4-10-2017

Past, Present, and Future of Ecological Integrity Assessment for Fresh Waters

Lauren M. Kuehne
University of Washington - Seattle Campus

Julian D. Olden
University of Washington - Seattle Campus

Angela L. Strecker
Portland State University, angela.strecker@wwu.edu

Joshua J. Lawler
University of Washington - Seattle Campus

David M. Theobald
Colorado State University - Fort Collins

Citation Details

Kuehne, L. M., Olden, J. D., Strecker, A. L., Lawler, J. J., & Theobald, D. M. (2017). Past, Present, and Future of Ecological Integrity Assessment for Fresh Waters. Frontiers In Ecology & The Environment, 15(4), 197-205.

Past, present, and future of ecological integrity assessment for fresh waters

Lauren M Kuehne[1]*, Julian D Olden[1], Angela L Strecker[2], Joshua J Lawler[3], and David M Theobald[4,5]

One of the most influential environmental laws in the US – the 1972 Clean Water Act – included the visionary objective of maintaining and restoring aquatic ecological integrity. However, the efficacy of the Act depends on how integrity is assessed. Reviewing the assessment literature for fresh waters over the past 40 years, we found evidence of methodological trends toward increased repeatability, transferability, and robustness of assessments over time. However, implementation gaps were revealed, based on the relatively weak linkages to freshwater policies, stakeholder involvement, emerging threats, and conservation opportunities. A related survey of assessment practitioners underscored the disparity between need versus availability of assessments that guide management policies. Technological changes in data collection and synthesis have clearly influenced assessments, and appear to have led to a reduced reliance on ecological response indicators and corresponding increases in stressor-based metrics. We recommend designing assessments around specific freshwater policies and regulations to improve applicability of assessment products for management and conservation.

Front Ecol Environ 2017; doi:10.1002/fee.1483

Restoring and maintaining "the chemical, physical, and biological integrity of the Nation's waters" (termed *ecological integrity*) was considered both an ambitious and visionary component of the 1972 US Federal Water Pollution Control Act (the "Clean Water Act" or CWA), and extensive conceptual and empirical work has since been dedicated to its measurement and assessment. In the early years, assessment programs focused narrowly on chemical contaminants (Karr and Dudley 1981), but broadened as researchers developed meaningful and defensible ways to incorporate physical and biological integrity into assessments (Karr 1993; Barbour *et al.* 2000).

In a nutshell:

- Since the passage of the 1972 US Clean Water Act, assessment of ecological integrity has underpinned conservation and management of fresh waters
- Our review indicates that although assessment methods have become more consistent and robust over time, there is little evidence of integration with management, conservation, and policy
- Increasing reliance on geographic information systems and remote-sensing is substantially influencing the metrics used in freshwater assessments, with consequences for the continuity and defensibility of results
- We recommend specifying the scale and design of assessments around specific freshwater policies and regulations to narrow the knowledge-to-action gap in freshwater conservation

Recent decades have witnessed broad acceptance of ecological integrity as a conservation goal, largely due to the benefits that it offered over related concepts of ecosystem health and biological diversity (Suter 1993; Angermeier and Karr 1996; Lackey 2001). Importantly, integrity was supported by legislative mandate and was an efficient way to communicate concern about environmental resources (Angermeier and Karr 1996; Woodley 2010). It was also seen as encompassing multiple concepts; for example, a system often has integrity when it is biologically diverse (Karr 1993). As a result, ecological integrity has been adopted as a management directive by diverse agencies in North America that oversee terrestrial as well as aquatic ecosystems, including the National Wildlife Refuge System (1997), Parks Canada (1998), and the US Forest Service (2012), among others (Figure 1).

The most often-cited definition of ecological integrity was articulated by Karr and Dudley (1981) not long after the CWA was passed, and is described as "the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitat of the region". In the years since the CWA's initial passage, however, the methods used to measure and describe the integrity of fresh waters have evolved in response to new conceptual frameworks, technologies, analytical approaches, and policy directives (Figure 1) (Yoder and Barbour 2009; Morgan and Hough 2015). Consequently,

[1]*School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA* *(lkuehne@uw.edu); [2]*Department of Environmental Science and Management, Portland State University, Portland, OR*; [3]*School of Environmental and Forest Sciences, University of Washington, Seattle, WA*; [4]*Conservation Science Partners, Fort Collins, CO*; [5]*Department of Fish, Wildlife and Conservation Biology, Colorado State University, Fort Collins, CO*

the time is ripe to evaluate progress in assessments of freshwater integrity and to identify knowledge or implementation gaps, which – if filled – would better support management and conservation.

We conducted a systematic review of peer-reviewed and gray literature related to assessment of wetlands, lakes, streams, riparian areas, and watersheds to understand how measurements of freshwater integrity are changing over time (WebTable 1). We used an attribute-based approach to analyze 89 assessments conducted since the passage of the CWA; these assessments evaluated the integrity or condition of at least one freshwater ecosystem at the scale of a watershed or greater. For each assessment, we characterized 60 attributes within four general categories – Assessment type, Methods, Indicators, and Applications (WebTable 2) – and analyzed dominant trends in each category. To identify how well assessments are meeting current management needs, we surveyed aquatic resource managers and researchers about the importance of using different assessment approaches (eg diagnostic, rapid) in their own work; their responses were compared to representation of these assessments in the literature to identify research and implementation gaps that may impede conservation, management, and policy actions.

Assessment of ecological integrity is foundational to conservation and management of fresh waters worldwide. However, given that assessment is closely tied to legislative directives and monitoring standards, we concentrated our review on work conducted within North America to produce a focused analysis of trends within a relatively narrow range of management and policy structures (Schröter *et al.* 2016).

