Quantile-based grading improves the effectiveness of a multimetric index as a tool for communicating estuarine condition

Chris S. Hallett [a,1]

[a] Centre for Fish and Fisheries Research, School of Veterinary and Life Sciences, Murdoch University, South Street, Murdoch 6150, Western Australia, Australia

E-mail: c.hallett@murdoch.edu.au Telephone: +31 (0)6 1887 5017

Abstract

Multimetric Fish Community Indices (FCI) were recently developed for assessing the ecological condition of shallow nearshore and deeper offshore waters of the Swan-Canning Estuary, Western Australia. The provisional system for classifying estuarine condition from FCI scores, which divided the possible range of scores (0-100) into four descriptive classes of equal breadth (good, fair, poor, very poor), was shown to be skewed towards producing fair to good grades. An alternative, alphanumeric (A-E) grading system, whose grade boundaries were defined by quantiles of the distribution of historical FCI scores, exhibited greater apparent sensitivity to decreases in ecological condition resulting from a harmful algal bloom than did the provisional classification scheme. These advantages of the quantile-based FCIs have led to their recent implementation as a monitoring and reporting tool by the primary environmental managers of the Swan-Canning Estuary, and their application to other permanently-open systems across Western Australia is currently being evaluated.

Keywords: fish, health, indicator, sensitivity, estuary, Australia.

1. Introduction

Effective indicators yield easily interpretable signals of ecosystem health or condition (Kurtz et al., 2001), thus providing invaluable decision support tools for environmental managers. They can also enable the ecological health of ecosystems to be simply communicated to politicians, stakeholders and the public, e.g. via report cards employing conceptually simple presentation techniques such as letter grades, colour coding and mapping (Longstaff et al., 2010).

Multimetric biotic indices are an example of such indicators and are employed globally to quantify the health of aquatic systems including rivers, lakes, estuaries and marine waters

[1] Present address: Sumatrastraat 260, 2585 CW, Den Haag, Netherlands.

(Birk et al., 2012; Rapport and Hildén, 2013). Multimetric approaches allow quantitative index scores to be converted to descriptive categories (e.g. the 'high' to 'bad' Status categories of the Water Framework Directive, or alphanumeric grades) for summarising ecosystem condition. Appropriate scoring thresholds between grades or classes must thus be determined, and can be achieved in a variety of ways. An optimal approach for determining grading thresholds will balance index sensitivity – the ability to distinguish between differing levels of ecological condition – and variability, an excess of which creates 'noise'. The resultant index is thus sensitive to ecologically significant changes occurring among biotic communities in response to diverse stressors (e.g. algal blooms, hypoxia, pollution), yet robust to natural, fine-scale spatio-temporal variability.

Hallett et al. (2012b) have developed the first fish community-based, multimetric indices for assessing the ecological condition of estuaries in Australia. These Fish Community Indices (FCI), which were first developed for the nearshore (<2 m depth) and offshore (>2 m depth) waters of the Swan-Canning Estuary, Western Australia (WA; Fig. 1 in Hallett et al. [2012b]), are broadly applicable to estuaries across WA and beyond. The nearshore and offshore FCIs comprised respective suites of 11 and 7 fish community metrics, including measures of species richness, diversity and abundance, trophic structure and life history function (for a full account of metric selection, reference conditions and FCI calculation, and the detailed rationale for these indices, see Hallett et al., 2012a; 2012b). Under the provisional condition classification system, the possible range of FCI scores (0-100) was subdivided arbitrarily into four classes of equal breadth (good, fair, poor, very poor). Preliminary validation demonstrated that these FCI classes were robust to natural and sampling-related variability, and sensitive to the effects of relatively short-term, localised environmental perturbations, exemplified by algal blooms (Hallett et al., 2012b). However, as the provisional classification scheme was considered to be skewed towards producing fair to good grades, a comparative evaluation was undertaken of the sensitivity and robustness of the provisional condition classifications, against those from an alternative, alphanumeric grading system whose grade boundaries were determined statistically from the distributions of observed, historical FCI scores.

2. Material and methods

Development of the alternative, quantile-based grading approach employed FCI scores calculated from the data used by Hallett et al. (2012a; 2012b) to select metrics, establish reference conditions and devise the original provisional classification scheme for the indices,

namely those data derived from historical samples of the nearshore and offshore fish communities collected throughout the Swan-Canning Estuary between 1977 and 2009. Note that the nearshore data set had previously been subjected to novel standardisation procedures to minimise biases arising from multiple gear types over that period (Hallett and Hall, 2012).

For each of the nearshore and offshore FCIs, an alphanumeric grading system was developed with five respective grades (A-E) representing very good to very poor ecological condition, whereby the respective boundaries for grades A and E comprised the 90th and 10th percentiles of the index scores from the historical data sets. Boundaries for grades B-D were determined by dividing the remaining 80% of historical index scores into three equal quantiles, each containing 26.67% of the observed historical scores. Under this scheme, the proximity of an index score to grade boundaries was also considered when determining condition grades. Scores within one point of a grade boundary were allocated an intermediate grade, denoted using the symbol '/', e.g. a mean score within one point over the boundary score between grades B and C would be denoted 'B/C', whereas a mean score within one point below the same boundary score would receive the condition grade 'C/B'.

