a species' statewide behavior, often across states with expansive latitudes and longitudes (e.g., Illinois, Florida, Dakotas). Species' *C*-values are less accurate further from their origin (Rothrock and Homoya 2005, Bourdaghs et al. 2006, Ervin et al. 2006), which is sometimes considered the center of the state they were assigned in (Wilhelm and Ladd 1988, Landi and Chiarucci 2010). An illustrative example of assigning scores to species with difficult bimodal behaviors is presented by Taft et al (1997),

> "On occasion, during the coefficient assessment phase of this project, we needed to evaluate taxa that demonstrate regional behavioral differences in Illinois, such as *Asclepias tuberosa* and *Oxalis violacea.* These species are occasional to common in degraded habitats in far southern Illinois, but in central and northern Illinois they are more restricted to remnant areas. In these instances, we assigned an intermediate value such as 5."

Similar examples are given elsewhere (Milburn et al. 2007). While there is no evidence that overall *site FQA values* are biased or inaccurate at the periphery states because of this affect, such an inquiry would be useful. An evaluation of error and imprecision when scores from neighboring states are used in studies from states for which scores are not yet available would also be important.

C-values and disturbance thresholds

The philosophy behind species scoring has been described by botanists in the Chicago Region with the following scenario:

> "Someone brings us a specimen of *Lepidium virginicum,* and asks, on a scale of 0-10, how confident we are that the specimen was taken from a remnant natural plant community. We would have to say that we have no confidence, but that in all likelihood it was taken from a highway verge or the edge of a parking lot. Someone else brings in *Aster borealis,* and we are virtually 100 % confident that it was taken

> from a remnant…one that was not terribly degraded. Another brings in *Galium triflorum;* we are fairly certain that it came from some sort of remnant wooded area, but we can say little one way or the other about how degraded it is. In the first case, we can assign *Lepidium virginicum* a *C* of 0, since we had no confidence that it came from a natural community. *Aster borealis,* on the other hand, can be assigned a *C* of 10 since we have virtually 100% confidence that it came from an intact natural community. *Galium triflorum* can be given a 5 *C*-value since we are certain that it came from a remnant natural community, but have little confidence that the area was not degraded." (Wilhelm and Masters 1995, for a similar description see, Rothrock and Homoya 2005)

This anecdote describes the general philosophy surrounding Conservatism well, but it also raises a less-apparent theoretical incongruence. Specifically, it exposes the two ways that a species' Conservatism value might reflect its occurrences across the spectrum of disturbed and pristine habitats in the landscape (Figure A.3). The first way (bottom panel) is that a species' Conservatism value could reflect the most disturbed habitat a plant occurs in. Regional botanists would essentially be delineating the lower limit of anthropogenic disturbance a species seems capable of tolerating, as registered by its occurrence in disturbed versus undisturbed habitats. This may be termed a *lowest disturbance threshold* approach. With the second philosophy (bottom panel), a *C*-value reflects the habitat disturbance level that a species is most frequently found in. This may be termed the *most likely occurrence* approach.

In the passage above, Wilhelm and Masters (1995) mainly invoke a 'lowest disturbance threshold' perspective. For example, the most degraded habitat that *G. triflorum* (*C*=5 Chicago Region) can occur in is still recognizable as a natural area (i.e., "some sort of remnant wooded area"). It can also be assumed from their description that if such a barely tolerable woodland were to incur further disturbance this species would be eliminated from it. At the same time, *G. triflorum* occurs in undisturbed woodlands (i.e., higher quality remnants) with some regularity. So it is not clear where it most frequently occurs. Similarly, even though *L. virginicum* grows in the most anthropogenically disturbed habitats that plants grow in (e.g., parking lots), and is among the most highly disturbance tolerant plants in the flora, it is not restricted to these most highly disturbed habitats and it also occurs with some frequency in higher quality areas. Indeed, some non-Conservative taxa have been noted for their seeming indifference to habitat quality and they readily occur in disturbed and undisturbed habitats (e.g., in Illinois, *Equisetum arvense*, Taft et al. 1997). Thus, the 'lowest disturbance threshold' perspective sets the lower habitat degradation limit and says little about occurrences in higher quality habitats.

To understand the *most likely occurrence* approach, consider the following section from the passage above, "in all likelihood it [*L.virginicum*] was taken from a highway verge or the edge of a parking lot". The words "in all likelihood" suggest that what defines *L. virginicum*'s *C*-value is that it most frequently occurs in highly disturbed areas. The authors are invoking a *most likely occurrence* perspective here. Bried et al. (2013) also invoke such a perspective: "[*C*-values] are assigned, typically for each state as a whole by mentally 'averaging' the species' behavior". It is also not uncommon for authors to incorporate both concepts in their works. For example, in Taft et al. (1997), *C*-values were defined as follow. "[where] native species are *most successful* in badly damaged habitats [they] were given *C*-values of 0. At the other end of the spectrum, species virtually restricted to natural areas received *C*-values of 10." Note that the description for species with *C*-values of 10 suggests a threshold to species occurrence by being "restricted", whereas the *C*-value description for a zero value implies a frequency based definition (i.e., "*most successful* in badly damaged habitats"). While most authors seem to imply a combination of the two philosophies in their definitions, the *most frequent occurrence* perspective is probably more heavily emphasized across the literature. The only authors who seem to explicitly take a single-concept perspective are Cohen, Carstenn et al. (2004), who considered their *C*-values as the mean or median habitat quality in which a plant occurs (i.e., a *most frequent occurrence* approach) . They also provide the only empirical test of whether a *most frequent occurrence* species scoring model fits their occurrence data, by comparing species scores with a measure of the average habitat quality they occurred in.

A threshold-based explanation may be less frequently emphasized because it is the less intuitive. For example, a highly Conservative species *must* reach its greatest occurrence-frequency-abundance in undisturbed remnants, because by this definition that is the only place it occurs. This may obscure that its distribution represents a threshold, rather than simply where it most frequently occurs. In most cases non-Conservative taxa actually are "most successful" in highly degraded areas, as this is where most weedy species prosper (e.g., *L. virginicum* from the example above). But, it is critical to remember that ***some non-Conservative species may just as frequently be found in remnant***s. An absolute threshold to species presence is of course impossible to precisely delineate and define, as Conservatism is based on botanists' experience and judgment. Thus, perfect compliance to a precise threshold should not be expected as aberrant Conservative individual occurrences will occasionally be found in degraded areas– often in small populations or single individuals in less-degraded micro habitats embedded within otherwise highly disturbed sites (personal observation).

Finally, it should be noted that a third, rather unorthodox *C*-value occurrence pattern has been presented (Mortellaro et al. 2012). Here, a species with a *C* value of 5 is described as being "obligate to native lands, with a 95% certainty it is growing in a natural area. However, the floristic quality of the natural area is likely low." This description uses a most likelihood occurrence approach (e.g., 95% certainty in a natural area), but with the further stipulation that the natural areas that a 5 *C*-value species occurs in are almost always degraded. This is a problematic criterion because it means that remnant pristine habitats could *only* contain species with *C*-values of 10. Such a scenario is highly unlikely when one considers that their study region contains approximately 150 species with *C*-values of 10. Thus, all of their pristine habitats habitat types (grasslands, forests, etc.) can only contain a species pool of perhaps a couple hundred possible species. Similarly, in most native habitats the most common matrix species (dominant plants with most of the habitat's total biomass, such as trees in forests and grasses in prairie) tend to only be moderately conservative, but they tend to be frequent in both degraded and pristine natural areas. Thus, this perspective on *C*-values is untenable.

In summary, these two concepts, the *most frequent occurrence* and the *disturbance threshold* approach are rather distinct (Figure A.3). In their Conservatism definitions, users should be mindful not to conflate them or use them interchangeably. However, even though they are mutually exclusive concepts, in reality, species distribution patterns are likely to follow an intermediate pattern somewhere between these two extremes. Some species are probably better described by one distribution or the other. Comparing whether threshold versus more frequent occurrences better describe species distributions with empirical data is needed, even if clearly determining a single best descriptor may ultimately be elusive.

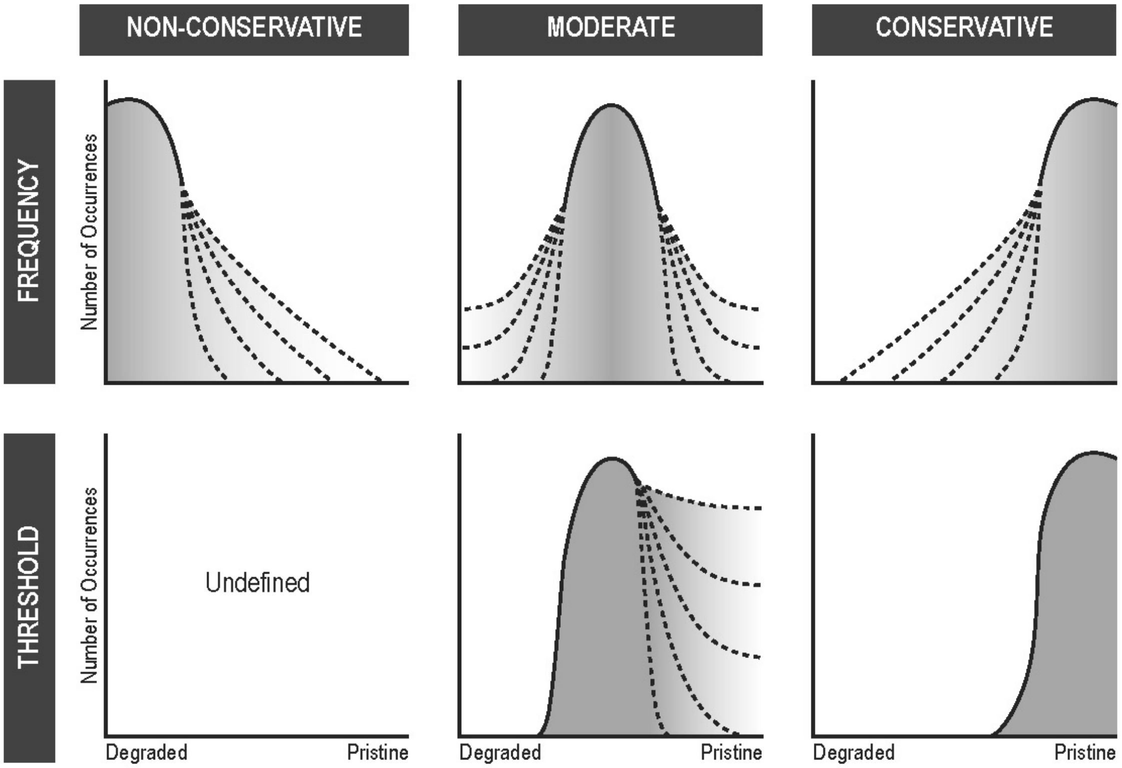


Figure A.3. Two potential species distribution patterns in degraded versus pristine remnant habitats for non-Conservative (left panel, 0-3), moderately Conservative (middle panel, 4-6) and highly Conservative species (right panel, 7-10)

Natural versus human disturbance

Anthropogenic (as opposed to natural) disturbances to habitats have distinct and well-established effects on different taxa groups (Karr 1981). Vegetation is considered a particularly sensitive measure of human stressors and disturbances, and their impacts on ecosystems (U.S.E.P.A. 2002b). For FQA specifically, the Conservative species are sensitive indicators, because they do not occur in anthropogenically disturbed habitats. This critical distinction between natural and human disturbance tolerance has occasionally not been made by FQA users. For example, (Miller et al. 2006) explained that their

"Conservatism values were assigned based on an individual plant species' fidelity to specific habitat types and tolerance to both ***natural*** and anthropogenic disturbance" (see also, Andreas et al. 2004). This has led to the misconception that Conservative species do not tolerate disturbance of any kind.

Intolerance to any disturbance would mean that Conservative species could not occur across vast regions of disturbance-adapted ecosystems. For example, there would be no Conservative species over significant parts of central North America, where landscape-scale fires occurred almost biennially (McClain et al. 2010). Or, at a smaller scale, certain habitat types, such as riparian systems that are frequently disturbed by flooding, could also not harbor Conservative species, even where they are pristine remnants (Rocchio 2007). Some authors have therefore taken care to clearly make this distinction in their definitions. In West Virginia, *Spiraea virginiana* was highlighted for its Conservatism (C=7), despite being "restricted to areas of deposition along banks of high gradient, highly scoured streams, where flood events are a necessary part of its life history strategy" (Rentch and Anderson 2006). Similarly, in Ohio, Andreas et al. (2004) described that *Justicia americana* (C=9) "grows almost exclusively on gravel bars of mainstem rivers", which naturally incur frequent disturbance. Furthermore, many highly Conservative species actually in decline in *protected, remnant habitats* lacking regular disturbance (e.g., fire, historic grazing) (Taft et al. 1997). Natural disturbance from flooding has been said to facilitate Floristic Quality increases after habitat degradation by filtering out highly competitive, less Conservative species over time and allowing Conservative species opportunities to re-establish (Stanley et al. 2005).

The importance of natural disturbance regimes for site Floristic Quality must not be overlooked or oversimplified. FQA rests on the assumption that Conservative plants decrease with anthropogenic disturbance, and either increase or are maintained under natural disturbance regimes. This framework assumes that natural disturbance regimes functionally differ from anthropogenic disturbances. However, one can clearly find instances where the two may be difficult to distinguish, and a dichotomy may not always be meaningful. For example, effects of moderate livestock grazing on grassland vegetation could approximate that of historic wildlife (Towne et al. 2005). In forests, selective timber harvests may mimic natural tree blow-downs, while clear-cuts could have effects similar to large-scale wind or ice-storm damage (but see discussion in, Lindenmayer and Laurence 2012).

Thus, a clear explanation of how or where natural and anthropogenic disturbance regimes differ in relation to FQA is fundamental to its understanding. Some authors have tried to define the distinction. Bried et al. (2012) describe “disturbances and stressors occurring outside of the environmental variation to which the species is evolutionarily adapted.” Roccio (2007) summarizes Wilhelm and Masters (1994) in describing human impacts since European settlement that cause “many ecological processes and disturbance regimes to function outside their natural range of intensity, frequency, or duration…[where] Conservative plants are not adapted to these human induced alterations and thus, are typically the first plants to disappear from a habitat impacted by human activities.” Rentch and Anderson (2006) distinguish between “disturbance as components of a natural disturbance regime, and degradation, which refers to habitat altering events of a largely anthropogenic origin, and that occur with a frequency and/or intensity that is beyond the long-term variability of the natural disturbance regime.” Finally, the evolved relationship among Conservative species, *natural* disturbance regimes, and remnant habitats has been summarized as follows:

> “Conservative floristic elements are those species that, through millennia, have become supremely adapted to an environment determined by a specific set of biotic and abiotic factors, interactions, and natural disturbances. These factors include local edaphics and extremes of drought, humidity, inundation, fires, temperature, and faunal interactions, etc. Though these factors have changed over time, the changes have been gradual enough and buffered sufficiently by system complexity

> to allow gene pools to adapt. When changes [to environmental conditions and disturbance regimes] occur rapidly, as they have in the post-settlement period, these conservative species on a given tract of land are reduced in accordance with the severity of the changes." (Wilhelm and Ladd 1988).

Although these explanations are somewhat cursory and abstract, they do contain common points. Anthropogenic disturbances retard *Conservative species* because they differ from natural disturbances quantitatively, in their frequency or intensity, or qualitatively, in that they are evolutionarily novel. In contrast, the evolutionary relationship between non-Conservative species and remnants is probably more varied, and it is certainly less clearly defined in the FQA literature (but see discussion of human habitats in *Breadth of habitat-types*).

In conclusion, the creators of ecological indicators struggle with the "difficulty in separating anthropogenic from natural sources of variation" (Niemi and McDonald 2004). Further research on the frequency, periodicity, type, and/or intensity of natural versus anthropogenic disturbance will be needed to clarify their differences. However, in regards to the utility of FQA, as long as species occurrences in natural versus anthropogenically disturbed habitats are distinct and apparent to the botanists charged with assigning Conservatism values in regional floras, FQA will be meaningful for its intended purpose.

Species Conservatism as an environmental indicator

Plants are well-studied environmental indicators (Adamus and Brandt 1990, Adamus et al. 2001). Their immobility allows them to register the cumulative effects of disturbances and environmental changes at sites over short and long terms, simply by their presence or absence. And, as a diverse group with species of varied sensitivities and tolerances to their physical environment, plant assemblages are especially useful for registering environmental fluxes and stresses at sites. Some authors argue that the "major challenge of ecological indicators continues to be the difficulty of discerning specific stressor-response relationships in a multiple stressor environment" (Niemi and McDonald 2004), but discerning individual stressor-species relationships has never been the goal of FQA. Instead, FQA is assumed to relate to the entire suite of site characteristics differentiating remnant versus disturbed habitats (environmental, biotic, etc.). *C*-values are not based on any kind of measured bioassay of species' environmental stress tolerance (e.g., pollutants or water quality variables in aquatic IBI's, Meador and Carlisle 2007). Indeed, Swink and Wilhelm (1994c) explain that Conservatism is based on the premise that it is "impossible to quantify all of the ways that people degrade habitats." They instead argued for assigning *C*-values based on species exclusiveness to remnant versus anthropogenic habitats, without consideration of the mechanisms, specific disturbances, or environmental conditions that underlie these occurrence patterns.

However, FQA users do occasionally use *C*- values (via site FQA scores) to measure relative levels of anthropogenic disturbance to habitats, and also to isolate effects from individual habitat stressors (e.g., grazing, fire exclusion, changes to environmental conditions; see, *Site FQA scores measure human disturbance and biological degradation*). In such instances, users should be clear to point out that a site's Floristic Quality is a *direct* measure of its species' remnant dependence, and it is only an *indirect* indicator of anthropogenic disturbances and environmental (physical/chemical) stresses to the habitat (although it has been shown to be an accurate indirect measure). This distinction is made because the likelihood of any one species being found in a given remnant could presumably be affected by more than the direct anthropogenic *physical disturbances* or *environmental stresses* that a habitat has incurred. Put another way, a Conservatism perspective based solely on sensitivity to changes in environmental conditions assumes that species occurrence is exclusively driven by extirpation from remnant habitats due to human changes to environmental conditions at a site. But, species' relative dispersal and colonization ability can also affect their remnant fidelity (see also, *Species Conservatism and restoration plantings*). For example, a disturbed remnant habitat may recover in environmental conditions to its pre-disturbance state. But, among the species lost from this habitat after the initial disturbance, some would be less likely to recolonize what should be an otherwise environmentally suitable habitat, and these would be the most Conservative of the species lost and the most remnant dependent (e.g., they are poor dispersers, have low fecundity, disrupted mutualistic associations, or because of an inability to overcome priority effects at the site). Thus, a species' remnant dependence may be dictated by any number of its life history or population characteristics (e.g., susceptibility to pathogens, rarity, generation time, population cycling, seed bank), or altered site conditions, that are not dictated by on-site physical disturbance and changed environmental conditions. A strict environmental indicator species would only reflect environmental and abiotic conditions at the site, whereas the remnant dependence of Conservatism is a broader, more integrated ecological and historical concept.

To summarize, FQA may be used as an indirect environmental indicator because:

> "Comprehensive, regionally calibrated lists of vascular plant species' [environmental disturbance] tolerances are generally lacking due to the number of species in the flora, the individualistic responses of plants to different types of disturbance, and the lack of experimental studies. Given this problem, the Coefficients of Conservatism provide a useful surrogate for experimentally derived species tolerances…[and] this tolerance (or sensitivity) as ranked by the C of C's is a useful tool for evaluating its ability to tolerate human disturbance." (Andreas et al. 2004)

Using species *C*-values as indicators of human disturbance or stress requires the caveat that they are not direct assays. Instead, they reflect the plurality of ecological factors dictating species distributions and remnant dependence. But, because site-level FQA scores incorporate the entire assemblage's disturbance/stress tolerances, they collectively provide enough redundancy in their responses to make an accurate proxy indicator of this type.

Species Conservatism and restoration plantings

It is not clear where habitat re-creations or restorations and their floras fit within Floristic Quality Assessment and how they relate to species Conservatism. Re-creations are highly-disturbed, non-remnant habitats, and by this definition they should not be amenable to Conservative species. At the same time, their purpose is to provide a habitat that will harbor a diverse and presumably Conservative native flora. Restorations may simply be a way to provide the habitat and conditions for Conservative species to establish in or they may be planted with these species. Planting artificially circumvents the dispersal and propagule limitation that may otherwise limit many Conservative species' distributions. It is not clear how likely Conservative species are to establish and persist in restorations compared to less Conservatism species, nor is it clear how one should interpret the ability of any Conservative species to establish in non-remnant habitats such as restorations, if their establishment is achieved by planting. Published definitions of *C*-value lists provide no guidance as to whether or not botanists considered occurrences in recreations and restorations in their species scoring.

Only a single study has looked for a relationship between *C*-values and plant likelihood of establishing in newly created habitats (Haan et al. 2012). They found that *C*-values of nine different wildflower species did not predict whether seedlings would establish when planted into roadside environments, although sample-size was small. However, because habitat restorations generally do not achieve the overall Floristic Quality of undegraded remnants, it can be assumed that Conservative species are generally not likely to establish either on their own (Foster et al. 2007, Spyreas et al. 2012), or when introduced into restorations. (Although some restorations can score very high in their FQA, and some Conservative taxa can establish in well-executed restorations (see *Restoration versus remnant* and *Restoration monitoring*)). Given the difficulty of habitat restoration due to the proliferation of non-Conservative often weedy plants in them (Matthews and Endress 2008, Matthews et al. 2009b, Matthews and Spyreas 2010), it would seem that

Conservative species behavior in them generally matches what might be expected from disturbed habitats.

Specific examples of difficult to establish Conservative species abound. Mead's milkweed (*Asclepias meadii*) is highly Conservative (IL=10), and although it was rather widespread in prairie remnants historically, it has proven nearly impossible to re-establish-restorations and remnants included. As a herbaceous plant that may live for more than one hundred years with very low seed output (Betz 1988), it may be an extreme and atypical case of Conservative plant behavior in restorations. Conservation practitioners are notoriously unsuccessful however at (re)establishing rare and endangered species populations in general, suggesting that this may be as much a characteristic of rare species recovery, as it is of Conservative species (Godefroid et al. 2011).

The autecological complexities, specific life-history requirements, and habitat requirements that many Conservative species could require may simply not be present in most restorations. Therefore, although dispersal barriers and propagule limitation can be overcome by planting, inappropriate site conditions may be insurmountable for species. The list of factors suggested as potentially impeding long-term establishment by Conservative taxa in restorations includes: local microsite availability, high resource competition, high seed predation, weedy competitors, and soil conditions such as a lack of specific mutualistic microbiotas (e.g., Orchidaceae, Fabceae), soil pathogens, pH, and nutrients (Bever et al. 2003, Foster et al. 2007, Heneghan et al. 2008, Middleton and Bever 2012).

In the future, authors describing *C*-value lists for statewide floras should be explicit with respect to whether occurrence in restorations has been considered in their scoring, and therefore, with whether species Conservatism is meant to include both remnant and re-created habitats. Future studies might retroactively test this for existing flora values by surveying regional habitat restorations to determine how likely species with higher C-values are to occur in them. Analyzing seed mixes from restoration plantings might also allow for comparisons of the relative likelihood of establishment among species.

Assigning Plant C-values

Species scoring is subjective

Species *C*-values are often criticized as being subjective, and therefore, potentially imprecise. It is certainly true that values are subjectively determined, as scoring is based on botanists' estimates of species values (see, *The scoring process*). However, subjectivity itself does not preclude utility, nor does it imply that FQA lacks scientific rigor. Some authors have gone so far as to argue that a higher level of reliability comes from botanists' experienced-based species evaluations than from what might be created with empirical data:

> "…Conservatism cannot be determined from some statistical sampling of an array of plant communities, ruderal or otherwise. The body of empirical observations that inform the experience of a field botanist over a long period of time yields more robust evaluations." (Mortellaro et al. 2012)

While it is an overstatement to say that species scoring could not benefit from more quantitative assessments, if only for quality assurance or evaluation purposes, even in their current state where they have been assigned based solely on botanist opinion, there is no doubt that site FQA scores successfully measure what they were intended to measure (see *Site FQA scores measure human disturbance and biological degradation*). FQA is not about individual species assigned values, but how well site scores work. Potentially imprecise or mis-assigned species scores are ameliorated when aggregated into FQA metrics (Wilhelm and Ladd 1988, Matthews 2003), in the same way that individual species' Ellenberg species values become more useful "when they are summarized as community- level indices … provide a more reliable measure of ecological change than do those based on individual taxa, because the overlap of ecological tolerances of multiple species is smaller than the amplitude of a single one" (as quoted from Diekmann 2003, LaPaix et al. 2009).

An argument repeatedly offered against claims that *C*-value subjectivity invalidates FQA is that values are set *a priori* to their use, and they are therefore consistent. Any biases in them are fixed, one-time occurrences. Once the numbers are assigned FQA users are thenceforth required to apply it objectively. Measuring site ecological condition is then done in a repeatable and dispassionate fashion (Swink and Wilhelm 1994b, Masters 1997, Taft et al. 1997, Francis et al. 2000, Herman et al. 2001, Andreas et al. 2004, Milburn et al. 2007). Therefore, "relative comparisons between sites (site 1 to site 2 to site 3), using Mean *C*, should not be distorted because any personal bias would be uniform across all sites (Herman 2005)."

Bried et al. (2012) summarizes these arguments:

> "Studies of floristic Conservatism have cautioned that subjectivity and disagreement in botanist opinions may ultimately affect interpretations of Floristic Quality (Bourdaghs et al. 2006, Cohen et al. 2004, Landi and Chiarucci 2010). Although differences of opinion may exist at the species level, floristic quality assessment is based on relative aggregate conservatism…Subjectivity is inherent to the thought process, but at least the thinking by a given botanist is evenly applied across species and potential bias gets loaded "up-front" in the overall assessment, meaning the subjective *C*-values are used objectively to estimate floristic quality (Andreas et al. 2004, Herman et al. 1997). Despite the subjectivity in assigning *C*-values, floristic quality assessment has repeatedly shown the negative linear response (to varying degrees) expected over a gradient of increasing human disturbance."

The scoring process

The process of determining Coefficients of Conservation for a region's flora is the single-most critical step in the development of Floristic Quality Assessment (Wilhelm and Ladd 1988). Despite this primacy, there are no codified or formalized protocols for how species *C*-values are assigned. No central body oversees scoring among states or regions, nor are there assurances for how, or if, scores are peer-reviewed. Species scoring is dependent on: the knowledge of botanists, how botanists are directed to score species (e.g., which criteria, see *Species Conservatism*), and the process itself.