Multi-decadal trends in assessment types and methods

The literature related to assessment of integrity in fresh waters accrued moderately through the mid-1990s, followed by rapid accumulation between 1995 and 2005 (Figure 1); the subset of assessments that met our criteria for review also followed this general pattern. Assessments have focused largely on streams or rivers (48%) (Figure 2), with only 1 in 10 incorporating more than one ecosystem type (typically stream and riparian areas). The spatial extent of assessments varied greatly but was generally balanced across scale categories, with only slight overrepresentation of smaller (58% sub-basin and basin) scales of assessment (as compared to larger-scale single or multiple ecoregions; Figure 2). The geographic extent of integrity assessments (in square kilometers) has steadily grown with time (R = 0.22, P = 0.04; correlation with publication year).

We also evaluated the extent to which assessments incorporate recommended design features or best practices. For example, two of the most fundamental recommendations in assessments are the use of (1) a reference method that allows condition to be evaluated in terms of departure from an ecological standard, and (2) methods that partition natural variability from human impacts; without adopting these practices, it is virtually impossible to credibly identify (and therefore mitigate) anthropogenic impacts on ecosystems (Innis *et al.* 2000; Barbour *et al.* 2004; Stein *et al.* 2009). Reviewing the knowledge accumulated over the past four decades through national workgroups, assessment standards, policy evaluations, and published studies (Figure 1), we identified attributes that describe these and other recommended practices, as well as design features that facilitate use of results for conservation and management (eg involvement of stakeholders, climate relevance) (WebTable 2). We then analyzed how often these practices and design features were incorporated over time; for instance, it might be expected that measurement of climate-relevant variables would be underrepresented across all assessments but improving in recent years with growing awareness of climate change.

We found highly encouraging trends in methodological practices that enhance the robustness, transferability, and repeatability of ecological integrity assessments. Overall, there was a very high prevalence both in use of reference methods and in methods for partitioning natural variability from human impacts (Figure 3). Recent declines in these

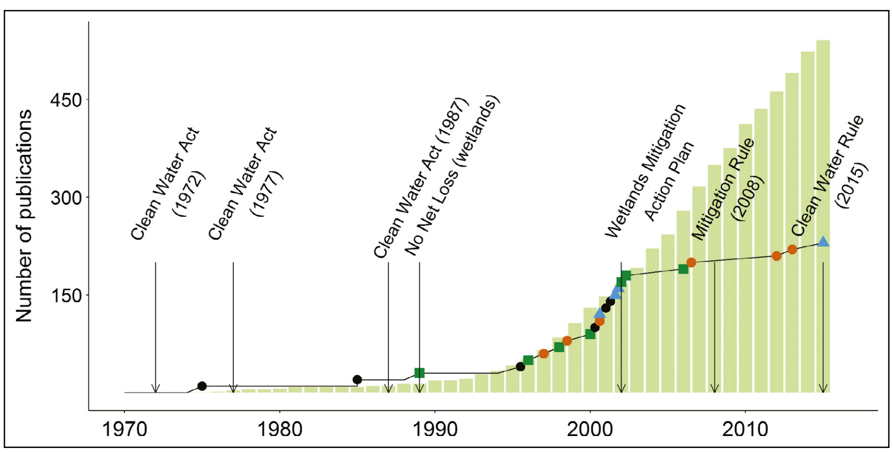

Figure 1. *Modifications in US freshwater policy (arrows, text) over time and cumulative research publications (bars) related to assessment of ecological integrity for freshwater ecosystems since passage of the 1972 Clean Water Act. Accumulation of knowledge and increased emphasis on ecological integrity assessment has also grown over time through the results of national workgroups (black circles), development of nationally applicable assessment standards (green squares), adoption of "integrity" as a management directive by federal agencies (orange circles), and evaluations of freshwater policies (blue triangles).*

trends are explained by greater numbers of watershed assessments; we found that such assessments incorporated reference conditions and accounted for natural variability far less frequently (45% and 69%, respectively) than all other types of assessments (83% and 90%).

The use of probabilistic (ie random) sampling for site selection was common and has grown considerably in recent years, which bodes well for the capacity to generalize results and compare results of assessments at different scales (Hughes *et al.* 2000). Likewise, the "repeatability" of assessments – defined as use of systematic and statistical approaches versus exclusive reliance on expert opinion in indicator and metric selection – is also increasing (Theobald *et al.* 2007; Stoddard *et al.* 2008). We also found greater use of collaborative datasets and synthetic modeling, indicating that assessment researchers are taking advantage of (and contributing to) advances in data standardization, data sharing, and more robust modeling frameworks to address data gaps (Hampton *et al.* 2013).

Less encouraging were the trends related to applications of assessment results, with little evidence of direct linkages to conservation or management (Figure 3). Less than one-quarter (24%) of assessment products or outcomes were related to any specific freshwater policy (state or national). This may be an indication that management needs and directives are not effectively considered at design stages, a practice that has been strongly recommended in earlier reviews and critiques (Kusler and Niering 1998; Innis *et al.* 2000; Stein *et al.* 2009). Stakeholder involvement is typically credited with greater acceptance and implementation of assessment results (Barbour *et al.* 2004; Allan *et al.* 2012), but our results showed that stakeholders were infrequently included in assessments (12%). To help account for the possibility that the reviewed frameworks were biased toward research rather than applied science (and therefore had potentially fewer opportunities or less need to engage stakeholders), we also considered whether assessments provided and discussed avenues for stakeholder involvement (eg recommending an iterative process of metric selection with public input). The percentage of assessments that included possibilities for stakeholder involvement was more than double (29%) the percentage where stakeholders were explicitly included. Although still somewhat low, the growing trend in both attributes that describe involvement of stakeholders is a promising indication that engagement with end users is being considered a routine part of the assessment process.

It is expected that assessments might shift in response to emerging threats and as new scientific, management, or conservation information becomes available. We evaluated how often assessments reflected two issues of critical importance to freshwater ecosystems: (1) projected changes in climate-sensitive metrics such as water temperature and streamflow, and (2) use of ecosystem service concepts to evaluate and communicate results (Naiman and Dudgeon 2011). We found that climate-sensitive

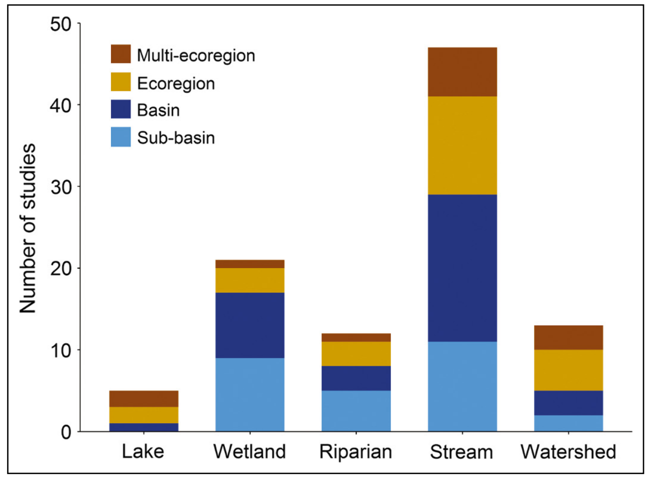

Figure 2. *Ecosystem focus and geographic scale of the freshwater assessments reviewed. The approximate geographic area (in square kilometers) associated with each scale category is sub-basin (10^3), basin (10^4), ecoregion (10^5), and multiple ecoregion (10^6).*

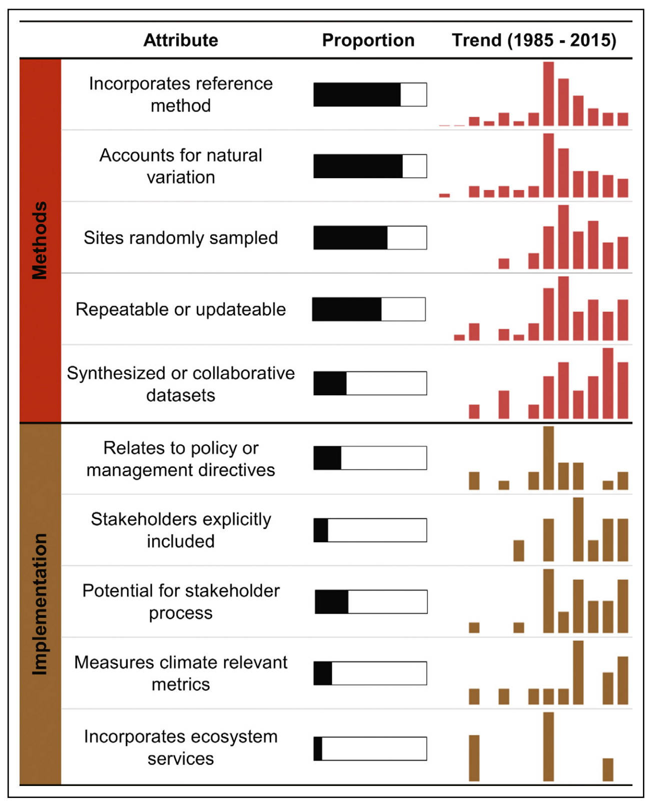

Figure 3. *Prevalence and trends in key attributes related to methodology (Methods) and how the assessment outputs apply to broader social, economic, or policy structures (Implementation). Out of the assessments reviewed, "Proportion" is the number in which the attribute was scored as present (horizontal black bar), and "Trend" shows how often the attribute was represented in assessments through time (1985–2015, pooled in 2-year increments, vertical bars).*

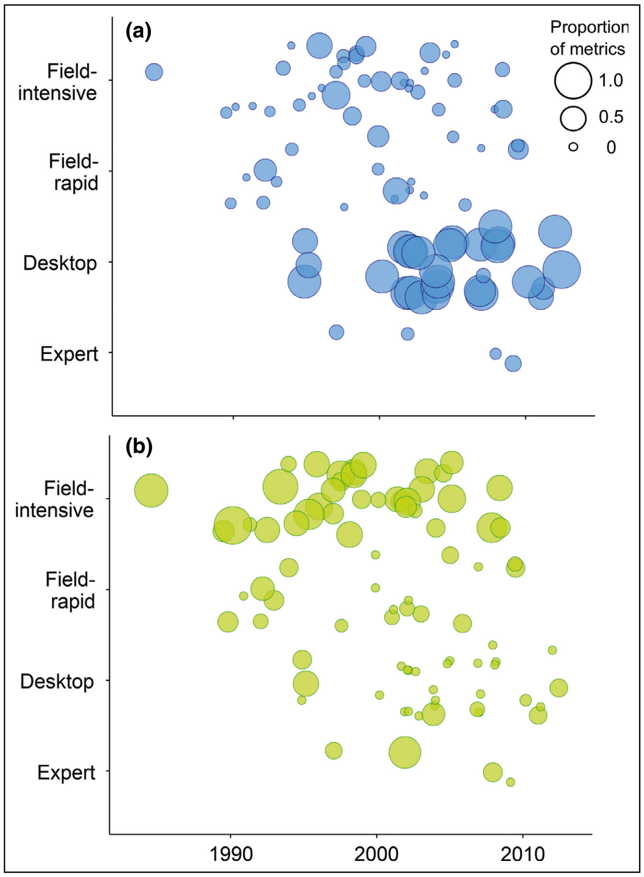

Figure 4. *The proportion of (a) landscape and (b) biological indicators represented over time in assessments based on field-intensive, field-rapid, desktop, and expert data collection efforts. For clarity, data points were jittered within categories. See WebFigure 1 for trends in chemical and physical metrics.*