The knowledge of the expert botanists consulted— those with experience and knowledge of where the plant species in the region occur— has never been questioned. Where gaps exist in their knowledge of particular species, other data sources are sometimes consulted, for example, descriptions of habitat preferences in local floras, herbarium specimen labels, published site species lists, and scores from neighboring states (Wilhelm and Ladd 1988, Andreas and Lichvar 1995, Andreas et al. 2004, Milburn et al. 2007). FQA in certain regions has been limited by the absence of a complete state flora list (e.g., Mississippi, personal communication Ervin). In such cases, only certain habitats' plants (e.g., wetlands) or regions in the state may be scored (Figure A.1).

The process of botanist consultation has most often had the botanists convene in person and agree on a Coefficients of Conservatism value for each species, with the help of a moderator, until the flora is complete. A couple of issues commonly arise during these sessions that must be managed by the facilitator (Gerould Wilhelm personal communication). First, a certain level of ego and pride in botanists must be overcome, so that they are willing to admit if they simply are not familiar enough with a taxon to give it an informed rating. Not every botanist will know or have familiarity with the distributions of every taxa. A second and more frequent task for the facilitator is to keep the group's understanding of the concept of Conservatism on track, as people will frequently want to use or incorporate criteria and concepts other than remnant fidelity as a basis for assigning scores (e.g., rarity, showiness, size, ecological amplitudes; see *Species Conservatism*).

In-person group discussion and consensus are the most common scoring method (Andreas and Lichvar 1995, Taft et al. 1997, Mushet et al. 2002, Rothrock and Homoya 2005). But they have been criticized because they may allow undue influence from dominant/assertive committee personalities (Landi and Chiarucci 2010). As an alternative, Cohen et al. (2004) had botanists assign scores independently, as opposed to by committee in person, they then combined the scores into a final value. This independent scoring process also simplified logistics by allowing each botanist to work at their own location, and at their own pace. Results from independent scoring also afforded the authors an opportunity to quantify expert disagreement among the species scores and to identify difficult to characterize species. Such species could then be focused on for further consultation, using other data sources, references, or botanists, or they could be flagged for future consideration. Forrest (2010) also used the median of independently assigned panel values. But where there was high disagreement on a species, they first used Peirce's criterion to eliminate *C*-value outliers. Peirce's criterion was described as "a rigorous method based on probability theory that can be used to eliminate data outliers or spurious data in a rational way." Milburn (2007) and Mortellaro et al. (2012) help to focus attention on difficult to characterize species by having their initial botanist panel assign a confidence level in each species' score so that species in need of further review could be prioritized based on their uncertainty.

While there are clear advantages to independently assigning C-values, Cohen et al. (2004) also conceded that "interactive group meetings confer certain important advantages" For example, scores must be defended and knowledge can be shared to inform a group consensus. Rocchio (2011) used a hybrid approach with one group of botanists working individually on scores, and another group meeting in person to deliberate:

> "A panel of botanical experts with field-based knowledge of the region and its flora is assembled. The panel convened for a one day workshop to review the process of how to assign *C*-values. Panel members subsequently assigned (individually) coefficients for those species with which they were familiar. Panel *C*-values were synthesized. A sub-panel (i.e. "Review Panel) of experts reconciled coefficient assignments for species which had wide disagreement across the panel."

This independent assignment process allows for increased deliberation by botanists, helps to avoid undue influence by dominant personalities, and benefits from logistical ease, while it also allows for feedback and argument (at least within the final review panel). At the same time, it puts considerable power within the hands of the few final arbiters on the review panel who can override the larger group of botanists. Mortellaro et al. (2012) used a different hybrid approach, where they first assembled a core team of knowledgeable senior botanists to score species. Then they convened a second group to peer review those assignments. Finally, both the core and review groups met to reach a consensus.

As a final example, Milburn et al. (2007) chose a highly iterative, multi-step process for *C*-value development, where there was five phases of comment and review by 4-5 different groups, before the final scores were settled upon. While this level of review and consideration is certainly welcomed, it would seem prohibitive in most cases. As a way to facilitate more efficient group dynamics within the botanist panel commenting session, while also incorporating the deliberative, iterative process, Cohen et al. (2004) suggested "approaches more commonly employed in the social sciences, such as the Delphi technique (Dalkey and Halmer 1963), to refine CC scores." This method would have each botanist's scores presented by a facilitator to the group ***anonymously***, along with the botanists written arguments for each score. The botanists then go back and review their original scores based on this meeting. This process of independent scoring and written argument, presented to the group anonymously, is then repeated until a consensus in converged upon. Or, alternatively, when the facilitator chooses the median group score where consensus for a species seems unlikely.

In summary, hybrid approaches incorporating individual scoring and group discussion seem a welcome compromise for the *C*-value scoring process that can avoid the weaknesses and incorporate the strengths of both techniques. Furthermore, having botanists identify their confidence with their taxa values or quantifying variation among panelists values are good ways to identify troublesome species for further review. Iterative approaches with multiple reviews of the scores are also suggested. Finally, project facilitators must continually remind botanists to base their scoring on remnant fidelity and to avoid conceptual drift in species Conservatism.

Assigning C-values at state vs. ecoregional levels

Several authors have suggested that physiographic or vegetative boundaries (e.g., ecoregions Bailey 1995) are more logical units by which to assign *C*-values (Cohen et al. 2004, Bourdaghs et al. 2006, Milburn et al. 2007). This may be assumed because, "Plants often have varying C-values in different geographic regions due to physiological and ecological variations within the range of each species." (Rocchio 2007), so they could behave more consistently with respect to their remnant fidelity within ecoregions. This could then allow for *C*-values to be more easily and accurately scored within ecoregions versus states. Bried et al. (2012) have gone so far as to speculate whether stratifying score assignment by both ecoregion ***and*** habitat type would further reduce subjectivity, bias, and improve accuracy.

At the same time, Bried et al. (2013) highlight that "Although statewide averaging may lose ecological precision, it [already] accounts inherently for the varied growing conditions under which a species may be observed… and [ecoregion based designations] themselves are subject to human error, and may not improve FQA." Species ranges are certainly not limited by ecoregional boundaries, and their boundaries may not even represent zones of high species turnover (McDonald et al. 2005). And, just as "State-based rankings...recognize that due to varying ecological tolerance across the species' range, score validity declines as distance from the origin of assignment increases" (Bried et al. 2012), ecoregional scoring would need to do the same.

Furthermore, the practical utility of assigning species *C*-values by political boundaries is clear, it facilitates use by agencies operating under political boundaries. It is also true that many of the professional botanists assigning scores are used to working along political boundaries (due to their employment, granting agencies, etc.), such that they may be more comfortable and have the most experience within them. "Botanists may not have the experience, or recollection of their experience, to assign ecologically specific *C*-values (Bried et al. 2013)."

Thus far, few FQA scorings have followed ecological boundaries, (Alberta in part, Forrest 2010, Chamberlain and Ingram 2012, Louisiana in part, Cretini et al. 2012), and no studies comparing resultant scores or FQA performance between these two types of units exist. Differences between state and ecoregion scores could be so slight that the relative accuracy gains might be meaningless compared to the ease of use afforded by statewide scoring— not to mention that statewide scores are what currently exist. Nonetheless, it has been suggested that future flora scoring panels assign values for both units (Rentch and Anderson 2006). Other plant scoring metrics besides FQA have been applied both ways (although not simultaneously). For example, in Europe Ellenberg values, and in North America, hydric indicator status values have been assigned for both ecological and political boundaries (Reed 1988, Zbigniew 2001, Godefroid and Dana 2007, Lichvar and Kartesz 2012). One potential benefit to ecoregional scoring is that it could help to extend FQA into states that lack *C*-values, but that share the same ecoregional type (Medley and Scozzafava 2009, Bried et al. 2013); for example, Ellenberg scores have been successfully extended beyond their region of origin (Hill et al. 2000, Godefroid and Dana 2007). A minimum first step could be studies comparing *C*-values from select groups of species scored for both types of geographic units to look for evidence of major *C*-value differences, and also to identify any difficulties in the process of scoring species both ways.

The numeric scale in species C-values

Another pitfall in scoring species has been a lack of adherence to the standard 0-10 ordinal scale. For example, it was not clear why, but Nichols et al. (2006) scored species with integers ranging from 1-5. Assumedly, if users wanted place these scores into the standard context they could multiply them by two, but this would leave no species scored as zero or odd numbers. A more frequent mistake is to assume that only non-native species may receive a zero value, this had led many lists for native species to range from 1-10 (Nichols 1999, Hatzenbeler et al. 2004, Bowles and Jones 2006, Ervin et al. 2006, Rentch and Anderson 2006). There is no reason that native species cannot receive zero scores, provided that botanists have no confidence that they will be found in a natural area of any kind (for example see *Lepidium virginicum* discussed in *C-values and disturbance thresholds*).

Standardization, transparency, and publishing

Some FQA developers have described how their *C*-value lists were created in a fair amount of detail and others have not. Indeed, some states have not even published, or made their state score lists publicly available. Thus, one of the failures in FQA development— which is related to the absence of any formalized protocols, requirement, or oversight for scoring— is that there is no procedure for testing, challenging, or modifying a species score after, or before, they are published and in use. Due to the primary importance of species scoring for FQA, in order for it to move forward as a field, the process of scoring species must have as much transparency in how species were scored as possible; both to inform users and increase understanding of FQA, and to allow for some minimum level of quality assurance in scoring or peer-review. Publishing more detailed descriptions of how floras were assigned must occur.

Site-Level Floristic Quality Scores: Usage, Performance, and Properties

In this section the properties and performance of site-level FQA scores are considered. As compared to species values, a site's Floristic Quality, as measured by Mean *C* or FQI, is most simply said to reflect the "aggregate Conservatism of plants inhabiting a site (Miller et al. 2006)." This means that FQA metrics register how biologically degraded a site is due to human disturbances (Taft et al. 1997). Additionally, site Floristic Quality has also come to be described within broader environmental assessment frameworks, and it is now often said to measure an site's Biological Integrity (Mack 2007, Spyreas et al. 2010), or its conservation value (Spyreas and Matthews 2006). Unlike the confusion surrounding Conservatism's definition (see

Terminology & Metrics), these are tenable descriptors. For example, Biological Integrity is largely premised on anthropogenic disturbance degrading habitat health, which makes it a clear conceptual analogy to Floristic Quality (for an exellent description of the difference between Floristic Quality versus Biological Integrity see, Taft et al. 2006). Site Floristic Quality and 'conservation value' are almost synonymous because undegraded habitats are so rare in most modern landscapes that Conservative floras are invaluable for conservation purposes. To summarize, Floristic Quality can rightfully be said to measure a site's aggregate Conservatism, it's human disturbance legacy, it's conservation value, or it's Biological Integrity.

How has FQA been used?

FQA scores were originally suggested for use in: 1) identifying valuable natural areas, 2) comparing Floristic Quality levels among or within sites, 3) monitoring Floristic Quality over time, and 4) determining restoration and management effectiveness (Swink and Wilhelm 1994a). We present examples of these uses, as well as some others that FQA has expanded to include.

Remnant habitat monitoring

Natural areas monitoring with FQA is notable for the variety of habitat types, landscapes, management regimes, and time-frames over which it has been used. For example, monitoring of a high-quality native sedge meadow showed drastic biological degradation to it over a relatively short period, this was putatively due to urban and agricultural sedimentation. Upon further investigation of the causal factors, the ultimate cause of degraded Floristic Quality were found to be losses in soil micro-topography and subsequent invasion by exotic species (Werner and Zedler 2002). Kowalski and Wilcox (2003) sampled wetlands upstream and downstream from a river dam years after the dam had been created and found no difference in either habitat's Floristic Quality; they suggested that the dam was functioning similarly to "an enormous beaver dam" in its effects on the upstream riverine habitats (a natural disturbance). Fiedler and Landis (2012) found lower Mean *C*-values in a prairie fen invaded by an exotic shrub, as compared to the uninvaded areas in the fen. Nichols (2001) re-sampled more than 130 freshwater lake floras after several decades time (up to 82 years), and found FQA increases in some lakes (suggesting recovery) and decreases in others. The habitat stressors varied by region and the lake's water source (stream-fed or spring-fed), suggesting that disturbances and the region or type of lake interacted with Floristic Quality changes. Another long-term comparison found that Midwestern North American dry prairies lost Floristic Quality over 55-years. Although specific causal factors were not directly tested for, they were assumed to be "reductions in patch size and changes in the surrounding landscapes as well as fire-suppression and invasion by taller and woody plants (Kraszewski and Waller 2008)." The longest study of this type looked at the effects of urbanization on the flora of Milwaukee County, Wisconsin (USA). Using flora records going back to 1850, they found that the county had lost at least 36% of its original richness. More importantly, they found a near perfect relationship between likelihood of being extirpated from the county and increasing *C*-value. Thus, the remaining flora was dominated by weedy, common natives and showed a "county-wide" Mean *C* decline (Leitner et al. 2008). Finally, IBI's (Indices of Biotic Integrity) frequently use FQA for monitoring and habitat assessment (see, *Site FQA scores measure human disturbance and biological degradation*).

Restoration monitoring

Habitat restoration projects (including habitat re-creations *de novo*) commonly use FQA to monitor their progress or success (Matthews and Endress 2008). Monitoring restorations with FQA has yielded insights into both the field of restoration ecology (e.g., Matthews et al. 2009a), and the performance and limitations of FQA metrics (Matthews et al. 2009b). Most often studies have compared restorations of different ages with one another; and older restorations have tended to have higher Floristic Quality, suggesting a positive long-term relationship between Floristic Quality and time in restorations (Mushet et al. 2002, Balcombe et al. 2005, Taft et al. 2006). However, studies continuously tracking restorations show a more complicated relationship. Restoration time-series show that trends (or trajectories) are highly unpredictable, at least over the timeframes studied thus far (i.e., the first 15 years). Some sites consistently increase in FQA scores, some increase and then decline precipitously (often in association with rapid invasion by exotic species or flooding events), and others remain constant at low levels (Matthews et al. 2009b). Furthermore, scores in some restorations show high volatility, while others are relatively stable or consistent in their trends (McIndoe et al. 2008, Matthews et al. 2009b). For monitoring purposes, score volatility in young restorations can be dampened by calculating scores at quadrat, rather than site or transect, levels (McIndoe et al. 2008). Restoration score unpredictability seems to lie in stark contrast to consistent FQA trends observed during unmanipulated succession (Rothrock et al. 2011, Spyreas et al. 2012).

Using FQA to track restorations is particularly common for legally mandated wetland mitigation monitoring (Matthews and Endress 2008). The unpredictable trajectory of Floristic Quality in restorations has significant implications for the policies governing this work. For example, such restorations are typically mandated for monitoring over 5-years, after which they are judged as failures or successes. The assumption is that if successful after 5-years, FQA scores will continue to increase or stabilize. However, the evidence shows that scores in nascent restorations cannot be extrapolated into the future. For wetlands in particular, restoration scores are tenuous at any point. Unfortunately, this means that many restorations considered successful based on early Floristic Quality judgments would probably be failures if revisited today (Matthews and Spyreas 2010).

FQA wetland restoration research has also been useful in informing debates as to the best ecological metrics for determining restoration success. There are some measures that are commonly used in restoration monitoring, such as species richness, that almost always indicate restoration success even when sites are obvious ecological failures. Conversely, measures such as native species dominance, native species density, and FQA are far more discriminating and useful (Matthews and Endress 2008, Matthews et al. 2009b). FQA measures may be more informative for long-term monitoring because their terminal values in restorations rarely achieve levels found in undisturbed, remnant, natural areas (Bowles and Jones 2006, Jog et al. 2006, Taft et al. 2006). FQA measures can nearly always differentiate a restoration from a habitat remnant, unlike many other widely employed metrics (Hossler et al. 2011) (see *Restoration versus remnant)* Critics of FQA have pointed out that for legally mandated uses, FQA numbers are vulnerable to manipulation by planting a few highly Conservative species to artificially inflate restoration scores over the short-term monitoring window. The abuse in this case would be planting species that are inappropriate for the site, region, or habitat, or that will almost certainly not persist over the long term. While such malfeasance has not been demonstrated or reported, it should certainly be a concern. Despite advantages of FQA for restoration monitoring, its use is still controversial in some governmental agencies, while it is readily adopted in others (Personal communication Jeff Matthews, Personal communication Brook Herman).

Identifying, ranking, and assessing natural areas

Another intended use for FQA was to provide a tool for experts and non-experts alike to be able to dispassionately identify and rank existing natural areas for their native biological integrity and regional conservation value (Swink and Wilhelm 1994c). To this end, FQA scores have been used for land acquisition and protection decisions. For example, for over two decades in Illinois (USA), site FQA scores have been included in proposals to confer permanent legal protection on nature preserves. Site scores provide quantitative, scientific support that an area is unique and valuable enough to require legal protection. Rocchio (2007) listed organizations ranging from state agencies like the Missouri Department of Conservation and the Illinois Department of Transportation, to county governments, to international conservation organizations such as The Nature Conservancy that use FQA indices to "prioritize high quality natural areas to include in their natural areas networks". In some cases, legal requirements for FQA can be remarkably strict; for example consider the regulation described in Chu and Molano-Flores (2013):

> "the Kane County Stormwater Ordinance in Illinois…requires the protection of wetlands in a development in such a way that the Floristic Quality Index (FQI), an index based on plant species richness and coefficient of conservatism values, calculated two years after development cannot be more than **two points** less than the original FQI. If the wetland has a lower FQI, the developer must provide wetland mitigation for the impacted wetland"

Perhaps the most widespread formal FQA use is by the U.S. Army Corps of Engineers. Beyond restoration monitoring and performance standards mentioned previously (*Restoration monitoring*), its use is legally mandated in several jurisdictional regions to assess the value of wetlands that will be destroyed by development, and to determine the extent to which destroyed wetlands must be mitigated (replaced) (Streever 1999, Herman et al. 2001, Matthews and Endress 2008). This includes hundreds of hectares in compensatory wetland restorations that the U.S. Army Corps of Engineers oversees (Herman et al. 1997, Hough and Robertson 2009). Jurisdictional districts may use strict FQI cut-off values (e.g., wetlands with FQI values <20) for deciding the area of replacement compensatory wetlands required to mitigate wetland destruction. Another federal agency use is in the U.S. Fish and Wildlife Service's Comprehensive Conservation Plan (CCP), which dictates the management and administration of wildlife refuges. They have dictated using FQA scores to set habitat targets, monitor remnants and restorations, and to guide research (U.S.F.W.S. 2013). It is also likely that the U.S. Environmental Protection Agency (EPA) will eventually use FQA for nationwide wetlands monitoring and assessment, because of its "proven usefulness as a biological indicator and consistent effectiveness compared to other metrics" (U.S.E.P.A. 2002a, Medley and Scozzafava 2009, U.S.E.P.A. 2010).

Some authors have also looked at the Floristic Quality of different areas within a single site to compare their relative quality and disturbance (e.g., Edgin et al. 2010). When scores are compared within a site, their meaning relative to one another is obvious— one area is better than the other. However, it is far more common to rank scores from different sites against one another. As is discussed later in this review, users should carefully consider the relative meaning of FQA score comparisons among different sites (especially ones of different habitat-types, regions, sample seasons, area, and sample intensity). For example, there is a growing trend to include FQA scores in published floristic inventories. This provides a coarse indication of an area's conservation value for the reader. Publishing raw scores in this way without commenting on their relative meaning or context suggests that *any* FQA score can be seen as absolute value, which could lead to misinterpretation. This also highlights that there is a research need to summarize the range and expected natural variability in FQA scores within states. Providing reference Floristic Quality values and percentile rankings to scores by region and habitat type would allow users to put scores into context (see *Precision, sensitivity, and variability in site scores*).

Basic research

Floristic Quality was not an originally suggested use for in fundamental research, perhaps because there are no ecological theories explicitly underlying it. However, FQA has increasingly been used for ecological and conservation based research. For example, Spyreas and Matthews (2006) used *C*-values to judge the degree to which Conservative plants exhibited nested distribution patterns across forest habitat patches. In the same way that island species are often nested due to extinction-colonization dynamics according to island size and isolation, human disturbance to forests has sorted plants across the landscape based on their *C*-values. Burke and Nol (1998) found that FQA values increased over forest-edge to interior gradients, suggesting their sensitivity to edge effects. Similarly, the presence of highly Conservative forb species was the only vegetation measure that was capable of separating linear from block shaped (low edge- to- interior ratio) remnant prairies from one another (Davis et al. 2008). Matthews et al. (2009a) assessed the relative importance of local versus landscape level abiotic factors in predicting restoration success using FQA. McNicoll and Augspurger (2010) found that prairie vegetation's Floristic Quality was higher above ground than in the seed bank, suggesting that prairie seed banks are reservoirs for a lesser subset of a community's Floristic Quality. Both Tulbure et al. (2007) and Spyreas et al. (2010) used FQA to assess exotic plant invasion's impact on native wetland floras. Laroche et al. (2012) assessed the utility of using white fringed orchid (*Platanthera leucophea*) presence as an indicator of sphagnum bog ecological integrity. While Mean *C* was one of the best of several dozen environmental variables at explaining this "indicator species'" presence, users should avoid the mistake of associating FQA with species indicators (Ejrnæs et al. 2008), as neither FQA values nor individual specie *C*-values were intended to be used in this way.

Studies have also used FQA to study multi-taxa group ecology, often across trophic levels. Cline et al. (2008) used FQA as a measure of a bio-control program's success, where Conservative plants became more abundant after a herbivorous introduced beetle had controlled a non-native invasive plant. Panzer and Schwartz (1998) found that FQA did not explain insect richness, rare insect occurrence, or presence of conservative insects in prairies. O'Neal et al. (2008) found that site Floristic Quality explained waterbird and waterfowl richness, density, habitat-use, or reproductive activity. The association of FQA with non-plant taxa is largely unexplored and justifies further research. Finally, FQA has also been used in a non-biological context as a way to quantify economic costs associated with ecosystem function in wetland restorations (Gutrich and Hitzhusen 2004).

Assessing ecological management

There is a well-established history of using FQA scores to measure and monitor the effectiveness of natural areas management. Many North American woodlands experience anthropogenic fire suppression, and the behavior of understory FQA scores reported after fire re-establishment has varied. Jackson (2009) found no change in FQI over a nine-year fire treatment in an oak-savanna. He attributed the lack of an effect to the fire’s mistimed seasonal application and inadequate intensity, which failed to control invasive exotic species and abundant woody plant re-sprouts. Taft (1999) found increases in both FQA metrics over a six-year fire treatment period, while a similar five-year woodland study found that only FQI scores increased (Wilhelm and Masters 1994). Fire caused both weedy and Conservative species to emerge from the seed bank in these studies to different extents. But, the FQI-only response in the second study occurred because both Conservative and non-Conservative species increased in numbers (increasing overall richness). The weedy species in this second study were expected to decline with continued management, and Mean *C* scores were then expected to increase in it as well. These examples also highlight that Mean *C* is sometimes seen as the less responsive metric when assessing short-term treatments, compared to FQI, but it could also be thought of as be being the more conservative on as well by being less sensitive to short-term fluctuations (see also, *Creating FQA scores from quadrat, plot, or site level sampling*).

FQA has also been used to assess fire management in grasslands. Bowles and Jones (2006) studied 33 remnant prairies over a 22-year period and found that the number of times a site had been burned correlated with its FQI, but not Mean *C*. Although, Mean *C* was negatively correlated with woody plant enchroachment (presumably due to their shading out prairie plants). Similarly, Rooney and Leach (2010) and Milbauer and Leach (2007) both found strong evidence for historic fire suppresion effects on prairie Floristic Quality. Finally, a six-year burn cycle treatment was found not to be frequent enough to maintain prairie Floristic Quality, compared to a two-year cycle that caused FQA increases over 16 years (Masters 1997).

Several studies have compared planting and site preparation methods in restorations using FQA. They have examined seed sowing rates (Goldblum et al. 2013), seed mix species ratios (McIndoe et al. 2008), and various planting techniques (Packard 1994, Foster et al. 2007, Middleton et al. 2010, Carter and Blair 2012). Conservative species were shown to be more successful, while non-Conservative species were less successful, when prairie restorations were inoculated with remnant prairie soils. This suggests a strong relationship between soil biotas and species Conservatism— perhaps between beneficial myrchorhizal fungi and Conservative taxa and inhibitory pathogens and non-Conservative species. Finally, nitrogen enrichment decreased and nitrogen sequestration increased FQA scores in a grassland restoration (McIndoe et al. 2008).

Another topic of interest has been the effect of wildlife and livestock presence on FQA. In tallgrass prairies where white-tail deer have become over-abundant due an absence of predators, FQA scores were shown to be higher in deer exclosures or where Conservative forb species were less browsed with decreased deer numbers (Anderson et al. 2006, Anderson et al. 2007). South Dakota and Minnesota (USA) pastures and grasslands managed for wildlife habitat were found to have lower Floristic Quality than areas managed as nature preserves (Smart et al. 2011). Finally, Brudvig et al. (2007) were not surprised that neither fire management nor light cattle grazing did not affect FQA in a tallgrass prairie due to the short treatment duration, although there were effects recorded from other vegetation measures.

Site FQA scores measure human disturbance and biological degradation

The most fundamental question concerning the validity of FQA, and one that is integral to nearly every use, is whether its metrics accurately and consistently measure its biological quality, or conversely the amount of biological degradation a habitat has incurred due to human disturbances and stressors. To be sure, the presence/absence of any single Conservative species does not provide enough resolution to do this (Stapanian et al. 2013). This section reviews how well *site-level* FQA scores work. To this point, in nearly every instance reported scores have been found to be highly predictive of both specific individual disturbances when they have been examined (Table A.1), and broader, more inclusive site human disturbance legacies.

Typical FQA users seek to separate degraded natural areas from those deemed high-quality. In order to test how well FQA does this, as with a simple correlation, there must be ways to rank sites *a priori* by what can be a rather vague criterion (this is sometimes referred to as the X-axis problem of determining site biological quality). Early studies testing the effectiveness of FQA metrics ranked sites based on qualitative judgments of their perceived quality. And, FQA scores generally corroborated these expert opinion based rankings (DeKeyser et al. 2003) Tests of ecological metrics and indicators have increasingly used more "objective" rankings of biological quality. To do this, more comprehensive site indices that quantify and summarize broad suites of human disturbances and stressors into a single value have been created. FQA has been found to accurately predict these quantitative habitat-quality/disturbance indices as well (Fennessy et al. 1988, Fennessy et al. 1998, Lopez and Fennessy 2002, Cohen et al. 2004, Herman 2005, Ervin et al. 2006, Miller and Wardrop 2006, Miller et al. 2006, Rocchio 2007) .

Similar powerful tests of FQA performance have come from studies with the primary objective of creating an Index of Biotic Integrity (IBI)— essentially a multi-metric tool for assessing habitat Biotic Integrity. To create an IBI, dozens of biological variables across dozens or even hundreds of sites in a region are screened for their correlation with a multivariate human disturbance gradient. The best predictive variables are then used to create a single predictive equation that is used as the IBI. Where the candidate list of biological variables for the IBI creation has included FQA metrics, two points have become apparent. First, FQA metrics nearly always have the highest correlation with the human disturbance gradient, among the dozens of potential biological variables tested (Wilcox et al. 2002, Mack 2006, Miller et al. 2006, Reiss 2006, Mack 2007, Mack et al. 2008, Raab and Bayley 2012, Wilson and Bayley 2012). Given that the goal of IBI's are to gauge effects from human stressors and disturbances to habitats (e.g., pollution, surrounding land use) with biological indicators, it is not surprising that FQA measures would perform well. But, their consistently superior performance is notable. For example, FQA metrics outperformed 50 other measures in assessing wetlands (Miller et al. 2006). The second point is that FQA metrics are especially valuable because they tend to add *unique* information to the IBI— information that is unlike that contributed by the other biological variables tested (e.g., as tested with redundancy analysis, Wilson and Bayley 2012). For these reasons, when they have been tested, FQA metrics are almost always included in final IBI equations.

Other studies provide indirect evidence for site FQA score effectiveness. For example, many single disturbance types have been shown to decrease site FQA on their own (Table A.1). Studies monitoring sites over time show that previously stable site scores quickly decline when new human disturbances occur. Alternately, where ecological management is implemented, site Floristic Quality has been shown to stabilize or increase (see *Assessing ecological management*). FQA metrics are also able to separate remnant habitats from created (*de novo*) restorations (but see, Allison 2002, Mushet et al. 2002, Andreas et al. 2004, Taft et al. 2006, Hossler et al. 2011), which reflects that these are newly created, and therefore, recently disturbed habitats. Finally, beyond site-level scores, declines in the Floristic Quality of regional floras due urbanization and land-use change causing Conservative species extirpations have also been demonstrated (Leitner et al. 2008).

In summary, the most difficult task in testing how well FQA works is objectively quantifying site biological quality, site degradation, and site disturbance legacies against which FQA scores may be tested. Evidence from multiple criteria, however, indicates that FQA consistently does what is purports to do, proving FQA metrics to be “remarkably robust” (Miller et al. 2006).

Table A.1. Examples of specific individual anthropogenic disturbances to which site FQA scores have been demonstrated to respond.

Forests/Woodlands	**Grassland**	**Wetland**	**Aquatic**
Logging (Francis et al. 2000)	Intensive herbicide use (Smart et al. 2011)	Intensive livestock grazing (Bowers and Boutin 2008, Boughton et al. 2010)	Pollutants and water quality degradation (Nichols 1999)
Pollutants and water quality degradation (forested wetlands) (Reiss 2006)	Intensive livestock grazing in tallgrass prairie (Smart et al. 2011)	Nutrient enrichment/pollution (Fennessy et al. 1998, Malik et al. 2012)	Decreased habitat buffer (Nichols 1999)
Livestock grazing (Gerken et al. 2013)	Nutrient enrichment (McIndoe et al. 2008)	Isolation (Fennessy et al. 1998).	
Surrounding land use urban-Urbanization (Gerken et al. 2013)		Habitat buffer (Fennessy et al. 1998, Kercher et al. 2004, Ervin et al. 2006, Miller and Wardrop 2006, Reiss 2006, Malik et al. 2012)	
		Pollutants and water quality degradation (Kercher et al. 2004, Malik et al. 2012);	
		Intensity and amount of surrounding anthropogenic land use (Herman 2005, Ervin et al. 2006, Miller and Wardrop 2006, Malik et al. 2012)	
		Population density surrounding (Bourdaghs et al. 2006)	
		Surrounding agricultural land use (Bourdaghs et al. 2006)	
		Hydrological modification – based on surrounding land use (Bourdaghs et al. 2006)	

Scores over time

The temporal dynamics in conservation metrics are often insufficiently understood, leading to their misuse (Niemi and McDonald 2004, Fleishman et al. 2006). FQA measures are no exception, as incorrect accounting for temporal trends in FQA scores may compromise research (Spyreas et al. 2012). This section considers short- trends in FQA scores. Long-term trends are addressed in Chapter 3 and in Introduction: *Understanding long-term successional trends in FQA and comparing site-scores over time*. Readers are also referred to *Seasonal variability in FQA sampling* for the relevance of intra-annual trends to FQA.

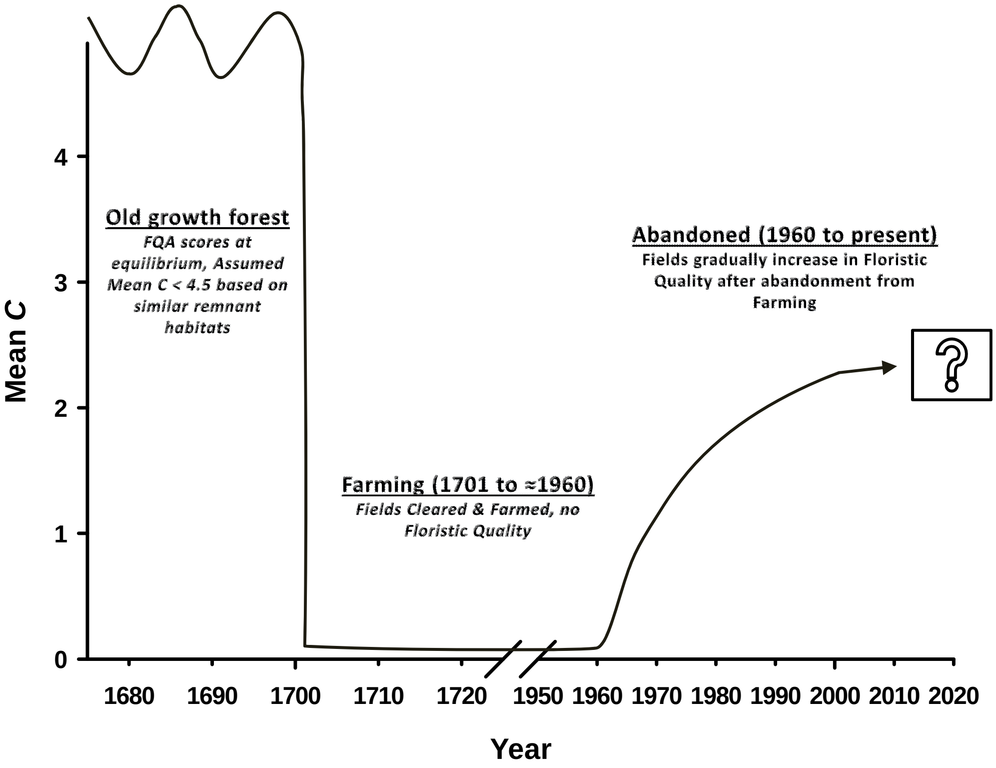


Figure A.4. Long-term trend in site Floristic Quality for a former forest that was converted to farmland, was then abandoned, and is now undergoing old-field succession in New Jersey (USA). Values over the past 50-years are based on Spyreas et al. (2012).

Short-term cyclic or inter-annual "nuisance variation"

The accuracy of FQA may be vulnerable to large inter-annual variation in sites. This is especially a concern for habitats that, despite being established remnants, frequently incur large (often re-occurring or cyclic) natural disturbances. For example, Wilcox et al. (2002) suggested that the year-to-year variability in lake levels that produces extreme changes in wetland plant communities bordering the Great Lakes (North America) could generate unacceptably large inter-annual variation in their FQA scores. Because such variation is unrelated to human disturbance, this would be a large inherent inaccuracy in FQA, although they did not test whether this resulted in FQA effects. Euliss and Mushet (2011) studied this question in North American "Prairie Pothole" wetlands. In certain zones in these wetlands, "Climate cycles result in temporal shifts between Conservative emergent, submerged, and free-floating species to less conservative ephemeral species, many of which are annuals." This led to yearly variation in Floristic Quality that was so large that it made it difficult to separate sites of different human disturbance levels from one another.

Such a finding raises serious concerns about using FQA, but inter-annual variation in other established habitats is frequently found to be minimal. For example, Herman (2005) found that average Mean *C* scores did not differ in wetlands sampled in consecutive years (see also, Lopez and Fennessy 2002). Furthermore, studies tracking Floristic Quality changes after (natural) disturbance from fire do not show dramatic yearly variation, beyond fires tendency to precipitate a general trend towards increasing scores (Wilhelm and Masters 1994). In prairies, Smart et al. (2011) argue that FQA scores are stable with proper ecological management as "most prairie plants are perennials and are observed during both wet and dry years." In some Great Lakes coastal wetlands, the natural flood regime and water fluctuations have actually been seen as vital for encouraging Floristic Quality *recovery* in disturbed wetlands. In this case, Conservative native plants were seen as requiring these historic disturbances to flourish. "The Saginaw Bay water levels are in constant flux, alternately inundating and exposing the coastal wet meadow zone…Certain grasses, sedges, and forbs have successfully adapted to this variable hydroperiod, allowing them to dominate wet meadow vegetation assemblages" (Stanley et al. 2005). Conversely, effects from lake-level fluctuations are described as being so severe in other Great Lakes wetlands that they regularly eliminate the remnant native community, and may allow opportunities for invasive exotic species to establish and eventually dominate, although an effect on FQA values in such instances has only been suggested and have yet to be the demonstrated (Tulbure et al. 2007).

Beyond cyclical or re-occurring cycles observed in remnant habitats such as these, the dramatic inter-annual FQA variation in very young, compositionally volatile, early-successional habitats is also a potential inconsistency to consider (McIndoe et al. 2008, Matthews et al. 2009b). Despite this short-term volatility, such habitats are typically undergoing an overall successional trend that is directional rather than cyclic, and such volatility becomes relatively minor after the earliest successional stages (Spyreas et al. 2012). In other words, this seems to be noise variation that is largely unavoidable for FQA users in young habitats. Clearly, users of an ecological indicator must be cognizant of whether a habitat was very recently disturbed (naturally or otherwise) and has not stabilized. And, if a site's longer-term trend is of interest, this will help to avoid over-emphasizing scores generated during volatile periods during, or soon after, large natural disturbances. Wilcox (2002) suggested that sampling some volatile habitats might "require differing [temporal] scales of measurement for years that differ in the length of time since the last major natural disturbance. " Thus, averaging scores from multiple years, perhaps before and after a disturbance, could be useful. Or, if before data is not available, sampling for at least two years to see if scores are unstable could be sufficient to screen for rapid post-disturbance changes.

If more sampling across years is not an option, it may often be possible to ameliorate volatility by sampling across the disturbance gradient in a site. For example, in the emergent wetlands example above, it has been pointed out that rather than calculating scores based on samples within discrete hydrologic zones (i.e., concentric rings in the emergent, submerged, floating plant zones surrounding circular wetlands, which are highly labile from year to year because of droughts, Nichols 1999, Euliss and Mushet 2011), by aligning sample transects perpendicular to the hydrologic gradient, scores spanning the upland to open-water zone would yield a fuller representation of the wetland and could moderate temporal volatility (Hargiss et al. 2007). However, this point has been debated:

> "(Hargiss et al. 2007) suggested that including information from upland communities in wetland plant community assessment methodologies can account for shifts of species between zones caused by hydrologic fluctuations and provide the ability to separate changes due to normal hydrologic change from changes due to human disturbance. However, studies have shown that significant changes in plant communities occur in wetlands as they cycle through hydrologic fluctuations that cannot be attributed to simple shifts in the spatial location of species into uplands or vice versa (e.g., Harris and Marshall 1963; van der Valk and Davis 1978; Welling et al. 1988; Euliss et al. 2004). For example, during dry periods, annual plants colonize exposed mudflats following germination from dormant seed banks in the wetland substrates and replace submersed and free-floating perennials in wetlands (van der Valk and Davis 1978; van der Valk 2005). This temporal pattern is a natural disturbance pattern in which annual plants that typically have low C-values periodically replace perennial species that are generally more conservative and have correspondingly higher C-values…Plants tolerant of moderately brackish or brackish semi-permanent wetland subclasses typically have high average C-values. During natural drawdowns, these species are replaced by species that have significantly lower C-values."

A different study on emergent wetlands across Florida (USA) generated Mean *C*-value calculated from inner, intermediate, or outer wetland zones and found that they correlated very highly with the Mean *C* calculated from the entire wetland's flora (r = 0.93, 0.74, 0.99, respectively) (Murray-Hudson et al. 2012). This may suggest that sampling any one hydrologic zone will represent the whole wetland's FQA score well, and placing the sample transects perpendicular to the hydrologic gradient may not be needed to solve the problem. However, this was a single year study, which was not necessarily conducted during a drought year. Therefore, water table fluctuation and weedy plant emergence in exposed zones may not have occurred. Until perpendicular and parallel sample techniques are compared directly in drought and non-drought years, the question of whether FQA score volatility could be moderated to acceptable levels in drought prone wetlands will remain unsettled. In addition, further study is needed to determine whether FQA scores vary annually in other naturally disturbance-prone habitats, such as in large river floodplains and tidal zones. Furthermore, this inquiry could be extended even further to include less-regular or cyclic habitat disturbances such as extreme drought, hurricanes, wind-storms, ice damage, herbivory-grazing, or saltwater intrusions.

Using FQA: Considerations of Methodology, Sampling, and Metric Properties

Using Mean C versus FQI

In studies comparing the ability of Mean *C* and FQI to measure habitat degradation and conservation value, some show Mean *C* to be the better predictor (e.g., Fennessy and Roehrs 1997, Francis et al. 2000, Rooney and Rogers 2002, Cohen et al. 2004, Herman 2005, Miller and Wardrop 2006, Bowers and Boutin 2008, Bried et al. 2013), while others have found FQI to be the stronger measure (e.g., Bourdaghs et al. 2006, Bowles and Jones 2006, Spyreas et al. 2010). In some cases, studies find that there is no difference between the two (e.g., Jog et al. 2006), although they both generally perform well (*Site FQA scores measure human disturbance and biological degradation*). Beyond their putative correlations with biological degradation, there are other performance factors to consider when choosing between FQA metrics. For example, while Mean *C* values might not respond as quickly as FQI values to ecological management or other site changes (see, *Assessing ecological management*), it has several other potentially desirable properties to distinguish it from Mean *C* (see *Species richness and area effects*). This section reviews the major usage considerations when using FQA.

Species richness and area effects

FQI

FQI has a species richness parameter directly in its calculation, while Mean *C* does not, making FQI more prone to differences in richness among sample areas. Sample area effects can lead to unintuitive FQI results in several ways. For example, a disturbed site that is species-rich can compare favorably in its FQI score with one that is less diverse, but pristine and undisturbed (e.g., Swink and Wilhelm 1994b, Francis et al. 2000, Rooney and Rogers 2002). This type of result may be seen as a flaw in the metrics (i.e., a more disturbed site scoring comparably with a less disturbed one), but users should understand that any metric that incorporates species richness will have this type of disturbance-richness relationship because richness is usually highest at intermediate disturbance levels (Fleishman et al. 2006). If disturbance to a site is discrete, localized, and not typical of the overall site, users might also consider excluding such an area from the sample to mitigate its effect on the FQA score.

Almost all biological indicators respond differently as a function of spatial scale (Niemi and McDonald 2004). In addition to the influence of disturbance on richness, FQA users must also consider the effects of scale on richness (especially for FQI). Both the size of the habitat patch, and the size of the sampled area in the patch may warrant attention. FQI generally exhibits a positive relationship with sample area because richness increases with area (Francis et al. 2000, Rooney and Rogers 2002, Matthews 2003, Matthews et al. 2005, Bourdaghs et al. 2006). The only circumstance where FQI values might not increase as more area is sampled is if the added area is highly disturbed. In this case, increased sample richness could come from weedy, non-Conservative species that lower the Mean *C* parameter value in the FQI equation enough to counterbalance the higher richness parameter value, leading to an overall decrease in the FQI value (but see, Miller and Wardrop 2006, Rocchio 2007). Nonetheless, in practice, the scenario of increasing sample area and decreasing FQI scores has only rarely been observed (Miller and Wardrop 2006).

Area-richness effects (usually positive) must be considered whenever comparing FQI site scores so that sample-area effects are not attributed to a habitat's character or biological quality. For this reason, FQA users should always report the size of the sampled area and how it was sampled. Because sample intensity/area is so infrequently equivalent across study sites, comparing *FQI values* is often not the best option for ecological studies. One option is to sub-sample sites and average FQI values per quadrat or sample plot, provided that the quadrats are the same size across sties, this would allow for sample area and richness to be controlled for and for FQI scores to be compared across sites (see, *Creating FQA scores from quadrat, plot, or site level sampling*). FQI values created from whole site lists can be the more useful metric when the cumulative or total conservation value of a site is of interest, rather than a per unit area assessment. For example, when considering a habitat parcel for acquisition or protection, it's species Conservatism, richness, and area can all be evaluated with a single value by using the FQI value of the whole habitat patch.

Mean *C*

Even with the option for quadrat-averaged FQI scoring, Mean *C* is still the generally preferred metric for among-habitat comparisons because of area-richness effects (Taft et al. 1997, Francis et al. 2000, Rooney and Rogers 2002, Matthews 2003, Matthews et al. 2005). Even so, the generally accepted assumption that site Mean *C* is independent of differences in diversity, sample area, or sample intensity (because it does not incorporate species richness directly into its calculation), must be verified if unbiased site comparisons are to be confidently made. Chu and Molano-Flores (2013) found that *both* FQI and Mean

C increased with sampled area in 12 urban wetlands. However, this study is somewhat anomalous in finding this relationship. For example, Francis, Austen et al. (2000) found that Mean *C* did not change as sample plot size increased in forests. Rooney and Rogers (2002) found that Mean *C* did not change as more quadrats were sampled in a forest, and sampled area, richness, and spatial coverage of the site all increased (see also, Herman 2005).

One study looked at Mean *C* and area at two scales (Bourdaghs et al. 2006). They compared Mean *C* scores generated from quadrats along a transect to those from a species list generated by a larger, single plot encompassing the transect, and found no difference between the two sampled areas in Mean *C*. Their second test compared Mean *C* scores across nested plots ranging from 16 m^2 to 0.25 m^2. Mean *C* did not change with plot size in two of three wetland habitat types tested, whereas it was *negatively* related to area in a third habitat-type tested. The authors suggested that this somewhat unexpected result was attributable to this rather unique habitat type (fen). Taft et al. (2006) have suggested that the occasional findings of an area effect on Mean *C* are probably attributable to comparing sites at the extreme ranges in patch sizes, where results are "driven by [those] very small sites that are less likely to sustain some conservative species." Thus, they conclude that when comparing sites that do not dramatically differ in sample area size and at the spatial scales typical of community-level studies, "Mean *C* is safely assumed…an area-independent metric." (Taft et al. 2006)

In addition to considering *sample area* effects on Mean *C*, the area-richness-Mean *C* relationship might also be affected by the size of the *habitat patch* itself. In other words, larger habitat patches may contain greater relative numbers of Conservative species, or they may contain fewer non-Conservative species. This patch size by Mean *C* relationship has been investigated in two ways; by comparing sites scores where the flora of the entire habitat patch was surveyed (total species list by Mean *C*), or by comparing scores from a constant sample area size (Mean *C* per unit area) among variably-sized habitats. Roccio

(2007) used the latter technique and found that Mean *C* was not correlated with wetland patch size. Spyreas and Matthews (2006) also did not find a correlation between Mean *C* and the size of forest patches, but they did find that Mean *C* was positively correlated with species richness (density) in the sampled areas. This suggests that Mean *C* did not increase as a function of a habitat patch's overall size, but rather because species-rich (species-dense) forests had greater *proportions* of Conservative species. Matthews (2003) found that wetland Mean *C* was not correlated with site richness (density), where wetland area was statistically controlled for. The positive Mean *C* by site richness (density) relationship observed in these two studies with large sample sizes has been attributed to Conservative species occurrences displaying a nested pattern among habitat patches. This suggests that the most Conservative species in a habitat tended to be the first ones eliminated with disturbance. Taft et al. (2006) describes a similar phenomenon as "community disassembly", where species loss due to human disturbance occurs in an orderly fashion with the most Conservative species being lost first. The net effect of such a process would lead to a positive correlation between site richness and Mean *C*, but only if the Conservative species attrition with disturbance was not balanced by colonization of equivalent numbers of less-Conservative species. Alternatively, (Bowles and Jones 2006) found a weak positive Mean *C* by native richness relationship in high-quality prairies. This was, however, probably an artifact of including different prairie types in the regression—where the driest prairies sampled had the very highest Mean *C* values, while also being an inherently less species diverse habitat types. Therefore, comparing dry prairie with wet-mesic prairie likely created this negative slope.

Few studies have examined the Mean *C*-richness-area relationship, where Mean *C* is calculated based on sampling the entire habitat patch's species (e.g., Matthews 2003). Results from studies sampled this way are important as they can yield other unique insights into this relationship. Bourdaghs (2004) found that Mean *C* was not correlated with total richness in any of three different wetland types studied. Another study looked at an entire watershed's 203 wetlands and did not find a Mean *C*-area relationship (Matthews et al.

2005). The subset of 107 of these that were forested, however, did have a positive Mean *C* by patch-size relationship. Assuming that smaller forests were not simply more likely to be directly disturbed (grazing, logging, etc), this area effect may have been attributable to the ecology dynamics of smaller habitats (e.g., fragmentation, edge effects, small population sizes). Indeed, the authors also found that perimeter to area ratios in the wetlands were strongly negatively related to Mean *C*. Strong edge effects, where local Mean *C* and FQI scores increased away from habitat edges bordering anthropogenic habitats, have been shown in other forests (Burke and Nol 1998), and in tallgrass prairie (Taft personal Communication). This suggests that habitat-area by Mean *C* effects could be expected where Conservative species are sensitive to edge effects or are intolerant of reduced habitat patch size.

The critical question for understanding Floristic Quality properties is, where such relationships are due to the effects of habitat fragmentation, edge-effects, etc., are they necessarily anthropogenic effects, or can lower in Mean *C* scores arise in habitats that are naturally smaller and more fragmented. The later could represent a confounding Mean *C* relationship with particular habitat's spatial ecology and landscape context. For example, habitat types that are inherently small and were spatially-discrete in pre-settlement landscapes (e.g., fens) might be naturally prone to lower Mean *C*, even when they are pristine. To this point, there is no evidence that smaller habitats such as fens and bogs are inherently deficient in Mean *C* scores (see Chapter 1), suggesting that where a Mean C by habitat-area relationship has been observed it is attributable to anthropogenic fragmentation effects on habitat patches.

To summarize, FQI is sample area, and therefore, species richness dependent. When choosing between the two main FQA metrics users must decide whether to incorporate diversity (FQI) into their assessment or to exclusively rely on species composition and species Conservatism (Mean *C*). When comparing site scores with one another using FQI, most users will want to control for sample area, sample intensity, and sample method.

Alternatively, they may use Mean *C,* as studies have shown that it tends not to suffer sampling biases in this way. However, because this is such a critical assumption for using Mean *C*, further study is still needed to ensure that it is reliably sample-area neutral. A few studies have found that Mean *C* by patch size effects occur alongside well-known landscape ecology factors (e.g., fragmentation, isolation, edge-effects) (see also, Matthews et al. 2009a). Since these largely originate as anthropogenic disturbances in these contexts, they do not suggest a bias in the performance of Mean *C*. Finally, observations that species-dense habitats have higher Mean *C*-values presumably reflects their retaining their Conservative species over time compared to their loss without replacement from disturbed habitats.

Circumscribing the site to be sampled

The term "site" is typically used to describe an FQA study area. What should, or should not constitute a site, sample-unit, or sample-area when using FQA has never been explicitly detailed. Taft et al. (1997) provides some of the most useful guidance on the topic:

> "Determining the extent and configuration of the survey unit [for FQA] often is not a trivial question. Where the unit of floristic analysis is an isolated habitat fragment, the sample area usually is readily apparent. In landscapes with more contiguous vegetation, however, determining the sample unit is less obvious and in many ways dependent on the questions and interests of the investigation…"

A strict definition of an acceptable FQA site may not be possible, but some parameters can be established. For example, sample data from two or more (discrete) "isolated habitat fragments" should not be combined into a single FQA calculation. A more complex question is how large or small can a site be? Simple guidelines as to a range in acceptable site size would be useful. For example, a "forest stand" is a common forestry term used to describe the appropriate size of a sample-area, which has been described as being between 10-100ha in size. When using FQA, this question has two parts: 1) spatial scales at which FQA metric functionality becomes compromised, and 2) scales at which its use is simply impractical. The size of a site will differ depending on whether the site is to be completely surveyed (assessed as a discrete habitat patch) or sub-sampled. In either case, there is no apparent reason that FQA functionality is compromised as sample area increases. Instead, upper area limits will likely be set by the practicality of surveying increasingly large areas (Taft et al. 1997). One study tracked the Floristic Quality of a county's entire flora over time, and while this is certainly a spatial scale beyond that of a site, it is a valid FQA use (Leitner et al. 2008). At the largest extreme, Floristic Quality scores generated from an entire state's flora are not comparable.

At the small end of the site spectrum, there is typically little interest in complete site floral inventories of the smallest habitat fragments as for their conservation value because they are too small to function as natural areas on their own. For example, while some protected natural areas can certainly be quite small (groundwater seeps habitats can be on the order of dozens of square meters), the floral assemblage of a discrete habitat patch the size of an automobile would draw little interest for its Floristic Quality score for this reason. While FQA is typically not considered practical for use on the smallest sites in instances like these, what the minimum size of sub-sampled habitats is (i.e., sample plots), that could lead to compromised FQA metric functionality, remains an important question warranting consideration (see *Site sub-sampling*).

To summarize, the "site" for FQA is a somewhat intractable term, with few guidelines to help define it. Non-contiguous habitats should not be combined into a single sample score. For whole-site floristic inventories, it is the practicality of sampling very large sites, and the limited usefulness of sampling very small sites, that limits when FQA can be used, as opposed to whether metrics function at these scales.