metrics were included within assessments relatively infrequently (16%) and with only a slightly upward trend through time. This suggests that existing assessment frameworks are not well positioned to account for changes in integrity due to projected climate impacts on freshwater organisms (Heino *et al.* 2009). There was very little representation of ecosystem services in assessments (7%), with no trend toward improvement, suggesting strong conceptual or methodological barriers between typical assessment practices and ecosystem service valuation (Liu *et al.* 2010; Schröter *et al.* 2016). This implementation gap contrasts strongly with assessment efforts in the European Union, where ecosystem services are being comprehensively assessed in conjunction with ecological condition for all Member States as part of the European Biodiversity Strategy to 2020 (Maes *et al.* 2016).

Changes in assessment indicators and metrics

Ecological integrity is interpreted through the lens of the metrics being measured, and analysis of the dominant metrics used can be a powerful way to detect pivotal changes in assessment practice. We therefore classified the individual metrics in each assessment as physical, chemical, biological, or landscape-oriented (WebTable 2), and calculated the proportion of metrics in each group over time and with respect to four types of data-collection effort: field-intensive (>0.5 day per site), field-rapid (<0.5 day per site), desktop (relies primarily on spatial and/or remotely sensed data), and expert (synthesis of expert knowledge) (WebTable 2).

We saw a distinct shift away from the prominence of biological metrics (pre-2000) and toward a greater reliance on landscape metrics (post-2000) to assess ecological integrity (Figure 4), while use of physical and chemical metrics remained more consistent over the same time period (WebFigure 1). Physical and chemical metrics had the broadest representation across all types of data-collection effort, whereas biological and landscape metrics were more closely associated with field-intensive and desktop assessments, respectively. Somewhat unexpectedly, landscape metrics were not often incorporated in field-rapid assessments, which instead relied predominantly on physical metrics. The number of expert assessments was small but suggests an emphasis on physical and chemical metrics.

The process of selecting metrics can be complex, but both qualitative and quantitative evidence support inclusion of multiple metric types that reflect a range of ecosystem processes as a foundation for accurate representation of conditions as well as ability to distinguish causes and mechanisms of degradation (Noss 1990; Brooks *et al.* 1998; Clapcott *et al.* 2012; Vander Laan *et al.* 2013). Inferences can also differ depending on whether the emphasis is on ecological metrics that respond to stress (hereafter "indicators"), ecosystem stressors, or a combination of both (Barbour *et al.* 2004; Wardrop *et al.* 2007). We therefore compared the metrics used in assessments against two checklists that we considered benchmarks or standards for freshwater indicators and stressors: Environmental Monitoring and Assessment Program (EMAP) Core Indicators for Surface Waters (McDonald *et al.* 2004) and the freshwater stressors synthesized by Dolédec and Statzner (2010). Using these checklist data we calculated the diversity of indicators and stressors separately for each assessment using the Shannon-Wiener index, and examined concurrent changes in diversity with time.

The diversity of indicators used in assessments grew from the mid-1980s to 2000, at which point indicator diversity becomes asymptotic before declining up to the present (Figure 5). Concurrently, the diversity of stressors in assessments showed an early and marked rise and then plateaued for more than a decade; a second increase in stressor diversity is indicated in recent years. The overall trend toward greater reliance on exclusively stressor-based assessments (which functionally differ from response-based assessments) has important implications for freshwater management and conservation.

Management thresholds are typically related to ecological response metrics (eg biocriteria, conservation targets); although the type and extent of stress on ecosystems can point toward potential sources of degradation, stressors are less likely to represent actionable metrics. The distribution of stressors may also be very poorly aligned with social, economic, or management constraints, and a focus on reducing threats may not have the desired or expected impact on conservation targets (Tulloch *et al.* 2015). For these reasons, while acknowledged as useful tools to prioritize areas of the landscape for management or conservation, the use of stressor-based assessments is strongly recommended within a well-defined and structured conservation strategy (Brooks *et al.* 2006; Tulloch *et al.* 2015).

Management needs versus availability of assessments

Integrity assessments are conducted for a variety of purposes, but ultimately are only useful to the extent that they support conservation, protection, and restoration of fresh waters. In other words, even if an assessment accurately reflects ecological condition, it is of limited use if the analyses or outputs fail to match practitioner needs (Bain *et al.* 2000; Stein *et al.* 2009). Through the literature review, we identified the five most common goals (ie reasons) for conducting an assessment (WebTable 2) and used these as a checklist to determine how often goals were met across all of the assessments reviewed. Although some individual assessments met all five goals, those that met two or three goals were most common (60%).

Using a standardized survey instrument, we then asked freshwater assessment experts and practitioners in state, federal, and academic institutions to rank the importance of these same five assessment goals in their work (WebTable 3); using a five-point Likert scale ranging from "Rarely" to "Very Often", respondents (n = 169, representing a 38% response rate) indicated how often they needed assessments designed to meet each goal.

All five assessment goals were ranked fairly high, with two-thirds or more of the respondents answering "Often" or "Very Often" (Figure 6). Assessments that recommended management policies were highest ranked, with four-fifths (79%) of respondents saying they had need of these assessments either "Often" or "Very Often", and diagnostic assessments were ranked the lowest (66%). When we compared respondents' rankings of goals to representation across assessments in the literature, the largest gap between availability and need was for assessments that recommend management policies (56% difference), followed by diagnostic assessments (29%). Gaps were low to minimal for assessments that monitored resources (18%), were rapid and simple to conduct (17%), and prioritized areas for restoration or protection (8%). Clearly, all types of assessments were valuable; however, the current knowledge gaps point toward management concerns (policy guidance and diagnosis) as opposed to more technical or methodological requirements (monitoring, rapidity, and prioritization).