Site sub-sampling

Sample area and sample intensity

While the sample area for whole site FQA surveys is straightforward and obvious, comparative vegetation studies of all types grapple with the appropriate size or intensity of area needed for adequate site sub-samples (Mueller-Dombois and Ellenberg 1974). Studies tend to employ species accumulation curves to ensure that an areas flora has been adequately characterized, and that it may be compared with other site samples. But, only a relatively small amount of a site's vegetation needs to be inventoried to produce a robust FQA measure of a site (Bourdaghs et al. 2006, Milburn et al. 2007, Medley and Scozzafava 2009). In a woodlands study, Francis et al. (2000) found a significant correlation between FQA scores in 0.39-ha plots nested within 0.78-ha plots, suggesting that these plots are large enough to characterize woodlands floras. Several other studies do not find score differences across nested plots of various sizes (especially for Mean *C*) (reviewed in *Species richness and area effects*). However, Bourdaghs et al. (2006) used species accumulation curves to show that ***FQI*** values in wetlands did not stabilize until about 20 m^2 was sampled (see also, Miller and Wardrop 2006). If species accumulation curves in other temperate habitats are any guide, the sample area needed to confidently characterize ***FQI*** in most other habitats is probably much larger (Peet R.K. et al. 1998, McCarthy 2003). Insuring sample sufficiency with species accumulation curves may be especially important when comparing FQI scores *among different vegetation types*, as different habitat types can accumulate species at different rates (i.e., different species-area relationships) (Matthews 2003). They may also be useful as a way to control for different-sized sample areas when comparing FQI scores.

Achieving adequate site coverage would seem to be less of a concern when using Mean *C* because of its insensitivity to area and richness. Bourdaghs et al. (2006) suggest that complete sample unit inventories are "not necessary for Mean *C*, as index scores are stable with small samples….and the point of diminishing returns is reached with minimal sampling for Mean *C*" (see also, Rothrock and Homoya 2005). Even though Mean *C* does not exhibit an area effect, when there are very few species contributing to a mathematical mean, as was discussed in previous sections, Mean *C* is prone to volatility, instability, and even overestimated values (young habitats, very small sample plots, exotic species invasions, low richness, etc.). Just as generating species accumulation curves would be useful for FQI to determine if an area's richness has been characterized, sample area comparisons would be useful in future studies to determine the sufficient sample area or species richness needed to generate stable and adequate Mean *C* estimates.

In conclusion, Species richness and area effects

FQIFrancis et al. (2000) make the following general recommendations for sample intensity and FQA:

> "…useful information can potentially be gleaned from quite small plots. The "optimal" plot size depends upon the precision of the estimates from different size plots relative to the time required to survey them….[but] it is likely to vary with habitat and the species accumulation curves in that habitat. For a study designed to get good information on a small number of habitat patches, we would recommend attempting to survey many plots within each patch, at a size that allows fairly rapid assessment of the flora. This would simultaneously allow an assessment of the variability within each woodlot and provide species/area accumulation curves. For a study designed to survey a large number of sites, fewer plots per site would be recommended, and a pilot study would be appropriate to estimate the optimal tradeoff between plot size and number of plots that would yield the best precision."

Creating FQA scores from quadrat, plot, or site level sampling

Once the site to be sampled is determined, users must decide whether to sample the whole-site or to sub-sample. Although FQA was originally presented as a technique to assess sites by simply obtaining a complete species list, scores are now just as often generated using systematic, plot-based, site sub-sampling. "FQI and Mean *C* have the added advantage to many diversity indices in that they also can be calculated from both plot-based and plot-less [i.e., whole-site] sample data" (Taft et al. 2006). One obvious benefit to creating scores from plot-based site surveys is that sampling (e.g., effort, area, plot layout) can be controlled to facilitate comparisons among sites. Taft et al. (2006) suggest that, "Sub-sampling, spatially and temporally, is a practical option, particularly where habitat integrity appears relatively uniform and the survey unit is too large to inventory completely within the time available."

However, depending on the size and diversity of the site, sub-sampling can actually take more time and effort than simply generating a site species list (plot set-up time, re-recording species numerous times in each quadrat, etc). Another potential weakness of plot-based sampling, is that important species may be missed and the ability to discriminate fine site Floristic Quality differences could decrease. In most cases, however, FQA performs well when species are over-looked, and its resolution in differentiating sites is good even when sites are under-sampled (see *Plant species misidentification and detection* and *Sample area and sample intensity*).

From a metric calculation perspective, there are two ways that users have aggregated plot-based sample data into FQA scores. First, species found in quadrats can be compiled into a single cumulative site sub-sample species list, where each species is counted only once (e.g., Foster et al. 2007). Alternatively, scores have been calculated for each quadrat at a site, and then quadrat scores are averaged (e.g., Bourdaghs et al. 2006, Bowles and Jones 2006). Quadrat averaging can be appealing because it can allow for comparison of FQI among sites that were sampled differently or are of different sizes (controlling the area-richness bias) (e.g., Bourdaghs et al. 2006, Johnston et al. 2010), *as long as the same sized quadrats are used*. However, because calculating scores this way will "count" the same species at a site multiple times, it should be remembered that it is functionally a frequency-weighted measure, as a species influences the score of every quadrat it occurs in. Because of this fundamental difference, users must be sure to indicate whether they are calculating scores using a simple species list or by quadrat-level averaging, as this has unfortunately not been reported in some studies (e.g., Spyreas et al. 2010).

From a performance perspective, McIndoe et al. (2008) have argued that quadrat-level averaging (at the site, plot, or transect level) is preferable because it emphasizes frequently occurring species and tempers influences from singleton, rare, ephemeral, or outlier species that disproportionally affect site scores (see also, *Abundance weighted measures*). These authors also suggested that quadrat-level averaging better elucidated the underlying temporal trend in their young, unstable restoration, because it dampened inter-annual "noise" variation in scores due to the differences in species composition from year to year that can characterize volatile, young habitats. Beyond young habitats, dampening variation via quadrat averaging may also be useful when sampling habitats or sample plots that are particularly depauperate in species (Matthews 2003, Spyreas et al. 2010). For example, sites invaded by a dominant exotic species may have had nearly every other species eliminated (often except for a single native shrub or tree specimen). The resultant plot score would yield a misleadingly high overall value (especially its Mean *C*-value) as the average of two plants— because other species that might have moderating effects on the score have been eliminated, leaving only the dominant exotic and a single moderately Conservative native. Similarly, misleadingly high Mean *C*-values can occur when experimental plots contain few species because they are very small, and therefore, not diverse, and prone to outlier species scores (see also Jones 2004, Jog et al. 2006, Spyreas et al. 2010). Quadrat-level averaging may help to ameliorate this problem, but in such instances, using FQI might lead to a more representative Floristic Quality value as it would better reflect this extremely low sample diversity (see also, *Sample area and sample intensity*).

A few studies have compared the performance of quadrat-averaged, versus sample species list, calculation methods directly. Spyreas et al. (2012) and Rothrock et al. (2011) found that Mean *C* values in old fields exhibited identical trends over time when calculated either way. However, scores calculated via quadrat-level averaging were slightly lower than those calculated from accumulated species lists (Spyreas et al. 2012), which suggested that the latter method decreased the influence of less frequent species that may have been more Conservative. Kercher et al. (2004) found that Mean *C*-values calculated from site level species lists could not statistically detect hydrological disturbances to wetlands, but when they were calculated from plot-level averages, values were far lower in the disturbed wetlands. Similarly, quadrat-averaged FQI scores were more responsive to a prescribed fire treatment than transect level cumulative list scores (Wilhelm and Masters 1994). This suggests that quadrat-averaged calculations could be quicker to respond to site changes than whole site species lists (see discussion in, *Abundance weighted measures*).

Sampling across ecological gradients, ecotones, or community types at the site

Another sampling consideration is how to deal with habitat heterogeneity, ecological gradients, and vegetation boundaries (ecotones) across sites. There is conflicting information on this topic in the FQA literature. Andreas et al. (2004) recommended that "Compositing of scores from different communities at a site should generally be avoided unless the study expressly makes note of it. Where compositing is done, scores from the individual communities should also be reported." Similarly, Rentch and Anderson (2006) warn that "pooling of results from communities within a single larger complex may give misleading results for reasons that have little to do with inherent floristic quality." Alternatively, some texts encourage sampling across ecotones and communities. "Plant surveys for determining FQI should include buffers and ecotones and disregard changes in community type (Wilhelm and Ladd 1988)." Such contradictory recommendations are likely attributable to different aims by these authors. Original FQA texts focused on inventorying entire sites for their conservation value, so complete sampling was encouraged. But, FQA is now often used to compare sites with one another. And, in such studies, site sub-samples need to encompass a relatively homogenous area, because artificially adding habitat heterogeneity and beta-diversity to the sample by crossing ecotones would likely bias experimental site comparisons. Therefore, in most site score comparison studies, sampling across community-types would be discouraged..

The question of how fine of community-type distinctions should be made when sampling, and when ecological gradients or ecotonal boundaries should be avoided, or course defies easy description and will probably vary depending on the user. Swink and Wilhelm (1979) suggest that where community or vegetation dissimilarities "are well marked both florististically and topographically" they can be accounted for fairly easily. They describe two scenarios; the first describes the sand dune and swale communities that border the Great Lakes:

> "…two rather discrete communities, the integrity of which seems inextricably dependent upon their somewhat anastomosing, wavy topography. Here the individual swells are so narrow, and so clearly related to the swell on adjacent sides of intervening swales, that the swell-and-swale scenario has come to be viewed as an entity in and of itself."

In this case, sampling and considering "the swells and swales as separate and discrete entities seems ludicrous and patently impractical." This is compared to a second scenario, sampling a site of about 8ha. Here a mixed-woodland grades downslope into a sedge-meadow. Evaluating these topographically dictated communities is quite workable because they are large enough to sample individually and because the "common [floristic, structural, and physiognomic] denominators are not difficult to derive from among upland woods and sedge meadows", making them dissimilar enough to delineate.

To summarize, the best way to handle natural ecological gradients, heterogeneity, and ecotones in sample units will not always be obvious. FQI users are advised to limit site-samples to areas of homogenous vegetation structure, and to avoid crossing obvious ecotonal boundaries (due to the richness effect). Mean *C* should be less affected by such considerations because of its insensitivity to beta-diversity gradients. The effect of combing communities and crossing ecological gradients on FQA is suggested as an important area for further study (Bried and Edinger 2009) (see also, Chapter 1). Finally, surveying areas of different human disturbance histories within a site is appropriate when characterizing the entire site's Floristic Quality, but site sub-samples should be constrained to areas that are homogeneous in their disturbance history to facilitate their comparison.

Plant species misidentification and detection

Species lists will inevitably be incorrect due to misidentifications, and/or incomplete due to plants being overlooked during sampling and the annual, seasonal, or spatial variation that affects their apparency and detection (Scott and Hallam 2002). How robust are FQA metrics to these two sampling issues? With regards to incomplete samples, estimates are that a typical, single-site visit FQA sample during the growing season will typically yield about an 80% complete species inventory (Taft et al. 1997). Between 40%-80% (depending on the size and richness of the site) of the species from a sample area-site are necessary to attain a stable Mean *C* score for prairie (Rothrock and Homoya 2005). Cohen et al. (2004) compared Mean *C*-values generated by including only abundant, common, or uncommon species, which totaled between 28%-41% percent of the total species identified. They found that correlations between these groups' Mean *C* and site disturbance were not significantly different from those calculated using all species (although R^2 values were lower). Similarly, Mean *C* calculated from either the inner, intermediate, or outer wetland zones separately correlated highly with Mean *C* calculated from the entire wetland's flora (r = 0.93, 0.74, 0.99, respectively) (Murray-Hudson et al. 2012). Therefore, effects from overlooked species do not seem problematic for FQA, and incomplete species lists may be expected to yield accurate FQA analyses (Lopez and Fennessy 2002, Bourdaghs et al. 2006).

Effects from plant misidentifications in ecological studies are less well known (Scott and Hallam 2002). Several authors have stressed that the botanical expertise needed to accurately identify large numbers of plants to the species level is the largest drawback and limitation to FQA (see also, *A recognized goal for FQA* is that nationwide (U.S.) comparisons become possible (Medley and Scozzafava 2009). However, the likelihood for continent-wide FQA integration is questionable, because FQA is based on region specific floral composition and behavior (see *Conservation* “hotspots” are often mapped and compared at global scales (Myers et al. 2000, Brooks et al. 2006). Such maps are often underpinned by ecological factors such as the global latitudinal diversity gradient (LDG) (Visser et al. 2014), and by anthropogenic land use patterns. However, comparisons of the conservation value of habitat patches across landscapes— the scales at which most conservation and restoration efforts happen— are not effectively made, because the means to do so are not well-developed (Stein 2002, Ejrnæs et al. 2008). Existing vegetation-based assessments of habitat conservation value at these scales have used endemic species, species diversity, phylogenetic diversity, rare species, indicator species, or indices that combine several metrics to rank areas (Mack et al. 2008, Kiera et al. 2009). But, such measures have not been readily adopted by conservation practitioners, either because they do not have a broad enough ecological scope (e.g., focus on specialized taxa or specific site properties), or because they are too difficult to generate, interpret, or compare across areas (Niemi and McDonald 2004, Pearman et al. 2006). Thus, generally applicable measures for assessing the conservation value of natural area vegetation, which are simple and flexible enough to be readily adopted by conservation practitioners, are wanting (Niemi and McDonald 2004, Taft et al. 2006).

It is in this context that Floristic Quality Assessment (FQA) has increased in use and influential over the past 30 years (especially in North America) (Appendix). A site’s plant species list is used to calculate its Floristic Quality value with two primary metrics. The first is the mean of *a priori* conservation values assigned to each species. Species conservation values are termed Coefficients of Conservatism, or *C-values* (further described in Methods), and their average for a site is the Mean *C*. The second metric incorporates site species richness and Mean *C* into the Floristic Quality Index (FQI). Floristic Quality metrics (Mean *C* and FQI) are said to measure an area's biological Integrity, conservation value, or habitat degradation level via the amount of human induced changes to that area's plant assemblage (Taft et al. 1997, Spyreas and Matthews 2006, Mack 2007). FQA metrics are popular for a number of reasons: they perform well, being characterized as "remarkably robust" for their stated objectives (Miller et al. 2006); they capture information that is unique among existing conservation measures (Appendix); they only require a plant species list for their calculation; their values can be compared over time (Spyreas et al. 2012); and their values can easily be compared between all types of sites within a given region (FQA regions are most commonly states/provinces, Appendix).

The ability to compare varied sites across large regions with a single metric has spawned many different uses. For example, they are used in research on site restoration success and failure across vast areas (e.g., Mushet et al. 2002, Balcombe et al. 2005, Matthews et al. 2009a), and to study the causes and patterns in statewide habitat degradation (Nichols 2001, Milbauer and Leach 2007, Kraszewski and Waller 2008, Rooney and Leach 2010, Smart et al. 2011). FQA is also influential because of its direct use as a habitat evaluation criterion by various agencies. For example, the Nature Conservancy uses FQA to "prioritize high quality natural areas to include in their natural areas networks" (Rocchio 2007). Some federal, state, county, and municipal agencies have legally mandated that FQA metrics be used to evaluate sites across their jurisdictions. For example, a fixed benchmark FQI value threshold of 20 has been used by state and federal agencies as the determinant criteria for mitigating wetland destruction, as regulated under the Clean Water Act (Herman et al. 1997, Streever 1999, Herman et al. 2001, Matthews and Endress 2008, Matthews et al. 2009b). Such use assumes that a value of 20 is an equivalent and absolute value, such that any area's FQI value can be held up against this benchmark and its meaning will be equivalent. For example, any site scoring below 20 is considered easily replaceable and might receive less protection (Wilhelm 1992, Matthews et al. 2005).

Use of FQA across large regions has therefore been encouraged as objective and repeatable; a quantitative criterion that has encouraged "uniformity in natural area evaluation, enabling planners, land custodians, ecologists and other practitioners to make standardized comparisons among various open land areas" (Wilhelm and Ladd 1988). Region-wide metric uses assume that site values only vary due to anthropogenic factors (e.g., livestock grazing, logging, exotic invasive species, ecological management regimes) (Spyreas and Matthews 2006, Milbauer and Leach 2007, Spyreas et al. 2010). Thus, Floristic Quality is otherwise assumed independent of location, even across large U.S. states and Canadian provinces (e.g., in Florida: Cohen et al. 2004, Ohio: Mack 2006, the Dakotas: Hargiss et al. 2007, Colorado: Rocchio 2007, Alberta: Raab and Bayley 2012).

Despite the assumption of regional neutrality, other conservation metrics and ecological indicators are inherently varied across regional ecological gradients (e.g., Pearman et al. 2006, Muratet et al. 2008). For example, plant species richness varies inversely with latitude at global scales, and it varies regionally in association with many other non-anthropogenic factors (e.g., productivity, climate, disturbance regime, glacial history, elevation, edaphic conditions, biogeography) (Mutke and Barthlott 2005, Sarr et al. 2005). As richness is a constitutive component of FQI, FQI is also expected to respond to spatial variation in these factors. On the other hand, Mean *C* is not tied to richness and it might be assumed to be independent of spatially varying ecological factors. The level to which FQA values vary due to regional ecological gradients that are unrelated to human disturbance, would need to be accounted for in their region-wide use (Bernthal 2003). If not, this would confound the use of FQA in making conservation decisions. For example, land acquisition and protection in regions where ecological characteristics naturally generate lower FQI values could be disfavored.

Studies have not yet determined if FQA values vary within a region. The largest existing study found latitudinal differences in FQI, Mean *C*, and species richness in ponds and lakes across Wisconsin (USA) (Nichols 1999). However, Nichols (1999) also pointed out that observed patterns could have simply been reflecting the different habitat-types sampled in different regions (oligotrophic versus eutrophic ponds; Chapter 1). Conversely, neither Cohen et al. (2004) nor Miller et al. (2006) found latitudinal differences in wetland FQA values. Finally, Johnston et al. (2010) found strong latitudinal variation in values from wetlands bordering the Great Lakes. Unfortunately, because they compared *C-values* from different states, their results must be viewed cautiously (i.e., because each state's floras have been assigned different *C-values* for a given species; Appendix). Additionally, because they only reported FQI values, it was not possible to determine if their observed patterns were due to differences in richness, Mean *C*, or both.

A considerable amount of data is required to make conclusions about statewide Floristic Quality patterns and regional trends in FQA metrics, and this has hindered their understanding thus far. If regional variation in FQA is to be understood, analyses *must* be conducted at statewide scales, because each species *C-value* is scored with respect to its *simultaneous behavior across the entire state* (see Methods). Therefore, because values in one area of a state might be counterbalanced by another part, assessments of regional FQA patterns must have complete statewide coverage. The following additional criteria for data exist if regional FQA patterns are to be understood: 1) habitat-type should be controlled for; 2) an unbiased site selection criterion with probabilistic sample design should be used (randomized, stratified-systematic, etc.) (Genet and Olsen 2006); 3) site sampling should be consistent (sample intensity, area, or season can affect FQI values, Appendix); 4) studies should examine more than wetlands or aquatic habitat-types.

). Wetland IBIs are currently being adapted for nationwide use and they may provide a template for how to develop a centralized system of this kind (U.S.E.P.A. 2010). Regardless of how such a nationwide undertaking unfolds, there is little doubt that

Conservatism values will be assigned for more regions and that FQA will expand in its geographic coverage and use.

FQA in hypo- and hyper-diverse). And, this obstacle may only get worse in the future, as “The number of individuals with taxonomic skills sufficient to correctly identify grasses, sedges, and other difficult taxa is decreasing (Noss 1996), as compensation for this expertise is inexcusably poor (Rooney and Rogers 2002).” This lack of expertise could not only hinder the use of FQA, but lead to high species misidentification rates and error in generated scores. No study has assessed species misidentification rates effects on FQA directly, but some studies can be instructive. Rooney and Rogers (2002) asked what would happen if a taxonomic group that is notoriously difficult to identify (the genus *Carex*, sedges) was removed from score calculations in 59 forest sites. Carices represent an appealing group to test in this way because they are difficult to identify— so much so that many vegetation studies simply lump them together to avoid their misidentification (essentially removing them from the sample). They are also comparatively overrepresented in many north temperate habitats (Kartesz 2014a). Removing Carices from species lists led to a slight decrease in forest plot Mean *C* (5.45 from 5.58), as well as the inevitable decrease in FQI due to decreased richness (37.6 to 29.9). However, it is also true that *Carex* might not be generally representative of species misidentifications because *Carex* might not represent the *typical* misidentified species. Future experiments might randomly mis-assign species scores or remove random species from lists and look for a critical level of misidentification where FQA becomes unreliable (e.g., 25% of species unidentified or misidentified).

Rather than misidentifications or missed species simply obscuring Floristic Quality resolution or increasing metric variability, Bernthal (2003) makes the following observation:

> “[FQA results] may be strongly affected by observer expertise, restricting the comparability of results between observers of different skill levels. The level of

> skill required for acceptable results is still unknown. More skilled observers are likely to identify more species and therefore generate higher FQI values. More skilled observers are also likely to find the more Conservative species, and would tend to generate higher Mean C values.“

This suggests that misidentification would produce directional biases, where better botanists would produce higher scores by finding more Conservative species. Alternatively, in the example of woodland Carices above, if they were on average less Conservative than the rest of the forest flora, overlooking them would lead to decreased Mean *C* scores. The exclusions of commonly misidentified or unapparent species could bias scores in either direction.

In conclusion, it seems that FQA measures can tolerate a considerable number of undetected species in a sample before their accuracy declines. But, it is not yet clear how important species misidentifications are. It is also not clear if missed species or misidentifications have the tendency to produce directional shifts in scores, or if they simply increase imprecision or variation in the Floristic Quality estimate. Unidentifiable species are typically dropped from the calculation. Or, occasionally taxa are identifiable to genera, and if all the possible species in their genus have the same *C*-value, then that value may be used. Or, if they are similar the average within the genus can be used. Finally, unidentifiable, or difficult to identify species can occasionally also be scored as the “least Conservative species in the genus that was likely to have been present in the region and that could have been confused with the specimen” (Francis et al. 2000). While only a small portion of their sample (between 2-3% of records and species) needed to be accounted for in this way, these authors suggest that this accommodation can help to reduce the imprecision that is introduced when specimens are ignored or not completely identifiable to species.

Seasonal variability in FQA sampling

In addition to the temporal variation associated with natural disturbance, intra-annual differences in plant community richness or composition could bias FQA scores. For example,

> "Many species will not be observable or identifiable by even the most skilled observer at certain times of the year. Repeated sampling over the course of a growing season will allow the closest approximation of the "true" Mean *C* and FQI values, but this is not likely to be feasible in many situations. (Bernthal 2003)."

Because site surveys are often based on a single visit, their Floristic Quality is calculated from the limited suite of plants identifiable on the survey date. It is then of considerable importance to know whether scores from any given season's sample differ from other seasons, and relatedly, whether more phenologically comprehensive species lists (generated from multiple samples across a year), differ from single site visit score. If any of these are true, comparisons of site scores from different calendar dates might be biased.

Although temporal bias in FQA applications is frequently mentioned, it is seldom evaluated (Bried et al. 2013). A few studies have looked at the effects of adding sample dates within a year on *FQI*– where the cumulative species list becomes more complete and also more phenologically comprehensive. FQI scores have been shown to differ between summer and combined summer-fall species lists, in several habitat types (Fennessy et al. 1988, Fennessy et al. 1998, Lopez and Fennessy 2002, Bernthal 2003, Andreas et al. 2004). Such differences are to be expected for FQI because of the increase in richness with additional sampling. Despite score changes, the relative rankings of these sites have remained consistent with additional sample dates. And, FQI scores have been found to be highly correlated with one another across seasons; between 92-94% from summer and summer-fall (Fennessy et al. 1988, Lopez and Fennessy 2002) and 95-96% between spring, summer, and fall (Fennessy et al. 1988). This suggests that FQI comparisons among sites sampled in different seasons might be possible if, “FQI accumulation curves” can be used to extrapolate scores across the season as species accumulate through a season’s sampling. This could be used to correct scores sampled in different sample seasons, or for comparing single- versus increasingly complete multiple-site visit surveys.

Single samples from across the calendar year have also been compared with one another (as opposed to single versus multiple site visit scores). Bried (2013) did not find a seasonal sample effect in Mean *C* scores for several community types tested. Nichols (1999) did not find significant differences for either FQI or Mean *C* scores of lake aquatic floras sampled in June-July versus August-September. Francis et al. (2000) found that species richness, and therefore FQI, was higher later in the season, but Mean *C* did not differ in spring versus summer woodland samples. Herman (2005) found the same pattern in wetlands. Matthews et al. (2005) found that Mean *C* generally did not vary in wetlands sampled any time between April and August. Only two studies have shown some kind of seasonal effect that might suggest bias from sample date. Wilhelm and Masters (1994) found that plots sampled later in the season at a woodland site had lower FQI values, but the sample size was only one. In the most expansive study on the subject, Matthews (2003) showed that FQI values from several hundred wetlands did differ by survey month, but this was generally attributable to higher species richness during the growing season. Mean *C* values did not differ statistically throughout the year, except that visual examination of trends suggested that January-March values were lower. Of course, FQA-based sampling during mid-winter in continental North America is of little practical utility.

Although seasonal variation in FQA scores does not seem commonplace (aside from FQI differences due to sampleable richness changes throughout the year), the instances where it does occur suggest four potential explanations. First, for early calendar months (e.g., January-March), the few species that *are identifiable* could simply be less Conservative on average. For example, many invasive exotic species are notable for their early emergence and later senescence compared to natives (Spyreas et al. 2004), suggesting a pattern where non-natives are simply more apparent than natives during these months. Second, the seasonality to site Mean *C* scores could reflect a bias in how C-values were assigned, where botanists may unwittingly favor the phenology of certain seasons, although this has not been tested. Third, the site(s) in question could have historically incurred disturbances in early months that did not occur in other months. This historic seasonally discrete disturbance could be mirrored in seasonal patterns in contemporary Floristic Quality at a site. For example, Taft (1996) suggested that because cattle were historically rotated off lush cool-season pastures in the spring into woodlots in summer, the intense disturbance from cattle grazing during summer months in many eastern and central North American forests could still be being echoed in the *summer* flora today. Those Conservative plants that were active and apparent to cattle during the summer would have been more impacted than Conservative plants emerging at other times of the year. Therefore, the Conservative "spring ephemeral" floras of today would have persisted because they were senescent when cattle were grazing in summers past, and would yield comparatively high spring FQA scores today compared to species with summer-fall phenologies. Other types of temporally discrete FQA effects have been shown by Wilhelm and Masters (1994) in a prescribed fire study. Forest plots sampled in the spring showed no effect from a burn treatment over time, while plots sampled in the fall (when the treatment fires were conducted) displayed a two-fold increase in FQA scores. This suggests a related fourth explanation, that the floras of certain seasons are more sensitive to human disturbance than other seasons, or similarly, the types of disturbances that tend to occur in some seasons have more severe than those in other seasons.