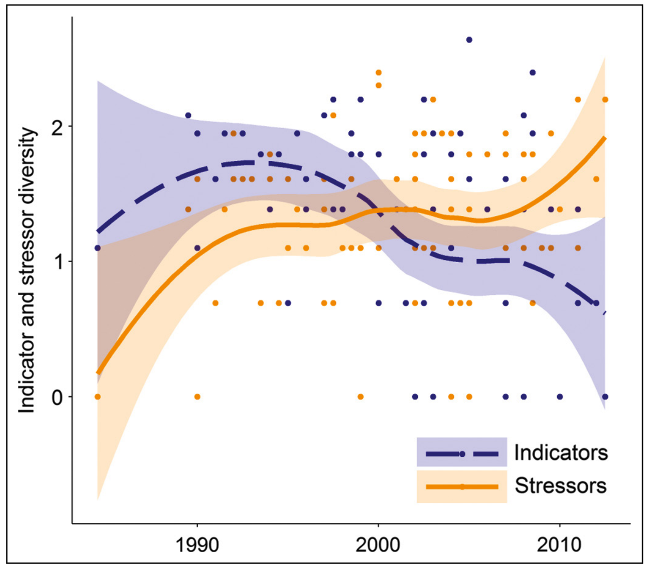

Figure 5. *Temporal trend in the diversity of ecological response indicators (dark blue circles) and stressors (orange circles) included in assessments over time as measured by the Shannon-Wiener diversity index; color-matched lines and shading depict LOESS smoothing curves and 95% confidence intervals. Twenty-one EMAP Core Indicators for Surface Waters (McDonald* et al. *2004) were the benchmark for examination of changes in indicator diversity over time. Fifteen freshwater ecosystem stressors compiled by Dolédec and Statzner (2010) were used to examine changes in stressor diversity over time (see WebTable 2 for additional details on the indicator and stressor checklists).*

Future of freshwater assessments

A fundamental goal of this review was to guide future assessment research, and examine where there is untapped potential to support conservation and protection of freshwater ecosystems. We have demonstrated ways in which integrity assessments have evolved since the passage of the CWA, and expect ongoing evolution into the future as new technologies create opportunities to reduce assessment costs while increasing extent, accuracy, and resolution (Table 1). However, it is important to also proceed thoughtfully to ensure that our basic understanding of integrity is not (unintentionally) redefined along the way.

We demonstrated encouraging methodological trends, specifically with respect to the use of reference methods, methods that account for natural variability, and of random site selection, which improves statistical inference. These practices – combined with more frequent use of statistical procedures for indicator selection (Whittier *et al.* 2007; Stoddard *et al.* 2008), reliance on nationally

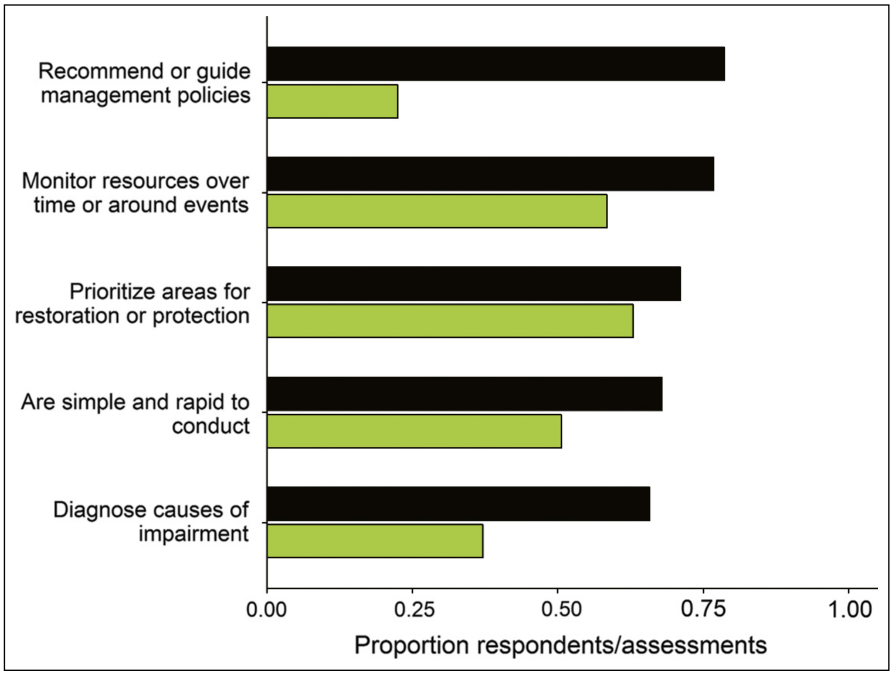

Figure 6. *Proportion of survey respondents who replied "Often" or "Very Often" with respect to their need for assessments designed around five common assessment goals (black bars), contrasted with how often those goals were achieved or represented across all assessments reviewed (green bars).*

standardized protocols (McDonald *et al.* 2004), and use of publicly available (rather than proprietary) data – appear to be supporting more systematic approaches and improved repeatability of assessments. An exception to this pattern is within watershed-type assessments; although methods that partition natural variability from human impacts are particularly critical at larger geographic scales (Esselman *et al.* 2011), our results indicate that fewer landscape assessments address this problem (Figure 3). Continued efforts to standardize existing methods (Detenbeck *et al.* 2000) and develop new approaches that incorporate both site-level and watershed processes are needed (Clapcott *et al.* 2012; Vander Laan *et al.* 2013), as well as statistical analyses that properly account for data complexity and the spatial structure of hydrological networks (Table 1) (Peterson *et al.* 2013).