In summary, as with inter-annual variation (see

Short-term cyclic or inter-annual "nuisance variation"), users should consider the possibility that before scores are compared from sites sampled on different calendar dates, vegetation surveys conducted in different season may be needed to be screened for potentially confounding temporal differences. It is clear that *FQI values* generated from different sample dates may not be compared with one another because richness often varies across seasons, just as FQA scores generated with different numbers of sample visits may not be compared. Further research is needed to determine whether FQI-species accumulation curves could account for such sample intensity difference and/or allow for across-season comparisons. Mean *C* scores are more useful when comparing sites surveyed at different times of the year because of their consistency across seasons, and they should yield reliable intra-annual comparisons samples within the growing season (i.e., April-October) (Francis et al. 2000, Bernthal 2003, Matthews 2003). Beyond concerns about sampling, this seasonal consistency in scores suggests an interesting FQA property, despite highly dissimilar community composition at a site from different sample dates, early season species will have similar *C*-values as species that emerge late in the season (Bried et al. 2013). In those instances where seasonal differences in Mean *C* are found, further study is needed to determine their cause. For example, knowing whether historic disturbances that were seasonally discrete yield temporal Mean *C* patterns long into the future would apprise users of potentially misleading across-season score comparisons.

Alternate FQA calculations

A considerable number of alternatives to the standard FQA metrics have been proposed, some with specific objectives, and others that are simply intended to improve upon the original metrics in general usage. An appealing aspect of FQA is its straightforward, simple formulas. And, because these alternative metrics tend to require additional equation parameters and data, any new metrics must be tested to ensure they outperform the standards, and that they are not redundant with them to justify the extra effort their calculation requires. In any case, users are urged to report scores from original metrics when the alternative ones are used (Andreas et al. 2004).

Non-native species

Alternative FQA metrics frequently incorporate non-native species in novel ways. Bowers and Boutin (Wilhelm and Ladd 1988, Bowers and Boutin 2008) went so far as to assign separate invasiveness values (Coefficients of Weediness [CW]) to all non-native species in southern Ontario's flora (beyond simply scoring them with zeros). Each non-native species was assigned a negative CW value (-1 to -3), and these are then aggregated into a Sum of Weediness score for a site. In their study, Sum of Weediness scores were highly correlated with human habitat disturbance, but other FQA measures performed even better (% Conservative species, native Mean *C*, native FQI), as did other common non-native species metrics (% non-native species, number of non-native species). In comparison, botanists in some regions have proposed negative non-natives species *C*-values (as opposed to zeroes), for use in the standard Mean *C* and FQI metrics (e.g., in their first version, Swink and Wilhelm 1979, Ladd 1993). Miller and Wardrop (2006) proposed a new metric, the adjusted FQAI (*I'*) that included a non-native species richness parameter. Its aim was to eliminate the richness bias in FQI and to penalize sites for their non-native species richness. But, FQAI has performed nearly identically to native Mean *C*. It was nearly as strongly correlated to site disturbance scores ($r = 0.85$ native Mean *C*, $r = 0.87$ adjusted FQAI), and it was 96% ($r = 0.98$) correlated with native Mean *C*, suggesting near-perfect redundancy. Similarly, Forrest (2010) found that site disturbance scores were better correlated with Mean *C* than adjusted FQAI ($r^2 = 0.66$ and 0.62 respectively).

The most frequently varied component of FQA metric calculations are the inclusion or exclusion of non-native species. Non-native species are either included with *C*-values of 0, or ignored, in the standard metrics. There is debate as to which of these is the best course, even within the standard FQA texts (Swink and Wilhelm 1994c, Taft et al. 1997). While both are considered valid and they yield robust measures, strong arguments have been made in support of either method, and these are reviewed below.

For most ecologists and conservation practitioners, the mere existence of exotic species at a site is seen as diminishing a site's conservation value (Martin and Blossey 2012). Swink and Wilhelm (1979, 1994c) also write that their presence represents a disturbance. FQA studies that include exotic species in metric calculations are more common. And, nearly every study comparing them has found that including exotic species in metrics achieves higher correlations with site human disturbance measures, often by a considerable degree (Forrest 2010). Or, where correlations do not improve, they are no worse than native-only FQA scores (e.g., Foster et al. 2007). Several authors argue that this is reason enough to include exotics (Cohen et al. 2004, Rocchio 2007).

Ervin et al. (2006) argued for non-native inclusion in evaluations of vegetation quality because "disregard for non-native species may result in overestimation of ecological integrity." This overestimation comes because the past site disturbance represented by exotic species presence is otherwise unaccounted for. Furthermore, exotic species inhabit space and resources at the expense of native species, representing the exclusion of past, current, or future native taxa. Exotic species in a natural community can "interfere with recovery processes"(Wilhelm and Ladd 1988, Taft et al. 1997), so their presence should be seen as depressing a site's current and future native Floristic Quality. Ervin et al. (2006) concluded that "insufficient empirical data exist to justify exclusion of exotic species from assessment methodologies. Thus, until data indicate otherwise, exotic species should be incorporated into any proposed method of quantifying wetland health."

Alternatively, Swink and Wilhelm (1994b, but see 1979 below, Taft et al. 1997) did not assign exotics a value and did not include them in metric calculations. Their reasoning suggested two explanations for this. First, there was the belief that assigning all exotics zeros was arbitrary, because exotic species actually have different remnant dependencies or human disturbance tolerances. "Some weeds, such as *Agrostis alba* or *Poa pratensis*, are virtually ubiquitous in the region, so they are almost inevitably detected somewhere in a site, if only along or near a trail. Such an occurrence does not signify significant degradation to an area." Other would then presumably be limited to natural areas of relatively higher quality. This highlights their rationale for changing from their 1979 edition, where introduced species were assigned *C*-values between -3 and 0 (Swink and Wilhelm 1979). This was originally done to reflect that some species are detrimental to sites (invasive), whereas others are innocuous. In the 1994 edition however, they concluded that quantitatively assigning how "negative" a species was to a sites conservation value was unrealistic, and a fundamentally different undertaking from assigning native species scores based on occurrences in, and dependency on, remnant habitats. Swink and Wilhelm's second point was that exotics species evolved disjunctly from local plant assemblages and were not present during remnant plant community's genesis and assembly– so why would it be appropriate to use them to compare, assess, or understand native plant community dynamics? Despite their shortcomings, definition of species Conservatism based on niche breadth (see *Breadth of habitat-types*) would seem to follow a similar rationale, where it would not make sense to score exotics because "non-native species by definition have no native fidelity to a particular [native remnant habitat] type (Cretini et al. 2012)."

Proponents of native-only calculations have also argued that if exotic species presence were harmful to site Floristic Quality, this would already be registered in depressed native FQA values. This provides a counterpoint to the "overestimation argument" for including exotics suggested above, which would make exotic species inclusion redundant and unnecessary (e.g., Mushet et al. 2002). Native-only calculations are also said to be preferable because they are the most straightforward application of FQA, containing no hidden information. But, even these transparency advocates have recommended using some kind exotic species index in site assessments, because of their recognized importance in determining site quality (Taft et al. 1997, Rooney and Rogers 2002). Finally, some studies have not found appreciable performance differences when including or excluding exotics. Foster et al. (2007), "calculated Floristic Quality using several other formulations…but report only one [natives only] because all formulations yielded similar results." Bourdaghs et al. (2006) found no significant differences between metrics that included or excluded exotics in the metrics' ability to discriminate sites. Thus, native and total Mean *C* are sometimes seen as performing similarly enough to be interchangeable (e.g., Sivicek and Taft 2011), due to their inevitably high correlation (Taft et al. 1997, Cohen et al. 2004).

In reality, different site characteristics and dynamics will probably lead to different FQA score by exotic species relationships. This means that different sites differentially benefit from including or excluding exotics. Sites with many non-native species at low abundances, versus sites dominated by one or a few dominant invaders, would depress scores differently. The latter would affect scores more indirectly by outcompeting and depressing the native Floristic Quality component (e.g., Frieswyk and Zedler 2006), while the former would directly decrease scores due to the inclusion of many zeroes in the calculation. Relationships between exotic species and measured Floristic Quality may also differ in high versus low-quality habitats. For example, Cohen et al. (2004) found that sites with low Mean *C* exhibited the largest differences between Mean *C* and native Mean *C* because of the correlation between exotic species richness and Mean *C* at their sites. Thus, at high quality sites there were fewer exotic species to create a gap between the different formulations. Such differences might also be expected between established remnant habitats and created habitats, or in younger versus older habitats. However, Spyreas et al. (2012) found that total and native-only calculations paralleled each other over 50 years of old-field succession, despite exotic species richness and dominance varying dramatically over that same time period. This suggests a surprisingly consistent exotic species effect on FQA scores during early- to mid- successional stages. The magnitude of differences between calculations at a site has also been suggested as a way to indicate if its flora has sustained detectable degradation by exotic species. For example, a difference of 0.5 or greater between native Mean *C* and overall Mean *C* is suggested to indicate sites with significant impacts to their Floristic Quality from exotic species (Rothrock and Homoya 2005). This is a provocative, but yet untested proposition.

In summary, while there are strong arguments for using either calculation method, both clearly perform well. Formulation choice may ultimately hinge upon user priorities. Is the primary objective quantifying the conservation value of the current native flora at a site? Or, is it quantifying how much disturbance the site has accrued and its overall ecological integrity? With the former, exotics are of no direct interest as only native plants have conservation value. With the later, exotics are certainly reflections of past, current, or future degrading factors and should be accounted for. Future research should aim to establish rules or patterns in when one formulation performs better than the other. Until then the recommendation that "calculations be made using all species (native and adventive) as well as native species only" is prudent (1997). Regardless of which calculation(s) is used, authors must indicate whether exotic species were included so that scores can be appropriately interpreted.

Abundance weighted measures

FQA's original authors very clearly intended to create a measure that could be used irrespective of relative plant abundances; "It avoids the use of frequency, dominance, physiognomy, or productivity of an individual plant species, primarily because positive or negative values with respect to plant community quality are either irrelevant or only ambiguously related to these factors" (Swink and Wilhelm 1994b). At the same time, this is one of the major criticisms of FQA, as some authors consider relative abundances to be integral reflections of a community's conservation value and disturbance levels that should not be ignored (Bowles and Jones 2006, Nielsen et al. 2007). Alternate FQA metrics that weigh species by their abundances have been created. Their performances have been found to be similar to standard FQA metrics in some studies, and different in others. Bowles and Jones (2006) found that weighted and un-weighted Mean *C* scores in remnant prairies were highly correlated with each other (Dry prairies, $r = 0.77$, $P < 0.001$; Mesic prairies, $r = 0.81$, $P < 0.001$), while non-significant correlations ranging from -0.36 to 0.22 were found between abundance-weighted and standard metrics in a young, restored prairie (Anderson et al. 2006). Hopple and Craft (2013) found nearly identical relationships between the abundance of species in different *C*-value classes (e.g., 0-2, 2-4), and species richness in these classes. This relationship was observed in both restored and remnant wetlands and suggests a high congruence between patterns in species *C*-value presence and abundance.

The critical question is how well do abundance-weighted versus standard FQA metrics register site disturbances and degradation? Anderson et al. (2006) compared the ability of a new abundance weighted FQA metric to discern browsing impacts from overabundant deer. They found that their Weighted Mean Fidelity (WMF) metric, was considerably more sensitive to high deer-browse levels than native Mean *C*, abundance weighted native Mean *C*, abundance weighted native FQI, species richness, and native FQI (from best to worst performing, respectively). Few species were completely eliminated from the site by deer, but species whose frequency increased under intense browsing had lower *C*-values compared to those that decreased, causing the abundance weighted differences observed. Another study found that WMF consistently increased over 13 years in a prairie restoration, nearly doubling by study's end, while Mean *C* only increased slightly, and in a circuitous fashion (Anderson et al. 2007). At the same time, species richness, evenness, and H' diversity were all unchanged at the site. This clear trend was seen as an indication that WMF was more useful than other metrics in tracking restoration progress. However, it is important to understand that their WMF only includes forbs in its calculation (perhaps to focus on plants that deer browse, or on the plant groups that often limit restoration success), while its performance was compared to measures that included all species. Therefore, it remains to be seen whether WMF versus standard whole flora metric comparisons perform equivalently— especially when one considers that native Mean *C* outperformed abundance weighted native Mean *C* in their study (Anderson et al. 2007).

Nearly every other study has found little or no performance difference between traditional and abundance weighted FQA measures– for example, in forests (Francis et al.

2000), prairies (Bowles and Jones 2006), wetlands (Cohen et al. 2004, Bourdaghs et al. 2006, Miller et al. 2006), and in multiple habitat comparisons (Rocchio 2007). Forrest (2010), found that abundance weighted metrics were considerably worse than native Mean *C* and native FQI at explaining site disturbances. Therefore, because they increase labor, data, and computational requirements without a clear performance improvement, abundance weighting tends to be discouraged for *general* use (Francis et al. 2000, Bourdaghs et al. 2006, Ervin et al. 2006, Rocchio 2007).

A primary reason for not originally including abundance in FQA was the great natural variation that occurs in species abundances. "Abundance and frequency are often artifacts of the season or year, and may fluctuate greatly. Some species which are 'dominant' (big or obvious) in spring can be scarcely evident in fall, replaced by species which were scarcely evident earlier in the year" (Wilhelm and Ladd 1988). Therefore, beyond questionable performance improvement and increased effort, this variation makes abundance-weighted site scores sampled in different seasons or sampled with different intensities less comparable with each other (Ervin et al. 2006). Over-emphasizing large, showy, or easy-to-detect species may not only increase sample-error and decrease resolution, but it may also bias FQA comparisons by inflating scores in certain types of sites (see discussion, *Plant species misidentification and detection).* Furthermore, deemphasizing diminutive or less-common species it is also seen as philosophically objectionable by original FQA authors:

> "It is not clear, furthermore, what significance one could attach to a perceived relative abundance of individual members of a discrete community... [FQA] is based on the mere presence of a plant and its rating coefficient, as its operative premises. The resultant set of coefficients reflects the degree to which Conservatism is present… [and this avoids deemphasizing] the majority of the floristic elements from the assessment rationale. **It would seem that a diminutive sedge with a coefficient of 7 is every bit as important to the biological and**

> **genetic diversity of a site as a large or common tree, especially since it is uncertain how abundant any particular organism should be.** Again, since perceived abundance and dominance may vary seasonally and annually, even in healthy systems, attaching fundamental significance to these factors in assessing natural quality can generate misleading conclusions. On any given site survey, the mere presence of a native plant speaks well enough for the conservatism it represents." (Wilhelm and Ladd 1988).

This passage also highlights that the process of *C*-value scoring does not consider abundance, which raises the questions of why FQA metrics would be expected to reliably perform with respect to abundance. *C*-values are only assigned with respect to presence-absence across remnant habitats because botanists have limited historical understanding of pre-settlement abundances— from the passage above, "how abundant any particular organism should be." Original authors have also suggested that abundances are redundant within FQA metrics. "An exaggerated abundance of any particular plant, valued high or low, often indicates that significant environmental alterations have occurred… [but] our own experience has shown that when certain few plants are inordinately abundant, the mean Floristic Quality is also depressed." (Wilhelm and Ladd 1988)

Despite these concerns, specific instances where weighted indices could be useful are acknowledged by many authors, including original authors FQA (see Figure 10, Swink and Wilhelm 1994b, Taft et al. 1997, Francis et al. 2000, Matthews 2003). For example, abundance weighted measures seem more likely to behave differently from un-weighted metrics in restorations (e.g., Poling et al. 2003, Anderson et al. 2006, Anderson et al. 2007). They are also suggested as useful when tracking changes within single sites over time, or where users wish to deemphasize rare species in their study (Anderson et al. 2006). These instances parallel performance differences in quadrat-calculated versus site-calculated FQA scores; where quadrat-calculated FQA scores are sometimes observed to be useful for tracking changes within sites (Swink and Wilhelm 1994b). This is not an unexpected parallel considering that calculating scores at the quadrat-level essentially weighs species by their frequency. If proven to be consistent, this would also imply that abundance weighting could be useful when assessing small sample areas, in low richness areas, in short-term studies, or in any other instances where a dramatic change in site species composition is not expected (Matthews 2003). Indeed, Poling et al. (2003) specifically proposed a new metric to measure Floristic Quality in small areas that is both abundance-weighted and calculated at the quadrat level (qFQI). However, its performance compared to other FQA metrics has not been assessed.

To summarize, abundance-weighted FQA measures do not consistently outperform original measures, and are not worth the extra effort for *general* FQA use. However, many authors concede instances where they may be useful, for example, in tracking changes within a single site, or where changes to local species abundance (especially dominant species) occur without changes overall species composition. In such instances, to ensure “maximum comparability between studies” users are urged to report results from original metrics as well as abundance-weighted measures (Andreas et al. 2004). In any case, users should remember when conceiving FQA, original authors placed far more confidence in species occurrences than on their abundances:

> “The collective data from all species inhabiting a site provides a concise, real measure of the extent to which the site represents a unique and irreplaceable element of our landscape. One can have much more confidence that an area is of natural quality if informed that there are 50 different plant species with a *C* of 5, than if informed that there are 50 individuals of a particular plant with a C value of 5. Similarly, the presence of 50 species with a C of 5 provides more insights into the potential quality of an area than being informed that a supposed "keystone" or "dominant" species has a cover value of 50%.” (Swink and Wilhelm 1994)

Subsets of floras

Subsets of site floras have also been used to measure site Floristic Quality. For example, as was discussed above, Anderson et al. (2006) only used forbs to calculate their WMF metric. Smart et al. (2011) calculated separate graminoid and forb FQA scores for prairies. They may have done this to try and isolate the relative effects of cattle grazing and herbicide use on Floristic Quality in prairies, as cattle preferentially graze grasses, and many herbicides are specific to certain groups of plants (e.g., forbs or grasses). Separate understory, shrub-sapling, and canopy tree layer FQA scores were calculated in a study of anthropogenic disturbance in forests (Nichols et al. 2006). The rational for this calculation was not provided, but tree FQA scores may have ultimately proven more useful at registering long-term disturbances and responses at broader spatial-scales because of the age of the tree canopy relative to other forest layers.

Another subsetting approach has been to bin species into *C*-value categories, to focus on either Conservative or non-Conservative elements of site floras. For example, the number of species with a *C*-value ≥ 5 or *C*-value ≥ 4 have been used alongside overall Mean *C* and FQI (DeKeyser et al. 2003, Hargiss et al. 2007, Mita et al. 2007, Euliss and Mushet 2011). Performance of these categorical variables was not compared to performance of the standard metrics in these studies. Matthews et al. (2009a) used species categories designated as non-Conservative (0-2), moderately Conservative (3-4), and Conservative (>5) to assess causes of wetland restoration success. Matthews and Endress (2009b) found that the number Conservative species (>5) was among the best indicators of wetland restoration success tested. Several wetland IBI's have included *C*-value category richness in their multi-metric Biological Integrity index; meaning that these variables were among the best from dozens tested (e.g., Miller et al. 2006). Bowers (2008) took a somewhat different approach and used the *percent* of Conservative species (4-10) rather than their richness as a variable (i.e., "PCC4-10"). They found that this had an even higher correlation with human habitat disturbance than standard Mean *C* or FQI. Score binning has also been used in conjunction with abundance weighting. For example, Andreas (2004) found that the percent cover (abundance) of human disturbance tolerant species (*C* = 0-2), separated wetlands based on their biological integrity, although it did not do this as well as Mean *C* (see also, Mack et al. 2000, Mack 2001). Miller et al. (2006) found that their binned, abundance weighted variables (% cover C ≤ 2 plants and % cover C ≥ 7 plants) had among the highest correlations with human disturbance in wetlands of 50 variables tested, on par with both Mean *C* and FQI.

In summary, some subset metrics perform well and others do not. Although those that do perform well are not clearly better or worse than original FQA metrics. Users are reminded to make clear that these are not the standard FQA metrics. And, future research should seek to describe these metric's site-level properties to verify that they meet the assumptions for use of typical FQA measures.

Other alternative measures

Bowers and Boutin (2008) assessed how well simply using the "sum of *C*-scores" as an FQA measure would work. The sum of species C-values at a site is a constituent component of both FQA metrics (where it is divided by the number of species to obtain Mean *C*, and multiplied by the square root of the number of species to obtain FQI). However, sum of *C*-values was outperformed by other FQA measures in their study (% Conservative species, native Mean *C*, native FQI). Other alternative metrics that are more or less intended to capture site Floristic Quality, but that contain no Conservatism values, have been proposed. For example, Ervin et al. (2006) tested FQA against four new formulas based on native species richness and plant Wetland Indicator Status (FAQWet index 1-4). None of these outperformed FQA in explaining site disturbance, but since they were intended to assess wetlands in regions where floras have not been assigned *C*-values, they could provide useful surrogates there. However, it is important that such metrics are not described using FQA terminology to avoid confusing users. For example, in the following sentence; "This index…was intended to serve as an alternative means of calculating ***floristic quality*** of wetlands where coefficients of conservatism are unavailable" (Ervin et al. 2006), inserting the caveat that these alternative metrics are intended to approximate Floristic Quality metrics would probably have been ideal since no Conservatism values were used in them.

Advancing Understanding of Site-Level FQA Scores

Precision, sensitivity, and variability in site scores

Ecological indicators require assessment of their statistical variability, precision, and accuracy (Niemi and McDonald 2004). FQA has been proven to be a robust measure (see, *Site FQA scores measure human disturbance and biological degradation*). But, establishing the precision of individual metrics is critical, especially because scores are frequently considered absolute measures, with exact thresholds ascribed to them (see, *Ranges and distributions in site scores*). Because the sensitivity of Floristic Quality metrics are not well-studied, some authors have questioned its ecological resolution. For example, even though FQA performed well in their study, Bowles and Jones (2006) suggested that its metrics may lack the ability to discriminate site degradation among the highest quality sites. "[FQA scores] may be most applicable in making comparisons between extremely low vs. high-quality vegetation." Similarly, despite high overall correlations with their site disturbance index (Mean *C*, $r = 0.82$ and FQI, $r = 0.75$), Miller and Wardrop (2006) suggested that Mean *C* and FQI metrics may differ in their precision at habitat quality extremes. But, they suggested that Mean *C* was suggested might be less effective at separating site differences at the *lower end* of the habitat quality spectrum.

Even though FQA metrics are continuous variables, due to this indeterminate precision question, Wilhelm and Ladd (1988) suggested that FQA based site rankings might best be done categorically:

> “It would be misleading to infer a profound difference in natural quality between two sites ranking, for example, 38 and 42. The rating system provides a general measure of site natural quality, but is not a precise numerical discriminator. Extensive application of this system to actual vegetative units has shown that discrete ranges of Index values are obtained and correlate closely with degrees of fundamental synecological integrity.”

In summary, further testing FQA metrics to determine their sensitivity and precision— their ability to make fine distinctions in site degradation levels— is one of the most commonly recognized FQA research needs (Taft et al. 1997, Nichols 1999, Francis et al. 2000, Bernthal 2003, Bowles and Jones 2006, Taft et al. 2006). This seems especially important at high- and low-quality habitat extremes. But, it is also important to point out that any study assessing the resolution of FQA metrics, must also recognize the imprecision in human disturbance measures used to rank sites (Bried et al. 2013). This underscores the challenge in establishing a true FQA sensitivity. Large studies, with very well documented site human disturbance histories, focused on the extremes of habitat quality, will be needed to better establish the sensitivity and precision in the future.

Ranges and distributions in site scores

A common question that users propose regarding scores is how to determine what constitutes a “good” site score. Some have used experience and judgment to delineate ranges or thresholds in scores that will signify a high-quality or “significant” remnant habitat:

> “We find that sites with a FQI of less than 20, based on “complete” inventory data, are usually severely degraded or derelict plant communities, or are very small habitat remnants. Sites with an FQI greater than 20 may be degraded but generally

> have potential for some level of recovery. Sites with indices greater than 35 are at least regionally noteworthy and often are sharply distinct from the predominant heavily degraded matrix areas in the landscape. Sites with indices greater than 45 are often also statewide-significant Natural Areas.“ (Taft et al. 1997)

Similarly, in the Chicago region, Swink and Wilhelm (1994b) first suggested that sites with a Mean *C* greater than 3.5 and an FQI greater than 35 had sufficient Floristic Quality to be “at least of marginal natural area quality”, while Mean *C* scores of 4.5 and FQI of 45 or greater were ”most certainly high-quality remnant natural areas.” Despite their basis in considerable experience, such valuations are subjective, and they are want for empirical validation. Unfortunately, these suggested threshold values have since been widely repeated and used to identify remnant or exceptional natural areas in other regions, before they have been empirically validated (Kowalski and Wilcox 2003, Goldblum et al. 2013). Furthermore, these cut-offs cannot be assumed applicable outside their region of origin, Indeed, the evidence suggests that may have never been intended to be used with such exactitude by their authors.

Instead, statistically determined score ranges could be used to more objectively identify site scores that are uniquely high across a region. This would make the determination of what constitutes a uniquely high-quality site much easier. This would also help to justify the use of strict score cutoffs where they are needed or are legally mandated, and it would avoid the appearance of arbitrariness to precise cut-off values, benchmarks, or ranges demarcating “exceptional” areas. Indeed, given limited conservation resources and the pressure to defend value-based criteria in conservation decisions, empirical statistically validated FQA valuations would be more defendable and more effective (e.g., Myers et al. 2000).

A first step in this process could be to determine what the highest attainable score in a habitat is— what can be expected from for a undisturbed high-quality habitat in the region. This could be the highest quality reference site score observed in a state, either for an existing site or for a historic reference. A discussion of the role of existing versus best possible historic, reference conditions is beyond the scope of this study, but such a choice would differ depending on the goals of the user and the available data. Once the upper limit for benchmark scores in a region is determined, the distribution and commonness of existing habitat scores on the landscape can then be used to statistically determine remarkable and less remarkable scores. For example, Nichols (1999) determined the median Mean *C* values of 554 lakes, reservoirs, impoundments, and ponds in Wisconsin. Using the median and overall distribution of scores, percent quantile ranges can then be generated and the determination of exceptional sites can be made (Malik et al. 2012).

To facilitate this process, Bernthal (2003) suggested that a "database of FQA site values, including a range of reference sites by ecoregion and habitat type" needed to be developed for each state. Ideally such data would contain scores that are: consistently generated (site sampling methods, score calculation, etc.); from sites that were selected systematically or randomly to reflect ambient conditions in the region and that provide an unbiased sample of the state's habitats. Another consideration to be made when populating such a database is whether to include restorations or "novel habitats" in these distributions, or whether to build them based solely on remnant habitats alone, as created habitats might constitute artificial scores (e.g., because they have been planted with high *C*-value species) (see *Restoration versus remnant scores* and *Interpreting novel species and novel habitats*). Furthermore, it made be necessary to calibrate "good" or "exceptional" site score thresholds differently depending on the habitat type and region within the state (see Chapter 1 & 2). Finally, it is important to emphasize that resultant quantiles would only be meaningful within states, as any absolute designations of good, high-quality, or regionally remarkable scores are only relevant within their state of origin. This is not only because FQA scores are not comparable among states, but because different regions have different land-use patterns (i.e. agricultural, urbanized), which would generate a different distribution in high versus low quality habitat scores.