We found a clear trend toward greater use of remotely sensed and geospatial data in integrity assessments. This type of data also tends to emphasize ecosystem stressors. Here we must acknowledge a caveat of our methods, which is that literature reviews may reflect a disproportionate use of novel technology as researchers grapple with how to incorporate new tools into existing practices. Regardless, we anticipate that use of landscape-level tools in assessments will continue to grow as the diversity of geospatial data increases, along with its spatial and temporal resolution (Table 1).

For freshwater ecosystems, overreliance on landscape indicators presents specific challenges, particularly if driven mainly by the desire to reduce assessment costs or time. Relationships between landscape metrics and freshwater outcomes of interest have arguably not been tested to an extent that allows consistent inferences (Leibowitz and Hyman 1999). For example, relationships between impervious land cover and urban stream indicators have been assessed in hundreds of studies, but predictive capacity remains highly variable depending on the watershed context and the response indicator chosen (Schueler *et al.* 2009). Our own review of the subset of documents that compared a desktop assessment with field-intensive measurements (WebTable 1) suggests that landscape approaches alone tend to reliably identify only areas with the highest and lowest integrity (eg Wardrop *et al.* 2007), or have considerably reduced accuracy (< 70% variance explained) in classifying sites using broad condition classes (eg impaired, good) (Falcone *et al.* 2010; Brown *et al.* 2012). The defensibility of assessment results plays strongly into management and policy decisions, and large gaps in explanatory and diagnostic power can undermine the ability of an assessment to support conservation action (Frazier 1998; Keiter 2004).

The use of landscape metrics has immense advantages related to geographic extent, access, and convenience; most importantly, landscape metrics can quantify human impacts at scales (eg catchment) that are difficult if not impossible to do through ground-based sampling. However, our results suggest exclusive reliance on landscape metrics as opposed to complementary approaches, which have been strongly recommended for fresh waters (Gergel *et al.* 2002). We suspect that pressures to reduce assessment costs and time while increasing spatial scale are driving these trends (Carletti *et al.* 2004), combined with challenges in integrating assessments across disciplinary (eg biological, geomorphic, landscape) boundaries (Gergel *et al.* 2002). Our review and analysis point toward research that can help optimize use of landscape tools and data in freshwater assessments. One option is to increase the number of desktop assessments that are validated using field measurements (see WebTable 1 for field-validated studies included in this review); by characterizing the differences in inferential power, users can consider the uncertainty associated with different types of assessments (Wardrop *et al.* 2007; Falcone *et al.* 2010). Another option – development and use of statistical approaches that incorporate multi-scale processes (eg machine learning techniques, Bayesian approaches) – provides new opportunities for assessments that integrate,

Table 1. Selected emerging innovations and the ways in which they may shift assessment practices in the future

	Capacity	*Examples*	*Impact*	*Considerations*
Data availability	Open data sharing	DataONE, USGS Data Portals	Facilitates greater diversity of indicators and stressors	Synthesis and standardization challenges
	Modeled data	Species distribution models, NorWest StreamTemp Regional Database	High-resolution information helps fill knowledge gaps	Model assumptions must be well understood and validated
Data collection	Monitoring arrays	National Ecological Observatory Network	Reduces time lags between assessment and management action	Geographic extent or coverage is typically small
	Increased temporal resolution	Temperature loggers, video feeds	Allows analysis of temporal variability in response to stressors	Current methods are built around discrete sampling events
	Increased spatial resolution	High-resolution (<30 m) DEM and satellite images	Improves predictive landscape models using stressors	Higher resolution information not necessarily aligned with management scales
	On-the-ground technologies	Underwater LiDAR, eDNA, autonomous vehicles	Increases resolution of physical habitat and biological data	Expense, understanding limitations of new technologies
Analysis	Non-linear modelling	Machine learning approaches	Enhances ability to model complex data	Accepted stressor relationships may need to be re-examined
	Expert elicitation	Fuzzy logic, Bayesian models	Supports systematic use of traditional or expert knowledge	Repeatability, defensibility, learning curve
	Data visualization	GUIs, Shiny package (R Statistical Computing)	Facilitates interaction of stakeholders with assessment process	Requires time and expertise

Notes: DEM = digital elevation model; eDNA = environmental DNA; GUI = graphical user interface; LiDAR = light detection and ranging; USGS = US Geological Survey.

rather than rely exclusively on, landscape data (Table 1) (Brown *et al.* 2012; Vander Laan *et al.* 2013).

Our analyses highlight relatively large gaps between integrity assessments, freshwater policies, and management needs, a challenge that has also been qualitatively noted in prior reviews and research (Kusler and Niering 1998; Bain *et al.* 2000; Hruby 2009). Quantitative approaches to assessing this issue are rarer; however, in a unique evaluation of an extensive monitoring program in the Pacific Northwest, Irvine *et al.* (2015) empirically tested and found only weak relationships between long-term monitoring data and original conceptual models. Likewise, Wagner *et al.* (2013) showed that freshwater fish monitoring programs generally lack the power to statistically detect the linear time trends that related to management objectives. These studies, in conjunction with our own analysis, support the idea of a fundamental disconnect between typical "status and trends" assessment and management-driven questions, which has been suggested in previous conceptual work (Noss 1990; Nichols and Williams 2006).

It is difficult to develop approaches that are unbiased and that – at the same time – guide management decisions and conservation policy; however, doing so is critical as more agencies and institutions adopt ecological integrity into management directives not only in North America but also worldwide (Keiter 2004; Dolédec and Statzner 2010). Based on our review, some specific recommendations to improve applicability of assessments include choosing policy-relevant indicators and outputs (Brooks *et al.* 1998; Dale and Beyeler 2001; Hruby 2009), using methods that clearly separate indicators (responses) from stressors (causal mechanisms), and ensuring that management questions closely inform the assessment from the beginning (Barbour *et al.* 2004; Stein *et al.* 2009; Schröter *et al.* 2016). In addition, because many freshwater protection and conservation programs in the US are specific to, or implemented at, the state level (Yoder and Barbour 2009), closer collaborations with state agencies in assessments (whether conducted within the confines of a state or at a regional scale) could increase policy connections and take advantage of underutilized conservation opportunities (Steiner *et al.* 1994).

We expect that our knowledge and capacity to measure ecological integrity will continue to grow over time, driven in part by changes in technology (Table 1). As we move into the future, it is important to learn from the past, to ensure that we can remain true to a consistent understanding of integrity, even as standards and modes of assessment change. This retrospective analysis indicates that integration of assessment results into management and conservation of fresh waters has not kept pace with methodological advances in assessing condition, and we recommend that researchers, managers, and assessment practitioners address this when developing the next generation of integrity assessments.

Acknowledgements

This work was funded by the Landscape Conservation Cooperative (LCC) Network through a national grant opportunity offered by the US Fish and Wildlife Service. JDO was also supported by the H Mason Keeler Endowed Professorship (School of Aquatic and Fishery Sciences, University of Washington). We gratefully acknowledge the participation and support of the network of regional LCC Directors and Science Coordinators who generously gave their time to offer feedback on the review process, as well as the many assessment practitioners who volunteered their expertise by completing the expert survey.

References

Allan JD, Yuan LL, Black P, *et al.* 2012. Investigating the relationships between environmental stressors and stream condition using Bayesian belief networks. *Freshwater Biol* **57**: 58–73.

Angermeier PL and Karr JR. 1996. Biological integrity versus biological diversity as policy directives: protecting biotic resources. In: Samson FB and Knopf FL (Eds). Ecosystem management. New York, NY: Springer.

Bain MB, Harig AL, Loucks DP, *et al.* 2000. Aquatic ecosystem protection and restoration: advances in methods for assessment and evaluation. *Environ Sci Policy* **3**: 89–98.

Barbour MT, Norton SB, Preston HR, and Thornton KW (Eds). 2004. Ecological assessment of aquatic resources: linking science to decision-making. Pensacola, FL: SETAC Press.

Barbour MT, Swietlik WF, Jackson SK, *et al.* 2000. Measuring the attainment of biological integrity in the USA: a critical element of ecological integrity. *Hydrobiologia* **422/423**: 453–64.

Brooks RP, O'Connell TJ, Wardrop DH, and Jackson LE. 1998. Towards a regional index of biological integrity: the example of forested riparian ecosystems. In: Sandhu S, Jackson L, Austin K, *et al.* (Eds). Monitoring ecological condition at regional scales. Proceedings of the Third Symposium on the Environmental Monitoring and Assessment Program (EMAP); Albany, NY; 8–11 Apr 1997. Dordrecht, The Netherlands: Springer.

Brooks RP, Wardrop DH, and Cole CA. 2006. Inventorying and monitoring wetland condition and restoration potential on a watershed basis with examples from Spring Creek Watershed, Pennsylvania, USA. *Environ Manage* **38**: 673–87.

Brown LR, May JT, Rehn AC, *et al.* 2012. Predicting biological condition in southern California streams. *Landscape Urban Plan* **108**: 17–27.

Carletti A, Leo GAD, and Ferrari I. 2004. A critical review of representative wetland rapid assessment methods in North America. *Aquat Conserv* **14**: S103–13.

Clapcott JE, Collier KJ, Death RG, *et al.* 2012. Quantifying relationships between land-use gradients and structural and functional indicators of stream ecological integrity. *Freshwater Biol* **57**: 74–90.

Dale VH and Beyeler SC. 2001. Challenges in the development and use of ecological indicators. *Ecol Indic* **1**: 3–10.

Detenbeck NE, Batterman SL, Brady VJ, *et al.* 2000. A test of watershed classification systems for ecological risk assessment. *Environ Toxicol Chem* **19**: 1174–81.

Dolédec S and Statzner B. 2010. Responses of freshwater biota to human disturbances: contribution of J-NABS to developments in ecological integrity assessments. *J N Am Benthol Soc* **29**: 286–311.

Esselman PC, Infante DM, Wang L, *et al.* 2011. An index of cumulative disturbance to river fish habitats of the conterminous United States from landscape anthropogenic activities. *Ecol Restor* **29**: 133–51.

Falcone JA, Carlisle DM, and Weber LC. 2010. Quantifying human disturbance in watersheds: variable selection and performance of a GIS-based disturbance index for predicting the biological condition of perennial streams. *Ecol Indic* **10**: 264–73.

Frazier TW. 1998. Protecting ecological integrity within the balancing function of property law. *Environ Law* **28**: 53–112.

Gergel SE, Turner MG, Miller JR, *et al.* 2002. Landscape indicators of human impacts to riverine systems. *Aquat Sci* **64**: 118–28.

Hampton SE, Strasser CA, Tewksbury JJ, *et al.* 2013. Big data and the future of ecology. *Front Ecol Environ* **11**: 156–62.

Heino J, Virkkala R, and Toivonen H. 2009. Climate change and freshwater biodiversity: detected patterns, future trends and adaptations in northern regions. *Biol Rev* **84**: 39–54.