Restoration versus remnant scores

A primary habitat restoration goal is to re-assemble plant communities in the image of remnants, which is an aspect of restoration success that FQA is well suited to measure (Matthews et al. 2009b, Matthews and Spyreas 2010). However, some have argued that because remnant plant communities are so ecologically integrated with their associated non-plant biotas (soil microbes, pollinators, etc.), environment, and histories, they form an irreplaceable whole that once degraded cannot be replaced with simple stock plantings (Swink and Wilhelm 1979, 1994b). Thus, highly Conservative plants require specific remnant habitat circumstances, such that they should not persist when planted into newly created habitats in sufficient numbers for restoration FQA scores to approach those of undegraded remnants. This irreplaceability argument— emphasizing that remnant FQA scores are more than the sum of their plant parts— is certainly an intriguing and non-reductive perspective. However, the simpler and more immediate question of whether FQA scores approach, or are capable of approaching, scores in remnants requires consideration. Original FQA authors suggested that restorations did not approach remnants based on the restorations they had seen: “Wetland or prairie reconstructions seldom exceed an FQI of 35, at least in the short term, and only do so with intensive efforts. The long-term potential or stability of many reconstructions has not been determined.” (see also, Swink and Wilhelm 1994c, Taft et al. 1997). Most subsequent studies have corroborated this assertion, as “terminal” restoration values rarely approach undegraded natural areas (Bowles and Jones 2006, Jog et al. 2006, Taft et al. 2006, Bowers and Boutin 2008), and FQA scores can distinguish restorations from remnants in most instances (Mushet et al. 2002, Andreas et al. 2004, Hossler et al. 2011).

On the other hand, some studies have observed restoration values high enough to suggest that parity with remnants levels can be achieved (Allison 2002), especially in a sub-set of meticulously managed restorations (Sperry 1994, Gardner 1995). In one prairie restoration, Anderson et al. (2007) reported plot level Mean *C* scores (4.8 - 5.4) that were higher than the highest-quality local remnant prairie (4.7). But, their restoration's FQI value was comparatively low due to a lack of appreciable species density:

> "… the high Mean C of our study site results from planting prairie species that generally have high coefficients of Conservatism and does not reflect the low species richness of our study site. Consequently, assessment of restorations requires several measurements (e.g. Mean C, species richness, diversity, and floristic quality measurements) to determine whether or not restoration goals are being achieved. Clearly, our site is lacking in species richness."

Another prairie restoration had site FQI (53.5), and Mean *C* (4) scores that were comparable with remnant prairies sampled (Slagle et al. 2008). But, these scores differ from remnants in their relatively low *plot-level* Mean *C*, FQI, and richness, leading the authors to conclude that "While the restoration methods used at the LREC [restoration] have promoted an overall level of Floristic Quality and presence of species of conservation value similar to that of remnants, native species are still missing in the system and individual plot data indicate their distribution is not yet similar to what would be seen at a remnant." Finally, a unique study of aquatic floras showed that man-made impoundments or reservoirs in Wisconsin (U.S.A) approached or exceeded scores in remnant lakes, leading to the conclusion that lake Floristic Quality is more related to water quality than to a the lakes origin (i.e., created versus remnant) (Nichols 1999). This example is particularly interesting because aquatic floras would have established without active introduction and restoration in these reservoirs. It is also true that the remnant lakes for comparison were somewhat degraded in most cases.

In summary, existing restorations are typically deficient in some aspect of their Floristic Quality, although some site level FQA scores can approach high-quality remnants in some instances. Caution is warranted when directly comparing restoration scores with remnants. For example, in many cases the highest quality remnant for a particular habitat type may be rather degraded. Or, the potential to overload plantings with high *C*-value species, and perhaps artificially inflate site values over the short-term, has been suggested. Furthermore, due to their somewhat artificially supported biotas, restored sites might warrant separate consideration when creating the distributions and thresholds needed to establish regional contexts for Floristic Quality (*Precision, sensitivity, and variability in site scores*).

Interpreting novel species and novel habitats

To this point, site FQA scores have been considered for created restorations and remnant habitats. But, how are habitats outside of these two categories to be considered? For example, reservoirs were singled out for their higher values compared to natural lakes (Nichols 1999). Spieles et al. (2006) describe a "borrow pit" dug over one-hundred years ago that became a high Floristic Quality wetland after it was abandoned. Novel, no-analog habitats such as these raise the question of how to evaluate habitats where no remnant reference exists (e.g., *Ranges and distributions in site scores*). Similar confusion can apply to specific plants species occurrences in novel habitats. For example, calciphilic plants with high *C*-values have colonized abandoned limestone quarries (Thompson and Green 2010). Therefore, where their native remnant habitat might have been a limestone bluff, dolomite prairie, or calcium-rich sand habitat they may now largely consist of artificial highly-disturbed habitats. Similarly, halophytic plants with high *C*-values that otherwise occur in brackish marshes, estuaries, or other alkaline native habitats such as groundwater seeps, have become widespread along highways due to road salting (e.g., *Solidago sempervirens* in midwestern North America). These species were assigned *C*-values based on their previous remnant habitat occurrences that they are no longer confined to.

There are several ways to deal with plant occurrences in novel-habitat types. The first could be to simply revise species scores to an average value between their native and novel habitat occurrences. The second more conservative approach is to score the species as in reference to their current lower habitat quality occurrences, and lack of remnant exclusivity. For example, this has been done for *Linum arenicola* (sand flax), an endemic

Florida calciphile that was given a score of one (1), because it is now commonly associated with disturbed areas containing exposed lime, such as tops of canal banks and road sides cut through lime-rock (Mortellaro et al. 2012). A similarly conservative approach is often used for species whose origin is in question, or where a species contains native and introduced genetic components (e.g. Phalaris arundinacea, Phragmites australis Spyreas et al. 2004, Bowers and Boutin 2008). Another option would be to allow species to have separate native and non-native plant *C*-value entries, reflecting occurrences in different habitats within a state:

> "When both native and non-native populations of a species (or native and non-native varieties) are known in Ohio, or where a species or variety is native in one part of Ohio but adventive elsewhere (e.g. *Campsis radicans*)… A few plants in the Ohio plant database are listed twice, once in normal typeface and once in CAPS [score of zero], and information is provided in the footnotes as to the status of native and non-native populations (e.g. *Najas marina, Phlox subulata, Pinus strobus*). When there are both native and non-native varieties of the same species, these may also be listed in the database twice, with the native variety in normal typeface and the non-native variety in CAPS (e.g. *Descurainia pinnata, Fragaria vesca, Symphoricarpus albus*). Finally, the issue of nativity in only one part of the state was often addressed in a footnote for that species (e.g. *Achillea millefolium…Thuja occidentalis*)." (Andreas et al. 2004)

Expanded Use and Future Considerations

A recognized goal for FQA is that nationwide (U.S.) comparisons become possible (Medley and Scozzafava 2009). However, the likelihood for continent-wide FQA integration is questionable, because FQA is based on region specific floral composition and behavior (see *Conservation* "hotspots" are often mapped and compared at global scales (Myers et al. 2000, Brooks et al. 2006). Such maps are often underpinned by ecological factors such as the global latitudinal diversity gradient (LDG) (Visser et al. 2014), and by anthropogenic land use patterns. However, comparisons of the conservation value of habitat patches across landscapes— the scales at which most conservation and restoration efforts happen— are not effectively made, because the means to do so are not well-developed (Stein 2002, Ejrnæs et al. 2008). Existing vegetation-based assessments of habitat conservation value at these scales have used endemic species, species diversity, phylogenetic diversity, rare species, indicator species, or indices that combine several metrics to rank areas (Mack et al. 2008, Kiera et al. 2009). But, such measures have not been readily adopted by conservation practitioners, either because they do not have a broad enough ecological scope (e.g., focus on specialized taxa or specific site properties), or because they are too difficult to generate, interpret, or compare across areas (Niemi and McDonald 2004, Pearman et al. 2006). Thus, generally applicable measures for assessing the conservation value of natural area vegetation, which are simple and flexible enough to be readily adopted by conservation practitioners, are wanting (Niemi and McDonald 2004, Taft et al. 2006).

It is in this context that Floristic Quality Assessment (FQA) has increased in use and influential over the past 30 years (especially in North America) (Appendix). A site's plant species list is used to calculate its Floristic Quality value with two primary metrics. The first is the mean of *a priori* conservation values assigned to each species. Species conservation values are termed Coefficients of Conservatism, or *C-values* (further described in Methods), and their average for a site is the Mean *C*. The second metric incorporates site species richness and Mean *C* into the Floristic Quality Index (FQI). Floristic Quality metrics (Mean *C* and FQI) are said to measure an area's biological Integrity, conservation value, or habitat degradation level via the amount of human induced changes to that area's plant assemblage (Taft et al. 1997, Spyreas and Matthews 2006, Mack 2007). FQA metrics are popular for a number of reasons: they perform well, being characterized as "remarkably robust" for their stated objectives (Miller et al. 2006); they capture information that is unique among existing conservation measures (Appendix); they only require a plant species list for their calculation; their values can be compared over time (Spyreas et al. 2012); and their values can easily be compared between all types of sites within a given region (FQA regions are most commonly states/provinces, Appendix).

The ability to compare varied sites across large regions with a single metric has spawned many different uses. For example, they are used in research on site restoration success and failure across vast areas (e.g., Mushet et al. 2002, Balcombe et al. 2005, Matthews et al. 2009a), and to study the causes and patterns in statewide habitat degradation (Nichols 2001, Milbauer and Leach 2007, Kraszewski and Waller 2008, Rooney and Leach 2010, Smart et al. 2011). FQA is also influential because of its direct use as a habitat evaluation criterion by various agencies. For example, the Nature Conservancy uses FQA to "prioritize high quality natural areas to include in their natural areas networks" (Rocchio 2007). Some federal, state, county, and municipal agencies have legally mandated that FQA metrics be used to evaluate sites across their jurisdictions. For example, a fixed benchmark FQI value threshold of 20 has been used by state and federal agencies as the determinant criteria for mitigating wetland destruction, as regulated under the Clean Water Act (Herman et al. 1997, Streever 1999, Herman et al. 2001, Matthews and Endress 2008, Matthews et al. 2009b). Such use assumes that a value of 20 is an equivalent and absolute value, such that any area's FQI value can be held up against this benchmark and its meaning will be equivalent. For example, any site scoring below 20 is considered easily replaceable and might receive less protection (Wilhelm 1992, Matthews et al. 2005).

Use of FQA across large regions has therefore been encouraged as objective and repeatable; a quantitative criterion that has encouraged "uniformity in natural area evaluation, enabling planners, land custodians, ecologists and other practitioners to make standardized comparisons among various open land areas" (Wilhelm and Ladd 1988). Region-wide metric uses assume that site values only vary due to anthropogenic factors (e.g., livestock grazing, logging, exotic invasive species, ecological management regimes)

(Spyreas and Matthews 2006, Milbauer and Leach 2007, Spyreas et al. 2010). Thus, Floristic Quality is otherwise assumed independent of location, even across large U.S. states and Canadian provinces (e.g., in Florida: Cohen et al. 2004, Ohio: Mack 2006, the Dakotas: Hargiss et al. 2007, Colorado: Rocchio 2007, Alberta: Raab and Bayley 2012).

Despite the assumption of regional neutrality, other conservation metrics and ecological indicators are inherently varied across regional ecological gradients (e.g., Pearman et al. 2006, Muratet et al. 2008). For example, plant species richness varies inversely with latitude at global scales, and it varies regionally in association with many other non-anthropogenic factors (e.g., productivity, climate, disturbance regime, glacial history, elevation, edaphic conditions, biogeography) (Mutke and Barthlott 2005, Sarr et al. 2005). As richness is a constitutive component of FQI, FQI is also expected to respond to spatial variation in these factors. On the other hand, Mean *C* is not tied to richness and it might be assumed to be independent of spatially varying ecological factors. The level to which FQA values vary due to regional ecological gradients that are unrelated to human disturbance, would need to be accounted for in their region-wide use (Bernthal 2003). If not, this would confound the use of FQA in making conservation decisions. For example, land acquisition and protection in regions where ecological characteristics naturally generate lower FQI values could be disfavored.

Studies have not yet determined if FQA values vary within a region. The largest existing study found latitudinal differences in FQI, Mean *C*, and species richness in ponds and lakes across Wisconsin (USA) (Nichols 1999). However, Nichols (1999) also pointed out that observed patterns could have simply been reflecting the different habitat-types sampled in different regions (oligotrophic versus eutrophic ponds; Chapter 1). Conversely, neither Cohen et al. (2004) nor Miller et al. (2006) found latitudinal differences in wetland FQA values. Finally, Johnston et al. (2010) found strong latitudinal variation in values from wetlands bordering the Great Lakes. Unfortunately, because they compared *C-values* from different states, their results must be viewed cautiously (i.e., because each state's floras have been assigned different *C-values* for a given species; Appendix). Additionally, because they only reported FQI values, it was not possible to determine if their observed patterns were due to differences in richness, Mean *C*, or both.

A considerable amount of data is required to make conclusions about statewide Floristic Quality patterns and regional trends in FQA metrics, and this has hindered their understanding thus far. If regional variation in FQA is to be understood, analyses *must* be conducted at statewide scales, because each species *C-value* is scored with respect to its *simultaneous behavior across the entire state* (see Methods). Therefore, because values in one area of a state might be counterbalanced by another part, assessments of regional FQA patterns must have complete statewide coverage. The following additional criteria for data exist if regional FQA patterns are to be understood: 1) habitat-type should be controlled for; 2) an unbiased site selection criterion with probabilistic sample design should be used (randomized, stratified-systematic, etc.) (Genet and Olsen 2006); 3) site sampling should be consistent (sample intensity, area, or season can affect FQI values, Appendix); 4) studies should examine more than wetlands or aquatic habitat-types.

). Wetland IBIs are currently being adapted for nationwide use and they may provide a template for how to develop a centralized system of this kind (U.S.E.P.A. 2010). Regardless of how such a nationwide undertaking unfolds, there is little doubt that Conservatism values will be assigned for more regions and that FQA will expand in its geographic coverage and use.

FQA in hypo- and hyper-diverse regions

FQA has only been applied to temperate regions with moderate plant diversity thus far. Its use assumes enough botanical knowledge to assign informed *C*-values to species, and it seems questionable whether it could be expanded to highly diverse areas such as humid tropical zones. As currently practiced, FQA assignments "… are impractical when dealing with the hundreds to thousands of species necessary to inform biodiversity and for taxonomic groups about which little knowledge exists" (Nielsen et al. 2007). Even *C*-value development in temperate North America has been hampered by a lack of knowledge of state floras (e.g., Mississippi, personal communication Gary Ervin). There has been one application of FQA to a sub-tropical region in Florida. However, its flora is comparatively well-studied and was not prohibitively diverse (Cohen et al. 2004, Reiss 2006).

Despite the impediments, there is a demonstrated need for metrics of this type in highly diverse habitats. For example, crucial differences in the conservation value of primary versus secondary tropical forests, they can be difficult to separate in some instances (Dent and Wright 2009, Gibson et al. 2011). Such inquiries could benefit from community composition based analytics like FQA. Its development in highly diverse areas might be facilitated in two ways. First, some level of abstracting species identification and *C*-value scoring to the level of genera would be necessary, because many records will not be identifiable to species. For example, where taxonomic expertise is not common in temperate regions (e.g., non-vascular plant taxa), *C*-values have been applied at the generic level (aquatic algae, Nichols 1999, mosses, Andreas et al. 2004). Also, the average *C*-value for the genus can be used in FQA calculations (see, *Plant species misidentification and detection*). Sivicek and Taft (2011) have suggested using functional group diversity to indicate Floristic Quality, where species diversity is unmanageable. But, this approach is sufficiently unique that it is probably best considered outside of FQA). Second, there would have to be some determination of the minimum percentage of samples requiring a positive species identification for an accurate site Floristic Quality measurement to be achieved. Some studies indicate that metric resolution can be high despite relatively low levels of species identification (Cohen et al. 2004, Rothrock and Homoya 2005).

At the other extreme, it might also be true that FQA could be ineffective when extended into species poor regions. Some regions may not have enough species diversity to produce site metric scores with sufficient resolution and sensitivity to discriminate degradation among sites. In such areas, alternative taxonomic groups might be considered. For example, bryophytes or lichens may be useful in augmenting vascular plants where they are abundant (e.g., boreal habitats), although they can present their own usage challenges (Boch S. et al. 2013). Regardless of concerns about the lack/overabundance of plant diversity when extending FQA to new regions, considerable portions of the globe have plant richness levels on par with existing FQA regions, as well as potentially having sufficient knowledge of their floras to make it's extension feasible (Kier et al. 2005).

FQA in hypo- and hyper-disturbed regions

FQA in North America was created with the expectation of a certain amount of regional anthropogenic disturbance:

> "Following the alteration and clearing of large portions of the landscape of North America after European settlement, much of the native biota has been restricted to sometimes small, often isolated, and usually somewhat modified remnants of the pre-settlement landscape. Many of these remnant sites are under pressure from continuing agricultural and urban development, but as natural areas they play an important role in maintaining landscape integrity and regional biodiversity." (Francis et al. 2000)

Thus far, FQA has only been applied in landscapes dominated by disturbance tolerant species. It could be difficult to develop for regions where degraded habitats are so rare that highly- and moderately-human disturbance associated plant species are not apparent, or where they occur infrequently (e.g., boreal regions, alpine zones). For example, some authors have mentioned the difficulty in assigning scores in some less disturbed habitat types and regions. For example, it could be difficult to scores the floras of peat bogs and lakes in northern Wisconsin, "Many species from low alkalinity lakes are assigned high C-values. Low alkalinity lakes are the least disturbed so we do not have as much experience determining the Conservatism of their plants. [However] We believe these species are highly sensitive to disturbance because of their rather tenuous existence" (Nichols 1999). Thus, determining species sensitivity in relatively pristine regions is difficult. In general however, the amount of the world's ice-free land area that is human-disturbed and "favourable to weeds" is so vast (Jenkins and Pimm 2003), that such a concern would be rare (Sanderson et al. 2002).

At the other extreme, some authors doubt FQA development is possible in long altered landscapes. For example, where "…defining the degree of naturalness of vegetation, especially in areas like [much of] Europe where the human activities have modified vegetation composition and structure for millennia (Andreas and Lichvar 1995)", assigning *C*-values could be difficult (Landi and Chiarucci 2010). The comparison is made with North America, where "most of the human impacts are only recent and it is relatively easier to define what is natural vegetation and what is not" (Andreas and Lichvar 1995). This problem has also been defined in terms of reference habitats, where finding a broad enough spectrum of habitats that includes undegraded remnant habitats could be difficult:

> "A general criticism of using ecological indicators is that finding an ecological "reference condition" is unlikely (Green 1979), particularly because few ecosystems remain under least-possible-impact conditions. If the least-possible impact conditions were found, it has been suggested that we could never be certain that other less impacted sites do not exist)." (Lopez and Fennessy 2002)

This concern seems overstated as FQA has successfully been applied in highly disturbed North American states. For example, across highly-agriculturalized Illinois, where over 99.99% of remnant prairies habitats have been destroyed (White and Madany 1978), there remains enough relatively undegraded patches that the remnant dependence of the prairie flora is easily discerned.

FQA and global climate change

The magnitude and speed at which global climate change (GCC) is occurring was not apparent when original FQA texts were written. Thus, the following directive might not have been made were it written today, “Once a framework of Coefficients of Conservatism is established for an area, the system provides a dispassionate, cost-effective and repeatable methodology. Anyone with a reasonable field knowledge of vascular plants, *now or a hundred years from now*, can apply these techniques and obtain comparable evaluations” (Wilhelm and Ladd 1988). As species ranges and distributions shift with GCC, species occurrences in remnant communities will change and potentially destabilize. Thus, Conservative species “ecological inertia” (sensu lato, Lopez and Fennessy 2002, Bowers and Boutin 2008) towards exclusive occurrence in remnant natural communities could be broken, as floras migrate and reassemble in novel ways. The very foundation that species *C*-values were assigned under could change as the nature of what a remnant native habitat is becomes less meaningful.

Alternatively, FQA metrics could become more effective under GCC. Highly Conservative species would seem to be the least able to migrate across a hostile anthropogenic matrix, so the likelihood for them to migrate and establish in novel habitats under the range shifts required with GCC seems improbable. This could increase their exclusivity to remnant habitats, compared to less-Conservative plants that are able to more easily migrate and form new communities. The net effect could make differences between Conservative less-Conservative plants even more stark. And, under the assumption that all species and communities would shift and re-assemble, site FQA scores could more become accurate at separating remnant from disturbed areas. Clearly the complexity of GCC effects on FQA is beyond the scope of this review, but attention must be paid to whether species *C*-values will necessitate regular updating, to remain accurate relative to existing landscape conditions.

Acknowledgements

I would like to thank the following people for editorial comments and proofreading, Katherine Chi, Valerie Sivicek, and Jeffrey W. Matthews. I would like to thank Tara Beveroth and Danielle Ruffatto for help creating figures.

References

Adamus, P., and K. Brandt. 1990. Impacts on quality of inland wetlands of the United States: a survey of indicators, techniques, and applications of community level Biomonitoring Data. U.S. Environmental Protection Agency.

Adamus, P., T. J. Danielson, and A. Gonyaw. 2001. Indicators for monitoring biological integrity of inland, freshwater wetlands: A survey of North American technical literature (1900-2000). U.S. Environmental Protection Agency, Washington D.C.

Allain, L., L. Smith, C. Allen, M. F. Vidrine, and J. B. Grace. 2006. A Floristic Quality Assessment system for the coastal prairie of Louisiana. Proceedings of the 19th North American Prairie Conference:1-18.

Allen, L., K. Anderson, E. DeVito, T. Halliwell, M. Leck, B. Olsen, B. Rawlyk, K. S. Walz, G. Wilhelm, and L. Sauer-Jones. Floristic quality assessment for the state of New Jersey. Plant stewardship index (PSI). Bowman's Hill Wildfower Preserve. Available online at http://www.bhwp.org/psi/index.php. Accessed 2011., New Hope, PA.

Allison, S. K. 2002. When Is a restoration successful? results from a 45-year-old tallgrass prairie restoration. Ecological Restoration **20**:10-17.

Anderson, E. 1971. Plants, man and life. University of California Press, Berkely, CA.

Anderson, K. J. 2007. Temporal patterns in rates of community change during succession. American Naturalist **169**:780-793.

Anderson, R. C., B. Dorick, and T. Crispino. 2007. Impacts of initial species richness and deer browsing on the quality of restored prairie in central Illinois. Ecological Restoration **25**:99-102.

Anderson, R. C., D. Nelson, M. R. Anderson, and M. Rickey. 2006. White-tailed deer (*Odocoileus virginianus* Zimmermann) browsing effects on quality of tallgrass prairie community forbs. Pages 63–68 *in* Proceedings of the Nineteenth North American Prairie Conference. University of Wisconsin Press, Madison, WI.

Andreas, B. K., and R. W. Lichvar. 1995. Floristic index for establishing assessment standards: a case study for northern Ohio. Report to U.S. Army Corps of Engineers Waterways Experiment Station. Vicksburg, MS:14p.

Andreas, B. K., J. Mack, and J. S. McCormac. 2004. Floristic Quality Assessment Index (FQAI) for vascular plants and mosses for the State of Ohio. Report to the Ohio

Environmental Protection Agency, Division of Surface Water, Wetland Ecology Group, Columbus, Ohio:219p.

Bailey, R. G. 1995. Descriptions of the ecoregions of the United States. U.S. Forest Service, Washington, D.C.

Balcombe, C. K., J. T. Anderson, R. H. Fortney, J. S. Rentch, W. N. Grafton, and W. S. Kordek. 2005. A comparison of plant communities in mitigation and reference wetlands in the mid-Appalachians. Wetlands **25**:130-142.

Bernthal, T. 2003. Development of a floristic quality assessment methodology for Wisconsin. Report to the U.S. Environmental Protection Agency Region V, Chicago, IL, USA:19p.

Betz, R. F. 1988. Ecology of Mead's milkweed (*Asclepias meadii* Torrey). Pages 187-191. *in* Proceedings of the Eleventh North American Prairie Conference. University of Nebraska Printing, Lincoln, NE.

Bever, J., P. A. Schultz, R. M. Miller, L. Gades, and J. D. Jastrow. 2003. Inoculation with prairie mycorrhizal fungi may improve restoration of native prairie plant diversity. Ecological Restoration **21**:311–312.

Boch S., Müller J., Prati D., Blaser S., and Fischer M. 2013. Up in the tree – the overlooked richness of bryophytes and lichens in tree crowns. PLoS ONE **8**.

Bonanno, G., and R. L. Giudice. 2010. Application of two quality indices as monitoring and management tools of rivers. Case study: the Imera Meridionale River, Italy. Environmental Management **45**:856-867.

Boughton, E. H., P. F. Quintana-Ascencio, P. J. Bohlen, D. G. Jenkins, and R. Pickert. 2010. Land-use and isloation interact to affect wetland plant assemblages. Ecography **33**:461-470.

Bourdaghs, M. 2004. Properties and performance of the Floristic Quality Index in Great Lakes coastal wetlands. University of Minnesota, Minneapolis, MN.

Bourdaghs, M., C. A. Johnston, and R. R. Regal. 2006. Properties and performance of the Floristic Quality Index in Great Lakes coastal wetlands. Wetlands **26**:718-735.

Bowers, K., and C. Boutin. 2008. Evaluating the relationship between floristic quality and measures of plant biodiversity along stream bank habitats. Ecological Indicators **8**:466-475.

Bowles, J. M., and M. Jones. 2006. Testing the efficacy of species richness and Floristic Quality Assessment of quality, temporal change, and fire effects in tallgrass prairie natural areas. Natural Areas Journal **26**:17-30.

Braun-Blanquet, J. 1932. Plant sociology - the study of plant communities. McGraw-Hill Inc., New York, NY.

Braun, L. E. 1950. Deciduous forests of eastern North America. Macmillan Publishing Co., Inc., NY, NY.

Brewer, J. S. 2010. A potential conflict between preserving regional plant diversity and biotic resistance to an invasive grass, *Microstegium vimineum*. Natural Areas Journal **30**:279-293.

Bried, J. T., and G. J. Edinger. 2009. Baseline floristic assessment and classification of pine barrens vernal ponds. Journal of the Torrey Botanical Society **136**:128-136.

Bried, J. T., S. T. Jog, and J. W. Matthews. 2013. Floristic quality assessment signals human disturbance over natural variability in a wetland system. Ecological Indicators **34**:260-267.

Bried, J. T., K. L. Strout, and T. Portante. 2012. Coefficients of Conservatism for the vascular flora of New York and New England: inter-state comparisons and expert opinion bias. Northeastern Naturalist **19**:101-114.

Brooks, T. M., R. A. Mittermeier, G. A. B. d. Fonseca, J. Gerlach, M. Hoffmann, J. F. Lamoreux, C. G. Mittermeier, J. D. Pilgrim, and A. S. L. Rodrigues. 2006. Global biodiversity conservation priorities. Science **313**.

Brudvig, L. A., C. Mabry, J. R. Miller, and T. A. Walker. 2007. Evaluation of central North American prairie management based on species diversity, life form, and individual species metrics. Conservation Biology **21**:864-874.

Brunet, J., and G. von Oheimb. 1998. Migration of vascular plants to secondary woodlands in southern Sweden. Journal of Ecology **86**:429-438.

Burke, D. M., and E. Nol. 1998. Edge and fragment size effects in the vegetation of deciduous forests in Ontario, Canada. Natural Areas Journal **18**:45-53.

Carroll, C., C. Dassler, J. Ellis, G. Spyreas, J. B. Taft, and K. Robertson. 2002. Plant sampling protocols. Pages 38p. + Figures, Tables, and Appendix *in* B. Molano-Flores, editor. Critical Trends Assessment Program Monitoring Protocols, Illinois Natural History Survey, Office of the Chief Technical Report 2002-2, Champaign, IL.

Carter, D. L., and J. M. Blair. 2012. Recovery of native plant community characteristics on a chronosequence of restored prairies seeded into pastures in west-central Iowa. Restoration Ecology **20**:170-179.

Chamberlain, S. J., and H. M. Ingram. 2012. Developing coefficients of conservatism to advance floristic quality assessment in the Mid-Atlantic region. Jounal of the Torrey Botanical Society **139**:416-427.

Chapman, M. G., A. J. Underwood, and K. R. Clarke. 2009. New indices for ranking conservation sites using 'relative endemism'. Biological Conservation **142**:3154-3162.

Chase, J. M., and M. A. Leibold. 2003. Ecological Niches: linking classical and contemporary approaches. University of Chicago Press Books, Chicago, IL.

Chen, H., H. Qian, G. Spyreas, and M. Crossland. 2010. Native-exotic species richness relationships across spatial scales and biotic homogenization in wetland plant communities of Illinois, USA. Diversity and Distributions **16**:737-743.

Christensen, N. L., and R. K. Peet. 1984. Convergence during secondary forest succession. Journal of Ecology **72**:25-36.

Chu, S., and B. Molano-Flores. 2013. Impacts of agricultural to urban land-use change on Floristic Quality Assessment indicators in Northeastern Illinois wetlands. Urban Ecosystems **16**:235-246.

Clavel, J., R. Julliard, and V. Devictor. 2011. Worldwide decline of specialist species: toward a global functional homogenization? Frontiers in Ecology and Environment **9**:222-228.

Cline, D., C. Juricek, R. G. Lym, and D. R. Kirby. 2008. Leafy spurge (*Euphorbia esula*) control with *Apthona* spp. affects seedbank composition and native grass establishment. Invasive Plant Science and Management **1**:120-132.

Cohen, M. J., S. Carstenn, and C. R. Lane. 2004. Floristic quality indices for biotic assessment of depressional marsh condition in Florida. Ecological Applications **14**:784-794.

Collins, B. S., and S. T. Pickett. 1987. Influence of canopy opening on the environment and herb layer in a northern hardwood forest. Vegetatio **70**:3-10.

Council, N. R. 2000. Ecological indicators for the nation. Washington, DC. National Academies Press. http://www.nap.edu/openbook/0309068452/html/.

Cretini, K. F., J. M. Visser, K. W. Krauss, and G. D. Steyer. 2012. Development and use of a Floristic Quality Index for coastal Louisiana marshes. Environmental Monitoring and Assessment **184**:2389-2403.

Curtis, J. T. 1959. The vegetation of Wisconsin: An ordination of plant communities. University of Wisconsin Press, Madison, WI.

Davis, A. S., and D. A. Landis. 2011. Agriculture. Pages 7-11 *in* D. Simberloff and M. Rejmanek, editors. Encyclopedia of Biological Invasions. University of California Press, Los Angeles.

Davis, J. D., S. D. Hendrix, D. M. Debinski, and C. J. Hemsley. 2008. Butterfly, bee and forb community composition and cross-taxon incongruence in tallgrass prairie fragments. Journal of Insect Conservation **12**:69-79.

De Cáceres, M., D. Sol, O. Lapiedra, and P. Legendre. 2011. A framework for estimating niche metrics using the resemblance between qualitative resources. Oikos **120**:1341-1350.

DeKeyser, E., D. R. Kirby, and M. J. Ell. 2003. An index of plant community integrity: development of the methodology for assessing prairie wetland plant communities. Ecological Indicators **3**:119-133.

Dent, D. H., and S. J. Wright. 2009. The future of tropical species in secondary forests: a quantitative review. Biological Conservation **142**:2833-2843.

Devictor, V., J. Clavel, R. Julliard, S. Lavergne, D. Mouillot, W. Thuiller, P. Venail, S. Villeger, and N. Mouquet. 2010. Defining and measuring ecological specialization. Journal of Applied Ecology **47**:15-25.

Devictor, V., R. Julliard, and F. Jiguet. 2008. Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. Oikos **117**:507-514.

Diamond, J. M. 1976. Island biogeography and conservation: strategy and limitation. Science **193**:1027-1029.

Diekmann, M. 2003. Species indicator values as an important tool in applied plant ecology– a review. Basic and Applied Ecology **2**:493-506.

Drayton, B., and R. B. Primack. 1996. Plant species lost in an isolated conservation area in metropolitan Boston from 1894-1993. Conservation Biology **10**:30-39.

Edgin, B. R., S. J. Adams, E. J. Stork, and J. E. Ebinger. 2010. Flora and species dominance of Green Prairie, Jasper County, Illinois- a glacial drift hill prairie on Illinoian TIll. Erigenia **23**:24-33.

Ejrnæs, R., J. Liira, R. S. Poulsen, and B. Nygaard. 2008. When has an abandoned field become a semi-natural grassland or heathland? Environmental Management **42**:707-716.

Ellenberg, H. 1979. Zeigerwerte der Gefäszpflanzen Mitteleuropas. 2nd ed., Scripta Geobotanica 9, Göttingen.

Engelman, L. 2005. Nonlinear models. Pages 541-594 SYSTAT 11 statistics II. SYSTAT, Port Richmond, California, USA.

Ervin, G., B. Herman, J. Bried, and D. C. Holly. 2006. Evaluating non-native species and wetland indicator status as components of wetlands floristic assessment. Wetlands **26**:1114-1129.

Euliss, N. H., and D. M. Mushet. 2011. A multi-year comparison of IPCI scores for prairie pothole wetlands: implications for temporal and spatial variation. Wetlands **31**:713-723.

Fennessy, S., R. Geho, B. Elifritz, and R. D. Lopez. 1988. Testing the Floristic Quality Assessment Index as an indicator of riparian wetland disturbance. .

Fennessy, S., M. Gray, R. D. Lopez, and J. Mack. 1998. An ecological assessment of wetlands using reference sites. Ohio EPA, Columbus, OH.

Fennessy, S., and J. Roehrs. 1997. A functional assessement fo mitigation wetlands in Ohio: comparisons with natural systems. Columbus, OH.

Fiedler, P. L., and D. A. Landis. 2012. Biotic and abiotic conditions in Michigan prairie fen invaded by glossy buckthorn (*Frangula alnus*). Natural Areas Journal **32**:41-53.

Filippi-Codaccioni, O., V. Devictor, Y. Bas, and R. Julliard. 2010. Toward more concern for specialisation and less for species diversity in conserving farmland biodiversity. Biological Conservation **143**:1493-1500.

Findlay, C. S., and J. Houlahan. 1997. Anthropogenic correlates of species richness in southeastern Ontarion wetlands. Conservation Biology **11**:1000-1009.

Fleishman, E., R. F. Noss, and B. R. Noon. 2006. Utility and limitations of species richness metrics for conservation planning. Ecological Indicators **6**:543-553.

Flory, S. L., and K. Clay. 2010. Non-native grass invasion suppresses forest succession. Oecologia **164**:1029-1038.

Forrest, A. 2010. Development, analysis and application of a Floristic Quality Index (FQI) in Alberta's Boreal and Central Parkland Natural Regions. M.Sc. Thesis. University of Alberta, Edmonton, Alberta, Canada, 117 pp.

Foster, B. L., C. A. Murphy, K. R. Keller, T. A. Aschenbach, E. Questad, and K. Kindscher. 2007. Restoration of prairie community structure and ecosystem function in an abandoned hayfield: a sowing experiment. Restoration Ecology **15**:652-661.

Francis, C. M., M. J. W. Austen, J. M. Bowles, and W. B. Draper. 2000. Assessing floristic quality in southern Ontario woodlands. Natural Areas Journal **20**:66-77.

French, S., and B. Robson. 2012. What is a conservative substitution? Journal of Molecular Evolution **19**:171-175

Fridley, J. D., D. B. Vandermas, D. M. Kuppinger, M. Manthey, and R. K. Peet. 2007. Co-occurrence based assessment of habitat generalists and specialists: a new approach for the measurement of niche width. Journal of Ecology **95**:707-722.

Frieswyk, C. B., and J. B. Zedler. 2006. Do seed banks confer resilience to coastal wetlands invaded by Typha x glauca? Canadian Journal of Botany **84**:1882-1893.

Futuyma, D. J., and G. Moreno. 1988. The evolution of ecological specialisation. Annual Review of Ecology and Systematic **19**:207-233.

Galatowitsch, S. M., N. O. Anderson, and P. D. Ascher. 1999. Invasiveness in wetland plants in temperate North America. Wetlands **19**:733-755.

Garcıa-Madrid, A. S., J. A. Molina, and P. Canto. 2014. Classification of habitats highlights priorities for conservation policies: the case of Spanish Mediterranean tall humid herb grasslands. Journal for Nature Conservation **22**:142-156.

Gardner, D. 1995. Prairie restoration on an east-central Illinois field. Erigenia **14**:18-31.

Genet, J. A., and A. R. Olsen. 2006. Assessing the quantity and quality of depressional wetlands in the Redwood River Watershed utilizing a probabilistic sampling design. Minnesota Pollution Control Agency, St. Paul, MN:1-60.

Gerken, M. E., J. R. Thompson, C. Mabry, and R. K. Kolka. 2013. An Investigation of water nutrient levels associated with forest vegetation in highly altered landscapes. The Journal of Soil and Water Conservation **68**:361-371.

Gibson, C. W. D., and V. K. Brown. 1991. The nature and rate of development of calcareous grasslands in Southern Britain. Biological Conservation **58**:297-316.

Gibson, D. J., G. Spyreas, and J. Benedict. 2002. Life history of *Microstegium vimineum* (Poaceae), an invasive grass in southern Illinois. Journal of the Torrey Botanical Society **129**:207-219.

Gibson, L., T. M. Lee, L. P. Koh, B. W. Brook, T. A. Gardner, J. Barlow, C. A. Peres, C. J. A. Bradshaw, W. F. Laurance, T. E. Lovejoy, and N. S. Sodhi. 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. Nature **478**:378-381.

Gilliam, F. S., N. L. Turrill, and M. B. Adams. 1995. Herbaceous-layer and overstory species in clear-cut and mature central Appalachian hardwood forests. Ecological Applications **5**:947-955.

Gleason, H. A. 1909. Some unsolved problems of the prairies. Bulletin of the Torrey Botanical Club **36**:265-271.

Godefroid, S., and E. D. Dana. 2007. Can Ellenberg's indicator values for Mediterranean plants be used outside their region of definition? Journal of Biogeography **34**:62-68.

Godefroid, S., C. Piazza, G. Rossi, S. Buord, A. D. Stevens, R. Aguraiuja, C. Cowell, C. W. Weekley, G. Vogg, J. M. Iriondo, I. Johnson, B. Dixonm, D. Gordon, S. Magnanon, B. Valentin, K. Bjureke, R. Koopman, M. Vicens, M. Virevaire, and T. Vanderborght. 2011. How successful are plant species reintroductions. Biological Conservation **144**:672–682.

Goldblum, D., B. P. Glaves, L. S. Rigg, and B. Kleiman. 2013. The impact of seed mix weight on diversity and species composition in a tallgrass prairie restoration planting, Nachusa Grasslands, Illinois, USA. Ecological Restoration **31**:154-167.

Gotelli, N. J., and A. M. Ellison. 2012. A Primer of Ecological Statistics. 2nd edition. Sinauer Associates, Inc., Sunderland, MA.

Gould, F. W. 1937. The present status of the Dame County prairie flora. University of Wisconsin, Madison, WI.

Grime, J. P. 1974. Vegetation classification by reference to strategies. Nature **250**:26-31.

Grime, J. P. 1979. Plant strategies and vegetation processes. Wiley, New York, New York.

Grondin, P., S. Gauthier, D. Borcard, Y. Bergeron, and J. Noel. 2014. A new approach to ecological land classification for the Canadian boreal forest that integrates disturbances. Landscape Ecology **29**:1-16.

Gutrich, J. L., and F. J. Hitzhusen. 2004. Assessing the substitutability of mitigation wetlands for natural sites: estimating restoration lag costs of wetland mitigation. Ecological Economics **48**:409-424.

Haan, N. L., M. R. Hunter, and M. D. Hunter. 2012. Investigating predictors of plant establishment during roadside restoration. Restoration Ecology **20**:315-321.

Hargiss, C. L. M., E. DeKeyser, D. R. Kirby, and M. J. Ell. 2007. Regional assessment of wetland plant communities using the index of plant community integrity. Ecological Indicators **8**:303-307.

Hatzenbeler, G. R., J. M. Kampa, and M. J. Jennings. 2004. A comparison of fish and aquatic plant assemblages to assess ecological health of small Wisconsin lakes. Lake and Reservoir Management **20**:211-218.

Hawkins, C. P. 2006. Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessments. Ecological Applications **16**:1277-1294.

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

Herman, B. 2005. Testing the Floristic Quality Assessment Index in natural and created wetlands in Mississippi, USA. Mississippi State University.

Herman, K. D., L. A. Masters, A. A. Penskar, A. Reznicek, G. Wilhelm, and W. Brodowicz. 1997. Floristic Quality Assessment: development and application in the state of Michigan (USA). Natural Areas Journal **17**:265-279.

Herman, K. D., L. A. Masters, M. R. Penskar, A. Reznicek, G. Wilhelm, W. W. Brodovich, and K. P. Gardiner. 2001. Floristic Quality Assessment with wetland categrories and examples of computer applications for the state of Michigan. Revised, 2nd Edition. Michigan Department of Natural Resources, Wildlife Division, Natural Heritage Program, East Lansing, MI:19p. + Appendices.

Hermy, M., O. Honnay, L. Firbank, C. Grashof-Bokdam, and J. E. Lawesson. 1999. An ecological comparison between ancient and other forest plant species of Europe, and the implications for forest conservation. Biological Conservation **91**:9-22.

Hill, M. O., D. B. Roy, J. O. Mountford, and R. G. H. Bunce. 2000. Extending Ellenberg's indicator values to a new area: an algorithmic approach. Journal of Applied Ecology **37**:3-15.

Hill, M. O., D. B. Roy, and K. Thompson. 2002. Hemeroby, urbanity and ruderality: bioindicators of disturbance and human impact. Journal of Applied Ecology **39**:708-720.

Hilsenhoff, W. L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist **20**:31-39.

Hobbs, R. J., S. Arico, J. Aronson, J. S. Baron, P. Bridgewater, V. A. Cramer, P. R. Epstein, J. J. Ewel, C. A. Klink, A. E. Lugo, D. Norton, D. Ojima, D. Richardson, E. W. Sanderson, F. Valladares, M. Vila, R. Zamora, and M. Zobel. 2006. Novel ecosystems: theoretical and management aspects of the new ecological world order. Global Ecology and Biogeography **15**:1-7.

Hobbs, R. J., E. Higgs, and J. A. Harris. 2009. Novel ecosystems: implications for conservation and restoration. Trends in Ecology and Evolution **24**:599-605.

Hobbs, R. J., and D. A. Norton. 1996. Toward a conceptual framework for restoration ecology. Restoration Ecology **4**:93-110.

Hodgson, J. G. 1986. Commoness and rarity in plants with special reference to the Sheffield flora: Part 1: the identity, distribution and habitat characteristics of the common and rare species. Biological Conservation **36**:199-251.

Hoekstra, J. M., T. M. Boucher, T. H. Ricketts, and C. A. Roberts. 2005. Confronting a biome crisis: global disparities of habitat loss and protection. Ecology Letters **8**:23-29.

Holt, R. D. 2009. Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. Proceedings of the National Academy of Sciences of the United States of America **106**:19659-19665.

Honnay, O., B. DeGroote, and M. Hermy. 1998. Ancient-forest plant species in western Belgium: a species list and possible ecological mechanisms. Belgian Journal of Botany **130**:139-154.

Honnay, O., M. Hermy, and P. Coppin. 1999. Effects of area, age, and diversity of forest patches in Belgium on plant species richness, and implications for conservation and reforestation. Biological Conservation **87**:73-84.

Hopple, A., and C. Craft. 2013. Managed disturbance enhances biodiversity of restored wetlands in the agricultural Midwest. Ecological Engineering **61**:505–510.

Hossler, K., V. Bouchard, M. S. Fennessy, S. D. Frey, E. Anemaet, and E. Herbert. 2011. No-net-loss not met for nutrient function in freshwater marshes: recommendations for wetland mitigation policies. Ecosphere **2**:1-36.

Hough, P., and M. Robertson. 2009. Mitigation under Section 404 of the Clean Water Act: where it comes from, what it means. Wetlands Ecology and Management **17**:15-33.

Illinois Department of Natural Resources. 1994. The Changing Illinois Environment: Critical Trends. Vol. 3: Ecological Resources. Illinois Department of Natural Resources, Springfield, IL ILENR/RE-94/05.

Iverson, L. R., A. Prasad, and D. M. Ketzner. 1997. A summary of the Illinois flora based on the Illinois Plant Information Network. Transactions of the Illinois State Academy of Science **90**:41-64.

Izco, J. 1998. Types of rarity of plant communities. Journal of Vegetation Science **9**:641-646.

Jackson, P. 2009. Temporal and spatial vegetation dynamics in a remnant oak savanna: Middlefork Savanna, Lake County, Illinois. Erigenia **22**:9-35.

Jenkins, C. N., and S. L. Pimm. 2003. How big is the global weed patch? Annals of the Missouri Botanical Garden **90**:172-178.

Jog, S., K. Kindscher, E. Questad, B. Foster, and H. Loring. 2006. Floristic quality as an indicator of native species diversity in managed grasslands. Natural Areas Journal **26**:149-167.

Johnston, C. A., D. M. Ghioca, M. Tulbure, B. L. Bedford, M. Bourdaghs, C. B. Frieswyk, L. Vaccaro, and J. B. Zedler. 2008. Partitioning vegetation response to anthropogenic stress to develop multi-taxa wetland indicators. Ecological Applications **18**:983-1001.

Johnston, C. A., J. B. Zedler, and M. Tulbure. 2010. Latitudinal gradient of floristic condition among Great Lakes coastal wetlands. Journal of Great Lakes Research **36**:772-779.

Johnston, C. A., J. B. Zedler, M. G. Tulbure, C. B. Frieswyk, B. L. Bedford, and L. Vaccaro. 2009. A unifying approach for evaluating the condition of wetland plant communities and identifying related stressors. Ecological Applications **19**:1739-1757.

Jones, W. M. 2004. Using vegetation to assess wetland condition: a multimetric approach for temporarily and seasonally flooded depressional wetlands and herbaceous-dominated intermittent and ephemeral riverine wetlands in the northwestern glaciated plains ecoregion, Montana. Natural Resource Information System, Montana State Library, Montana Natural Heritage Program. Prepared for the Montana Dept. Env. Quality and U.S. E.P.A.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries **6**:21-27.

Karr, J. R. 1987. Biological monitoring and environmental assessment: a conceptual framework. Environmental Management **11**:249-256.

Kartesz, J. T. 2014a. *Carex*. The Biota of North America Program (BONAP). North American Plant Atlas. (http://www.bonap.org/napa.html). Chapel Hill, N.C. [maps generated from Kartesz, J.T. 2013. Floristic Synthesis of North America, Version 1.0. Biota of North America Program (BONAP).

Kartesz, J. T. 2014b. Moisture index. The Biota of North America Program (BONAP). North American Plant Atlas. (http://www.bonap.org/napa.html). Chapel Hill, N.C. [maps generated from Kartesz, J.T. 2013. Floristic Synthesis of North America, Version 1.0. Biota of North America Program (BONAP).

Keddy, P. A., and P. MacLellan. 1990. Centrifugal organization in forests. Oikos **59**:75-84.

Kercher, S. M., Q. J. Carpenter, and J. B. Zedler. 2004. Interrelationships of hydrologic disturbance, reed canary grass (Phalaris arundinacea L.), and native plants in Wisconsin wet meadows. Natural Areas Journal **24**:316-325.

Kier, G., J. Mutke, E. Dinerstein, T. H. Ricketts, W. Kuper, H. Kreft, and W. Barthlott. 2005. Global patterns of plant diversity and floristic knowledge. Journal of Biogeography **32**:1107-1116.

Kiera, G., H. Krefta, T. M. Leeb, W. Jetz, P. L. Ibisch, C. Nowickic, J. Mutkea, and W. Barthlotta. 2009. A global assessment of endemism and species richness across island and mainland regions. Proceeding of the National Academy of Sciences **106**:9322-9327.

Kindscher, K., and L. Tieszen. 1998. Floristic and soil organic matter changes after five and thirty-five years of native tallgrass prairie restoration. Restoration Ecology **6**:181-196.

Kirkman, L. K., K. L. Coffey, R. J. Mitchell, and E. B. Moser. 2004. Ground cover recovery patterns and life-history traits: implications for restoration obstacles and opportunities in a species-rich savanna. Journal of Ecology **92**:409-421.

Kowalski, K. P., and D. A. Wilcox. 2003. Differences in sedge fen vegetation upstream and downstream from a managed impoundment. American Midland Naturalist **150**:199-220.

Kraszewski, S. E., and D. M. Waller. 2008. Fifty-five year changes in species composition on dry prairie remnants in south-central Wisconsin. Journal of the Torrey Botanical Society **135**:235-243.

Kruckeberg, A. R. 1984. California serpentines: flora, vegetation, geology, soils and management problems. University of California Press, Berkeley, California.

Kucharik, C. J., J. A. Foley, C. Delire, V. A. Fisher, M. T. Coe, J. Lenters, C. Young-Molling, N. Ramankutty, J. M. Norman, and S. T. Gower. 2000. Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance and vegetation structure. Global Biogeochemical Cycles **14**:795-825.

Ladd, D. 1993. The Missouri floristic quality assessment system. The Nature Conservancy, St. Louis, MO.

Landi, S., and A. Chiarucci. 2010. Is floristic quality assessment reliable in human-managed ecosystems? Systematics and Biodiversity **8**:269-280.

LaPaix, R., B. Freedman, and D. Patriquin. 2009. Ground vegetation as an indicator of ecological integrity. Environmental Reviews **17**:249-265.

Laroche, V., S. Pellerin, and L. Brouillet. 2012. White fringed orchid as indicator of sphagnum bog integrity. Ecological Indicators **14**:50-55.

Leitner, L. A., J. H. Idzikowski, and G. S. Casper. 2008. Urbanization and ecological change in Milwaukee County.*in* D. M. Waller and T. P. Rooney, editors. The Vanishing Present: Wisconsin's changing lands, waters, and wildlife. University of Chicago Press, Chicago, IL.

Lichvar, R. W., and J. T. Kartesz. 2012. North American Digital Flora: National Wetland Plant List, version 3.0 http://wetland_plants.usace.army.mil.

Liira, J., T. Sepp, and O. Parrest. 2007. The forest structure and ecosystem quality in conditions of anthropogenic disturbance along productivity gradient. Forest Ecology and Management **250**:34-36.

Lindenmayer, D. B., and W. Laurence. 2012. History of hubris: cautionary lessons in ecologically sustainable forest management. Biological Conservation **151**:11-16.

Lopez, R. D., and M. S. Fennessy. 2002. Testing the floristic quality assessment index as an indicator of wetland condition. Ecological Applications **12**:487-497.

MacAllister, B. A., and S. B. Nemeth. 2009. Floristic Quality Assessments: public works technical bulletin 200-2-65. U.S. Army Engineer and Research Development Center Construction Engineering Research Laboratory. U.S. Army Corps of Engineers., Champaign, IL.

Mace, G. M. 2005. An index of intactness. Nature **434**:32-33.

Mack, J. 2006. Landscape as a predictor of wetland condition: an evaluation of the landscape development index (LDI) with a large reference wetland dataset from Ohio. Environmental Monitoring and Assessment **120**:221-241.

Mack, J. 2007. Developing a wetland IBI with statewide application after multiple testing iterations. Ecological Indicators **7**:864-881.

Mack, J. 2009. Development issues in extending plant-based IBIs to forested wetlands in the Midwestern United States. Wetlands Ecology and Management **17**:117-130.

Mack, J., N. Avdis, E. Braig, and D. Johnson. 2008. Application of a vegetation-based Index of Biotic Integrity for Lake Erie coastal marshes in Ohio. Aquatic Ecosystem Health & Management **11**:91-104.

Mack, J. J. 2001. Vegetation Indices of Biotic Integrity (VIBI) for wetlands: ecoregional, hydrogeomorphic, and plant community comparision with preliminary wetland aquatic life use designations. Final report to U.S. EPA. Wetland Ecology Group, Division of Surface Water, Ohio Environmental Protection Agency, Columbus, Ohio. 90 p. http://www.epa.state.oh.us/dsw/wetlands/.

Mack, J. J., M. Micacchion, L. D. Augusta, and G. R. Sablak. 2000. Vegetation Indices of Biotic Integrity (VIBI) for wetlands and calibration of the Ohio Rapid Assessment Method for wetlands version 5.0. Final report to U .S. EPA. Wetland Ecology Group, Division of Surface Water, Ohio Environmental Protection Agency, Columbus.

Malik, R. N., Z. K. Shinwari, and H. Waheed. 2012. Linkages between spatial variations in riparian vegetation and Floristic Quality to the environmental heterogeneity: A case study of River Soan and its associated streams, Pakistan. Pakistan Journal of Botany **44**:187-197.

Marcum, P., D. T. Busemeyer, L. R. Phillippe, and J. E. Ebinger. 2010. Vascular flora and woody plant structure and composition at Gooseberry Island Nature Preserve, Kankakee County, Illinois. Castanea **75**:341-352.

Marks, P. L. 1983. On the origin of the field plants of the northeastern United States. The American Naturalist **122**:210-228.