Hruby T. 2009. Developing rapid methods for analyzing upland riparian functions and values. *Environ Manage* **43**: 1219–43.

Hughes RM, Paulsen SG, and Stoddard JL. 2000. EMAP-Surface Waters: a multiassemblage, probability survey of ecological integrity in the USA. In: Jungwirth M, Muhar S, and Schmutz S (Eds). Assessing the ecological integrity of running waters. Proceedings of the international conference; Vienna, Austria; 9–11 Nov 1998. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Innis SA, Naiman RJ, and Elliott SR. 2000. Indicators and assessment methods for measuring the ecological integrity of semi-aquatic terrestrial environments. *Hydrobiologia* **422**: 111–31.

Irvine KM, Miller SW, Al-Chokhachy RK, *et al.* 2015. Empirical evaluation of the conceptual model underpinning a regional aquatic long-term monitoring program using causal modelling. *Ecol Indic* **50**: 8–23.

Karr JR. 1993. Defining and assessing ecological integrity: beyond water quality. *Environ Toxicol Chem* **12**: 1521–31.

Karr JR and Dudley DR. 1981. Ecological perspective on water quality goals. *Environ Manage* **5**: 55–68.

Keiter RB. 2004. Ecological concepts, legal standards, and public land law: an analysis and assessment. *Nat Resour J* **44**: 943–88.

Kusler J and Niering W. 1998. Wetland assessment: have we lost our way? *National Wetlands Newsletter* **20**: 1–9.

Lackey RT. 2001. Values, policy, and ecosystem health. *BioScience* **51**: 437–43.

Leibowitz SG and Hyman JB. 1999. Use of scale invariance in evaluating judgement indicators. *Environ Monit Assess* **58**: 283–303.

Liu S, Costanza R, Farber S, and Troy A. 2010. Valuing ecosystem services. *Ann NY Acad Sci* **1185**: 54–78.

Maes J, Liquete C, Teller A, *et al.* 2016. An indicator framework for assessing ecosystem services in support of the EU Biodiversity Strategy to 2020. *ISS Environ Sci Tech* **17**: 14–23.

McDonald M, Blair R, Bolgrien D, *et al.* 2004. The US Environmental Protection Agency's Environmental Monitoring and Assessment Program. In: Wiersma GB (Ed). Environmental monitoring. Boca Raton, FL: CRC Press.

Morgan JA and Hough P. 2015. Compensatory mitigation performance: the state of the science. *National Wetlands Newsletter* **37**: 5–13.

Naiman RJ and Dudgeon D. 2011. Global alteration of freshwaters: influences on human and environmental well-being. *Ecol Res* **26**: 865–73.

Nichols J and Williams B. 2006. Monitoring for conservation. *Trends Ecol Evol* **21**: 668–73.

Noss RF. 1990. Indicators for monitoring biodiversity: a hierarchical approach. *Conserv Biol* **4**: 355–64.

Peterson EE, Ver Hoef JM, Isaak DJ, *et al.* 2013. Modelling dendritic ecological networks in space: an integrated network perspective. *Ecol Lett* **16**: 707–19.

Schröter M, Albert C, Marques A, *et al.* 2016. National ecosystem assessments in Europe: a review. *BioScience* **66**: 813–28.

Schueler TR, Fraley-McNeal L, and Cappiella K. 2009. Is impervious cover still important? Review of recent research. *J Hydrol Eng* **14**: 309–15.

Stein ED, Brinson M, Rains MC, *et al.* 2009. Wetland assessment alphabet soup: how to choose (or not choose) the right assessment method. *Soc Wetland Sci Bull* **26**: 20–24.

Steiner F, Pieart S, Cook E, *et al.* 1994. State wetlands and riparian area protection programs. *Environ Manage* **18**: 183–201.

Stoddard JL, Herlihy AT, Peck DV, *et al.* 2008. A process for creating multimetric indices for large-scale aquatic surveys. *J N Am Benthol Soc* **27**: 878–91.

Suter GW. 1993. A critique of ecosystem health concepts and indexes. *Environ Toxicol Chem* **12**: 1533–39.

Theobald DM, Stevens Jr DL, White D, *et al.* 2007. Using GIS to generate spatially balanced random survey designs for natural resource applications. *Environ Manage* **40**: 134–46.

Tulloch VJ, Tulloch AI, Visconti P, *et al.* 2015. Why do we map threats? Linking threat mapping with actions to make better conservation decisions. *Front Ecol Environ* **13**: 91–99.

Vander Laan JJ, Hawkins CP, Olson JR, and Hill RA. 2013. Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams. *Freshw Sci* **32**: 801–20.

Wagner T, Irwin BJ, Bence JR, and Hayes DB. 2013. Detecting temporal trends in freshwater fisheries surveys: statistical power and the important linkages between management questions and monitoring objectives. *Fisheries* **38**: 309–19.

Wardrop DH, Kentula ME, Stevens DL, *et al.* 2007. Assessment of wetland condition: an example from the Upper Juniata Watershed in Pennsylvania, USA. *Wetlands* **27**: 416–31.

Whittier TR, Hughes RM, Stoddard JL, *et al.* 2007. A structured approach for developing indices of biotic integrity: three examples from streams and rivers in the western USA. *T Am Fish Soc* **136**: 718–35.

Woodley S. 2010. Ecological integrity and Canada's national parks. *The George Wright Forum* **27**: 151–60.

Yoder CO and Barbour MT. 2009. Critical technical elements of state bioassessment programs: a process to evaluate program rigor and comparability. *Environ Monit Assess* **150**: 31–42.

Supporting Information

Additional, web-only material may be found in the online version of this article at http://onlinelibrary.wiley.com/doi/10.1002/fee.1483/suppinfo