Martin, L. J., and B. Blossey. 2012. Invasive plant cover impacts the desirability of lands for conservation acquisition. Biodiversity and Conservation **21**:1987-1996.

Martin, P. H., C. D. Canham, and P. L. Marks. 2009. Why forests appear resistant to exotic plant invasions: intentional introductions, stand dynamics, and the role of shade tolerance. Frontiers in Ecology and Environment **7**:142-149.

Masters, L. A. 1997. Monitoring Vegetation. Pages 279-301 *in* S. Packard and C. Mutel, editors. The tallgrass restoration handbook: for prairies, savannas, and woodlands. Island Press, Washington, D.C.

Matlack, G. R. 1994. Plant species migration in a mixed history forest landscape in eastern North America. Ecology **75**:1491-1502.

Matthews, J. A. 1979. A study of the variability of some successional and climax plant assemblage-types using multiple discriminant analysis. Journal of Ecology **67**:255-271.

Matthews, J. W. 2003. Assessment of the Floristic Quality Index for use in Illinois, USA, wetlands. Natural Areas Journal **23**:53-60.

Matthews, J. W., and A. G. Endress. 2008. Performance criteria, compliance success and vegetation development in compensatory mitigation wetlands. Environmental Management **41**:130-141.

Matthews, J. W., A. L. Peralta, D. N. Flanagan, P. M. Baldwin, A. Soni, A. D. Kent, and A. G. Endress. 2009a. Relative influence of landscape vs. local factors on plant community assembly in restored wetlands. Ecological Applications **19**:2108-2123.

Matthews, J. W., and G. Spyreas. 2010. Convergence and divergence in plant community trajectories as a framework for monitoring wetland restoration progress. Journal of Applied Ecology **47**:1128-1136.

Matthews, J. W., G. Spyreas, and A. G. Endress. 2009b. Trajectories of vegetation-based indicators used to assess wetland restoration. Ecological Applications **19**:2093-2107.

Matthews, J. W., G. Spyreas, and C. M. Long. In Review. A null model approach for testing the validity of plant species' Coefficients of Conservatism. Ecological Indicators **XX**:xxx-xxx.

Matthews, J. W., P. A. Tessene, S. M. Wiesbrook, and B. W. Zercher. 2005. Effect of area on isolation on species richness and indices of floristic quality in Illinois, USA wetlands. Wetlands **25**:607-615.

McCarthy, B. C. 2003. The herbaceous layer of eastern old-growth deciduous forests. Pages 163-176 *in* F. S. Gilliam and M. R. Roberts, editors. The Herbaceous Layer in Forests of Eastern North America. Oxford University Press, New York.

McClain, W. E., T. L. Esker, B. R. Edgin, G. Spyreas, and J. E. Ebinger. 2010. Fire history of a post oak (Quercus stellata Wang.) woodland in Hamilton County, Illinois. Castanea **75**:461-474.

McDonald, R., M. McKnight, D. Weiss, E. Selig, M. O'Connor, C. Violin, and A. Moody. 2005. Species compositional similarity and ecoregions: Do ecoregion boundaries represent zones of high species turnover? Biological Conservation **126**:24-40.

McIndoe, J. M., P. E. Rothrock, and R. T. Reber. 2008. Monitoring tallgrass prairie restoration performance using Floristic Quality Assessment. Proceedings of the Indiana Academy of Science **117**:16-28.

McLachlan, S. M., and D. R. Bazely. 2001. Recovery patterns of understory herbs and their use as indicators of deciduous forest regeneration. Conservation Biology **15**:98-110.

McNicoll, M., and C. Augspurger. 2010. A comparison of vegetation and seed bank community structure in a sand prairie in Illinois, U.S.A. American Midland Naturalist **164**:136-150.

Meador, M. R., and D. M. Carlisle. 2007. Quantifying tolerance indicator values for common stream fish species of the United States. Ecological Indicators **7**:329-338.

Medley, L., and M. Scozzafava. 2009. Moving toward a national Floristic Quality Assessment: Considerations for the EPA National Wetland Condition Assessment. National Wetlands Newsletter **31**:6-9, 22.

Meier, A. J., S. P. Bratton, and D. C. Duffy. 1995. Possible mechanisms for loss of vernal-herb diversity in logged eastern deciduous forest. Ecological Applications **5**:935-946.

Meiners, S. J., S. T. A. Pickett, and M. L. Cadenasso. 2001. Effects of plant invasions on the species richness of abandoned agricultural land. Ecography **24**:633-644.

Meiners, S. J., S. T. A. Pickett, and M. L. Cadenasso. 2002. Exotic plant invasions over 40 years of old field successions: community patterns and associations. Ecography **25**:215-223.

Meiners, S. J., T. A. Rye, and J. R. Klass. 2008. On a level field: the utility of studying native and non-native species in successional systems. Applied Vegetation Science **12**:45-53.

Middleton, E. L., and J. D. Bever. 2012. Inoculation with a native soil community advances succession in a grassland restoration. Restoration Ecology **20**:218-226.

Middleton, E. L., J. D. Bever, and P. A. Schultz. 2010. The effect of restoration methods on the quality of the restoration and resistance to invasion by exotics. Restoration Ecology **18**:181-187.

Milbauer, M., and M. Leach. 2007. Influence of species pool, fire history, and woody canopy on plant species density and composition in tallgrass prairie. Journal of the Torrey Botanical Society **134**:53-62.

Milburn, S. A., M. Bourdaghs, and J. J. Husveth. 2007. Floristic Quality Assessment for Minnesota Wetlands. Minnesota Pollution Control Agency, St. Paul, Minn.

Miller, S. J., and D. H. Wardrop. 2006. Adapting floristic quality assessment index to indicate anthropogenic disturbance in central Pennsylvania wetlands. Ecological Indicators **6**:313-326.

Miller, S. J., D. H. Wardrop, W. M. Mahaney, and R. P. Brooks. 2006. A plant-based index of biological integrity (IBI) for headwater wetlands in central Pennsylvania. Ecological Indicators **6**:290-312.

Mita, D., E. DeKeyser, D. Kirby, and G. Easson. 2007. Developing a wetland condition prediction model using landscape structure variability. Wetlands **27**:1124-1133.

Molano-Flores, B. 2002. Critical trends assessment program monitoring protocols. Illinois Natural History Survey, Office of the Chief Technical Report 2002-2, Champaign. 38pp.

Morgan, P. A., and F. T. Short. 2002. Using functional trajectories to track constructed salt marsh development in the Great Bay Estuary, Maine/New Hampshire, U.S.A. Restoration Ecology **10**:461-473.

Mortellaro, S., M. Barry, G. Gann, J. Zahina, S. Channon, C. Hilsenbeck, D. Scofield, G. Wilder, and G. Wilhelm. 2012. Coefficients of Conservatism values and the

Floristic Quality Index for the vascular plants of South Florida. Southeastern Naturalist **11**:1-62.

Mueller-Dombois, D., and H. Ellenberg. 1974. Aims and Methods of Vegetation Ecology. John Wiley & Sons, New York, New York.

Muratet, A., E. Porcher, V. Devictor, A. G., J. Moret, S. Wright, and N. Machon. 2008. Evaluation of floristic diversity in urban areas as a basis for habitat management. Applied Vegetation Science **11**:451-460.

Murray-Hudson, M., C. R. Lane, S. North, and M. T. Brown. 2012. Macrophyte species distribution, Indices of Biotic Integrity, and sampling intensity in isolated Florida marshes. Wetlands.

Mushet, D. M., N. H. Euliss, and T. L. Shaffer. 2002. Floristic quality assessment of one natural and three restored wetland complexes in North Dakota, USA. Wetlands **22**:126-138.

Mutke, J., and W. Barthlott. 2005. Patterns of vascular plant diversity at continental to global scales. Biol. Skr. **55**:521-531.

Myers, N., R. A. Mittermeier, C. G. Mittermeiter, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature **403**:853-858.

Nelson, J. C., A. Redmond, and R. E. Sparks. 1994. Impacts of settlement on floodplain vegetation at the confluence of the Illinois and Mississippi rivers. Transactions of the Illinois State Academy of Science **87**:117-133.

Nichols, J. D., J. E. Perry, and D. A. DeBerry. 2006. Using a floristic quality assessment technique to evaluate plant community integrity of forested wetlands in southeastern Virginia. Natural Areas Journal **26**:360-369.

Nichols, S. A. 1999. Floristic Quality Assessment of Wisconsin lake plant communities with example applications. Lake and Reservoir Management **15**:133-141.

Nichols, S. A. 2001. Long-term changes in Wisconsin lake plant communities. Journal of Freshwater Ecology **16**:1-13.

Nichols, S. A. 2008. Shifting plants in Wisconsin lakes. Pages 211-227 *in* D. M. Waller and T. P. Rooney, editors. The Vanishing Present: Wisconsin's Changing Lands, Waters, and Wildlife. University of Chicago Press, Chicago, IL.

Nielsen, S. E., E. M. Bayne, J. Schieck, J. Herbers, and S. Boutin. 2007. A new method to estimate species and biodiversity intactness using empirically derived reference conditions. Biological Conservation **137**:403-414.

Niemi, G. J., and M. E. McDonald. 2004. Application of ecological indicators. Annual Review of Ecology and Systematics **35**:89-111.

Norris, W. R., and D. R. Farrar. 2001. A method for the natural quality evaluation of central hardwood forests in the Upper Midwest, USA. Natural Areas Journal **21**:313-323.

Noss, R. F. 1985. On characterizing presettlement vegetation: How and why. Natural Areas Journal **5**:5-19.

Noss, R. F. 1996. The naturalists are dying off. Conservation Biology:1-3.

O'Neal, B. J., E. Heske, and J. D. Stafford. 2008. Waterbird response to wetlands restored through the Conservation Reserve Enhancement Program Journal of Wildlife Management **72**:654-664.

Packard, S. 1994. Successional restoration: thinking like a prairie. Restoration and Management Notes **12**:32-39.

Packham, J. R., and A. J. Willis. 1976. Aspects of the ecological amplitude of two woodland herbs, *Oxalis acetosella* L. and *Galeobdolon luteum* Huds. Journal of Ecology **64**.

Panzer, R., and M. W. Schwartz. 1998. Effectiveness of a vegetation-based approach to insect conservation. Conservation Biology **12**:693-702.

Panzer, R. D., R. Stillwaugh, R. Gnaedinger, and G. Derkovitz. 1995. Prevalence of remnant-dependence among the prairie inhabiting insects of the Chicago region. Natural Areas Journal **15**.

Pearman, P. B., M. R. Penskar, E. H. Schools, and H. D. Enander. 2006. Identifying potential indicators of conservation value using natural heritage occurrence data. Ecological Applications **16**:186-201.

Peet R.K., Wentworth T.R., and White P.S. 1998. A flexible, multipurpose method for recording vegetation composition and structure. Castanea **63**:262–274.

Peterken, G. F. 1974. A method for assessing woodland flora for conservation using indicator species. Biological Conservation **6**:239-245.

Peterken, G. F., and M. Game. 1984. Historical factors affecting the number and distribution of vascular plant species in the woodlands of central Lincolnshire. Journal of Ecology **72**:155-182.

Poling, T. C., M. G. Banker, and L. M. Jablonski. 2003. Quadrat level floristic quality index reflects shifts in composition of a restored tallgrass prairie (Ohio). Ecological Restoration **21**:144-145.

Povilitis, T. 2002. What is a natural area? Natural Areas Journal **21**:70-74.

Qian, H., J. D. Fridley, and M. W. Palmer. 2007. The latitudinal gradient of species-area relationships for vascular plants of North America. The American Naturalist **170**:690-701.

Qian, H., and R. E. Ricklefs. 2004. Geographical distribution and ecological conservatism of disjunct genera of vascular plants in easter Asia and eastern North America. Jounal of Ecology **92**:253-265.

Qian, H., and R. E. Ricklefs. 2007. A latitudinal gradient in large-scale beta diversity in for vascular plants in North America. Ecology Letters **10**:737-744.

Raab, D., and S. E. Bayley. 2012. A vegetation-based Index of Biotic Integrity to assess marsh reclamation success in the Alberta oil sands, Canada. Ecological Indicators **15**:43-51.

Rabinowitz, D. 1981. Seven forms of rarity. Pages 205-217 *in* H. Synge, editor. The biological aspects of rare plant conservation, Wiley, New York.

Reed, P. 1988. National list of plant species that occur in wetland: 1988 national summary. Biological Report 88 (24). Washington, D.C. U.S. Department of the Interior, U.S. Fish and Wildlife Service.

Reiss, K. R. 2006. Florida Wetland Condition Index for depressional forested wetlands. Ecological Indicators **6**:337-352.

Rentch, J. S., and J. T. Anderson. 2006. A floristic quality index for West Virginia wetland and riparian plant communities. http://community.wvu.edu/~jsr008/FQI_WV.pdf. 67 pp. Acessed 2013.

Ricklefs, R. E. 2008. Disintegration of the ecological community. The American Naturalist **172**:741-750.

Ricklefs, R. E. 2012. Naturalists, natural history, and the nature of biological diversity. The American Naturalist **179**:423-435.

Ricotta, C. 2004. A parametric diversity measure combining the relative abundances and taxonomic distinctiveness of species. Diversity and Distributions **10**:143-146.

Robinson, G. R., M. E. Yurlina, and S. N. Handel. 1994. A century of change in the Staten Island flora: ecological correlates of species losses and invasions. Bulletin of the Torrey Botanical Club **121**:119-129.

Rocchio, J. F. 2007. Floristic quality assessment indices for Colorado plant communities. Colorado Natural Heritage Program & Colorado State University:111p. + Appendices.

Rocchio, J. F. 2011. Western Washington Floristic Quality Assessment. Washington Department of Natural Resources. & U.S. Environmental Protection Agency, Region 10, Seattle, WA.

Rooney, T. P., and M. Leach. 2010. Replacing hay-mowing with prescribed fire restores species diversity and conservation value in a tallgrass prairie sampled thrice: a 59-year study. American Midland Naturalist **164**:311-321.

Rooney, T. P., and D. A. Rogers. 2002. The modified floristic quality index. Natural Areas Journal **22**:340-344.

Rothrock, P. E., and M. A. Homoya. 2005. An evaluation of Indiana's Floristic Quality Assessment. Proceedings of the Indiana Academy of Science **114**:9-18.

Rothrock, P. E., R. T. Reber, and M. A. Misurac. 2011. Floristic Quality Assessment along an old-field chronosequence. Proceedings of the Indiana Academy of Science **1-2**:12-17.

Sampson, F., and F. Knopf. 1994. Prairie conservation in North America. BioScience **44**:418-421.

Sanderson, E. W., M. Jaiteh, M. A. Levy, K. H. Redford, A. V. Wannebo, and G. Woolmer. 2002. The Human Footprint and the last of the wild. BioScience **52**:891-904.

Sarr, D. A., D. E. Hibbs, and M. A. Huston. 2005. A hierarchical perspective of plant diversity. The Quarterly Review of Biology **80**:187-212.

Scholes, R. J., and R. Biggs. 2005. A biodiversity intactness index. Nature **434**:45-49.

Schwegman, J. E., G. B. Fell, M. D. Hutchinson, G. Paulson, W. M. Shephard, and J. White. 1973. Part 2: The natural divisions of Illinois. Comprehensive Plan for the

Illinois Nature Preserve system, Illinois Nature Preserves Commission, Springfield, IL.

Scott, W. A., and C. J. Hallam. 2002. Assessing species misidentification rates through quality assurance of vegetative monitoring. Plant Ecology **165**:101-115.

Singleton, R., S. Gardescu, P. L. Marks, and M. A. Geber. 2001. Forest herb colonization of postagricultural forests in central New York State, USA. Journal of Ecology **89**:325-338.

Sivicek, V. A., and J. B. Taft. 2011. Functional group density as an index for assessing habitat quality in tallgrass prairie. Ecological Indicators **11**.

Slagle, M. W., P. W. Sweeney, K. M. Hiser, and H. Wells-Sweeney. 2008. Vegetation survey of prairie reconstructions at Litzsinger Road Ecology Center.*in* Proceedings of the 20th North American Prairie Conference, University of Nebraska-Kearney, Kearney, Nebraska.

Smart, A. J., M. J. Nelson, P. J. Bauman, and G. E. Larson. 2011. Effects of herbicides and grazing on floristic quality of native tallgrass pastures in eastern South Dakota and southwestern Minnesota. Great Plains Research **Fall**:181-189.

Smith, P. G. R., and J. B. Theberge. 1986. A review of criteria for evaluating natural areas. Environmental Management **10**:715-734.

Sperry, T. M. 1994. The Curtis Prairie restoration, using the single-species planting method. Natural Areas Journal **14**:124-127.

Spieles, D. J., M. Coneybeer, and J. Horn. 2006. Community structure and quality after 10 years in two central Ohio mitigation bank wetlands. Environmental Management **38**:837-852.

Spyreas, G., J. Ellis, C. Carroll, and B. Molano-Flores. 2004. Non-native plant commonness and dominance in the forests, wetlands, and grasslands of Illinois, USA. Natural Areas Journal **24**:290-299.

Spyreas, G., and J. W. Matthews. 2006. Floristic conservation value, nested understory floras, and the development of second-growth forest. Ecological Applications **16**:1351-1366.

Spyreas, G., S. J. Meiners, J. W. Matthews, and B. Molano-Flores. 2012. Successional Trends in Floristic Quality. Journal of Applied Ecology **49**:339-348.

Spyreas, G., B. W. Wilm, A. E. Plocher, D. M. Ketzner, J. W. Matthews, J. Ellis, and E. J. Heske. 2010. Biological consequences of invasion by reed canary grass (*Phalaris arundinacea*). Biological Invasions **12**:1253-1267.

Stanley, K. E., P. G. Murphy, H. H. Prince, and T. M. Burton. 2005. Long-term ecological consequences of anthropogenic disturbance on Saginaw Bay coastal wet meadow vegetation. Journal of Great Lakes Research **31**:147-159.

Stapanian, M. A., J. V. Adams, and B. Gara. 2013. Presence ofIndicator plant species as a predictor of wetland vegetation integrity: a statistical approach. Plant Ecology **214**:291-302.

Stein, B. A. 2002. States of the Union: Ranking America's Biodiversity. NatureServe, Arlington, Virginia.

Streever, W. J. 1999. Examples of performance standards for wetland creation and restoration in Section 404 permits and an approach to developing performance standards. U.S. Army Engineer Research and Development Center, Vicksburg, MS.

Swink, F., and G. Wilhelm. 1979. A method for environmental assessment of open land. Pages 850-861 Plants of the Chicago Region. The Morton Arboretum, Lisle, IL.

Swink, F., and G. Wilhelm. 1994a. Coefficients of Conservatism. Pages 8-9 Plants of the Chicago Region. Indiana Academy of Science, Indianapolis, IN.

Swink, F., and G. Wilhelm. 1994b. Floristic Quality Assessment. Pages 11-18 Plants of the Chicago Region. Indiana Academy of Science, Indianapolis, IN.

Swink, F., and G. Wilhelm. 1994c. Plants of the Chicago Region. Indiana Academy of Science, Indianapolis, IN.

Taft, J. B. 1994. Ecology, distribution, and rareness patterns of threatened and endangered praire plant in Illinois.*in* Proceedings of the 4th Central Illinois Prairie Conference, Millikin University, Decatur, Illinois.

Taft, J. B. 1996. Reading the signs: plants as indicators of site history. Illinois Steward **5**:20-24.

Taft, J. B. 1999. Vegetation ecology and fire effects on community structure, composition, and diversity in a barrens remnant and a flatwoods in southern Illinois. University of Illinois, Urbana-Champaign, Urbana, IL.

Taft, J. B., C. Hauser, and K. Robertson. 2006. Estimating floristic integrity in tallgrass prairie. Biological Conservation **131**:42-51.

Taft, J. B., G. Wilhelm, D. Ladd, and L. A. Masters. 1997. Floristic Quality Assessment for vegetation in Illinois, a method for assessing vegetation integrity. Erigenia **15**:1-24 + Appendix.

Ter Braak, C. J. F., and N. J. M. Gremmen. 1987. Ecological amplitudes of plant species and the internal consistency of Ellenberg's Indicator values for moisture. Vegetatio **69**:79-87.

Theobald, D. M. 2013. A general model to quantify ecological integrity for landscape assessments and US application. Landscape Ecology **28**:1859-1874.

Thompson, R., and S. Green. 2010. Vascular plants of an abandoned Limestone quarry in Garrard County, Kentucky. Castanea **75**:245-258.

Thomson, J. W. 1940. Relic prairie areas in central Wisconsin. Ecological Monographs **10**:685-717.

Towne, E. G., D. C. Hartnett, and R. C. Cochran. 2005. Vegetation trends in tallgrass prairie from bison and cattle grazing. Ecological Applications **15**:1550-1559.

Tu, W. G., X. F. Gao, N. Wu, and S. H. Liu. 2009. A novel method for floristic quality assessment using the vegetation of the Jiuding Mountain, Sichuan, China as an example. Nordic Journal of Botany **27**:61-68.

Tulbure, M. G., C. A. Johnston, and D. L. Auger. 2007. Rapid invasion of a Great Lakes coastal wetland by non-native *Phragmites australis* and *Typha*. Journal of Great Lakes Research **33**:269-279.

Twedt, D. J. 2005. An objective method to determine an area's relative significance for avian conservation. Pages 71-77 *in* J. C. Ralph, and T. D. Rich, editors. Bird conservation implementation and integration in the Americas: Proceedings of the Third International Partners in Flight Conference. Volume 1 Gen. Tech. Rep. PSW-GTR-191. U.S. Dept. of Agriculture, Forest Service, Pacific Southwest Research Station, Albany CA, USA.

U.S.E.P.A. 2002a. Methods for evaluating wetland condition: developing metrics and Indexes of Biological Integrity. Office of Water, U.S. Environmental Protection Agency, Washington, DC. EPA-822-R-02-016.

U.S.E.P.A. 2002b. Methods for evaluating wetland condition: using vegetation to assess environmental conditions in wetlands. Office of Water, U.S. Environmental Protection Agency,Washington, DC. EPA-822-R-02-020.

U.S.E.P.A. 2010. Draft ecological indicators for the 2011 National Wetland Condition Assessment. U.S. Environmental Protection Agency, Washington, DC.

U.S.F.W.S. 2013. Neal Smith National Wildlife Refuge Comprehensive Conservation Plan, Prairie City, IA.

van Breugel, M., J. S. Hall, D. Craven, M. Bailon, A. Hernandez, M. Abbene, and P. van Breugel. 2013. Succession of ephemeral secondary forests and their limited role for the conservation of floristic diversity in a human-modified tropical landscape. PLoS ONE **8**.

Vaughn, K. J., and T. P. Young. 2010. Contingent conclusions: year of initiation influences ecological field experiments, but temporal replication is rare. Restoration Ecology **18**:59-64.

Visser, V., W. D. Clayton, D. A. Simpson, R. P. Freckleton, and C. P. Osborne. 2014. Mechanisms driving an unusual latitudinal diversity gradient for grasses. Global Ecology and Biogeography **23**:61-75.

Von Holle, B., H. R. Delcourt, and D. Simberloff. 2003. The importance of biological inertia in plant community resistance to invasion. Journal of Vegetation Science **14**:425-432.

Wallace, M. P. 2001. A floristic comparison of dry-mesic upland forest communities in Illinois. University of Illinois at Urbana-Champaign, Urbana, IL.

Werner, K. J., and J. B. Zedler. 2002. How sedge meadow soils, microtopography, and vegetation respond to sedimentation. Wetlands **22**:451-466.

White, J. 1978. Illinois Natural Areas Technical Report, Volume 1. Survey Methods and Results. Illinois Department of Natural Resources, Springfield, IL.

White, J., and M. H. Madany. 1978. Classification of natural communities in Illinois. Pages 310-405 (Appendix 330) *in* J. White, editor. Illinois Natural Areas Technical Report, Volume 1. Survey Methods and Results. . Illinois Department of Natural Resources, Springfield, IL.

Whitney, G., and D. R. Foster. 1988. Overstorey composition and age as determinants of the understorey flora of woods of central New England. Journal of Ecology **76**:867-876.

Whittaker, R. H. 1978. Classification of Plant Communities. Springer, Hague, Netherlands.

Wiens, J., and C. H. Graham. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annual Review of Ecology, Evolution, and Systematics **2005**:519-539.

Wilcox, D. A., J. E. Meeker, P. L. Hudson, B. J. Armitage, M. G. Black, and D. G. Uzarski. 2002. Hydrologic variability and the application of Index of Biotic Integrity metrics to wetlands: a Great Lakes evaluation. Wetlands **22**:588–615.

Wilhelm, G. 1977. Ecological assessment of open land areas in Kane County, Illinois. Kane County Urban Development, Geneva, Illinois.

Wilhelm, G. 1992. Technical comments on the proposed revisions to the 1989 Wetland Delineation Manual. Erigenia **12**:41-49.

Wilhelm, G., and D. Ladd. 1988. Natural area assessment in the Chicago Region. Transactions of the 53rd North American Wildlife & Natural Resources Conference:361-375.

Wilhelm, G., and L. A. Masters. 1994. Floristic changes after five growing seasons in burned and unburned woodland. Erigenia **13**:141-150.

Wilhelm, G., and L. A. Masters. 1995. Floristic Quality Assessment in the Chicago Region and application computer programs. The Morton Arboretum, Lisle, IL.

Wilson, M. J., and S. E. Bayley. 2012. Use of single versus multiple biotic communities as indicators of biological integrity in northern prairie wetlands. Ecological Indicators **20**:187-195.

Wisheu, I. C., and P. A. Keddy. 1992. Competition and centrifugal organization of plant communities: theory and tests. Journal of Vegetation Science **3**:147-156.

Withers, M. A., M. W. Palmer, G. L. Wade, P. S. White, and P. R. Neal. 1998. Changing Patterns in the number of species in North American Floras. Page 104 *in* T. D. Sisk, editor. Perspectives on the land use history of North America: a context for understanding our changing environment, U.S. Geological Survey, Biological Resources Division, Biological Science Report USGS/ BRD/BSR-1998-0003.

Woods, A. J., J. M. Omernik, C. L. Pederson, and B. C. Moran. 2006. Level III and IV Ecoregions of Illinois. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Western Ecology Division, Corvallis, OR.

Yarnell, R. A. 1976. Early plant husbandry in eastern North America. Pages 265-273 *in* C. E. Cleland, editor. Cultural change and continuity: essays in honor of Jame Bennett Griffin. Academic Press, New York.

Zbigniew, D. 2001. Assessment of light and soil conditions in ancient and recent woodlands by Ellenberg indicator values. Journal of Applied Ecology **38**:942-951.

Zedler, J. B., and J. C. Callaway. 1999. Tracking wetland restoration: do mitigation sites follow desired trajectories? Restoration Ecology **7**:69-73.