The provisional and alternative classification approaches then were evaluated by comparing their effects on the sensitivity and robustness of FCIs calculated from an independent, 'validation' data set, namely nearshore and offshore fish community data collected throughout the Swan-Canning Estuary during the austral summer and autumn of 2011 and 2012 (see section 2.2 of Hallett et al. [2012b] for details of sampling procedures). First, the sensitivities of the two approaches were evaluated for the nearshore FCI by comparing the ecological condition assessments for samples collected prior to, during and after a bloom of the dinoflagellate *Karlodinium veneficum*, which occurred in the Canning Estuary (CE) zone during May 2011 (Hallett et al., 2012b). The sensitivity of the offshore index could not also be assessed in this manner due to a lack of repeated sampling of the offshore waters of this zone during and after the bloom. The robustness of each of the approaches then was evaluated, for both the nearshore and offshore indices, by examining temporal patterns in the condition assessments for individual estuary zones and for the whole system across the validation period. Finally, the optimal scheme was considered to be that which resulted in indices that are (i) most sensitive to human stressors (here exemplified by algal blooms), (ii) robust to the effects of natural variability and (iii) informative, visual and easily understood by the wider community.

3. Results and discussion

Although modifying the grade/class boundaries for an index such as this does not strictly alter its sensitivity (i.e. the response of index scores to degradation), alternative grading systems may change its 'apparent sensitivity', or the ability of the index to communicate effectively the degree of perturbation. The true sensitivity of the FCIs or any other similar measure is a characteristic of the quantitative index scores, and not of the resulting qualitative condition classifications/grades. The former are based directly and objectively upon fish species abundance data collected via field sampling, such that a decrease in index scores reflects a putative response of the fish community to a decline in the ecological condition of the estuary: the larger the decrease in index score, the larger the indicated decline in ecological condition. An index is insensitive only if its scores exhibit little or no response to a measurable ecological perturbation. In contrast, condition grades are a subjective interpretation of what the index scores tell us about ecological condition, and are dependent on the grading scale employed. For example, suppose one were to develop a theoretical 0-100 scoring scheme that had only two grades/classes ('high', 'low') separated by a boundary score of 50 points, and a second scheme with ten grades separated by boundaries every 10 points. Two samples which returned respective index scores of 95 and 51 before and after an ecological perturbation would both receive the same 'high' classification under the former scheme but would be separated by four grades under the latter. In such an instance, the sensitivity of the index to the ecological perturbation has not changed, but the ability of our classification/grading scheme to effectively communicate the magnitude of the perturbation (i.e. the 'apparent sensitivity' of the index) has.

Given the above distinction, the apparent sensitivities of the provisional and quantile-based classification schemes differed markedly. The provisional system was skewed towards fair to good classifications, with the large majority of both nearshore and offshore historical samples being categorized as fair (Fig. 1a and b). Similarly, ~90% and 80% of the respective nearshore and offshore scores from the 2011-2012 validation data sets fell in the top two categories (good, fair), with almost no samples allocated to very poor condition (Table 1). This contrasts with the extensively-modified nature of this estuary (NLWRA, 2002), and suggests that an assessment of very poor condition would be made only on the rare occasions that an extremely low index score (<25) was observed. The provisional classification scheme thus appears relatively insensitive to differing levels of ecological stress, reducing its utility as a management tool.

In contrast, the quantile-based grading system possessed greater apparent sensitivity to ecological condition and was far less skewed than the provisional classification scheme, with

all five grades being awarded regularly across nearshore and offshore historical samples (Fig. 1c and d). Whereas only 10% of the 190 samples in the nearshore validation data set received poor or very poor classifications under the provisional scheme (with only one sample being classed as very poor), the bottom two quantile-based grades accounted for ~25% of samples in the same validation data set (Table 1).

The greater apparent sensitivity of the quantile-based grading scheme is confirmed by patterns in the condition grades observed across nearshore sites in the CE zone before, during and after the *K. veneficum* bloom of May 2011. Under the quantile-based grading system, the overall ecological condition of the CE consistently received a grade B across repeated sampling occasions prior to the bloom, with each individual site being graded A or B (Figs. 2a and b). Following the onset of the bloom, the ecological condition of some sites close to the centre of the bloom then decreased to a D or E grade and the overall condition of the zone declined (Fig. 2c). After the collapse of the bloom the condition of the CE zone subsequently recovered to its pre-bloom grade of B (Fig. 2d). In contrast, the provisional system classified the overall condition of the CE zone as fair throughout this period (Hallett et al., 2012b), thus failing to adequately capture the ecological significance of this notable bloom event, during which peak densities of *K. veneficum* cells exceeded management thresholds and triggered a management response (K. Trayler, Swan River Trust, pers. comm.).

The lack of skew and greater apparent sensitivity of the quantile-based grading system are a result of its condition grades being more numerous (5 vs 4) and of uneven breadth, compared to the provisional condition classes. Consequently, the lowest grade (E) was defined by scores of <45.5 and <36.8 for the nearshore and offshore waters, respectively (Table 1), as opposed to scores of <25 for the lowest class under the provisional scheme. Moreover, there was a far narrower range of scores representing each of the intermediate grades B, C and D (Fig. 1c and d), compared to those representing 'fair' and 'poor' condition under the provisional system (Fig. 1a and b), thus enabling ecologically significant changes in estuary condition to be communicated with greater resolution under the new system.

A potential weakness of a grading scheme that provides a high degree of resolution of spatial and temporal differences in ecological condition is that it may exhibit a high degree of variability in response to natural variability among fish communities, as small changes in index scores might cause frequent changes in condition grades. However, while the condition grades of a few sites in the CE zone changed between sampling occasions in the weeks preceding the bloom, the overall condition for the zone did not change in this time, indicating that the quantile-based grading system is also relatively robust (cf. Figs. 2a and b). Similarly,

the nearshore and offshore condition grades for each estuarine zone were relatively consistent across repeated sampling occasions, both within and between consecutive seasons in 2011-12, with the majority of grades staying the same or changing by only one grade (Table 2).

The quantile-based FCI grading system thus provides a monitoring and communication tool which is sensitive enough to distinguish the ecologically significant changes in estuarine condition that result from both chronic and acute human stressors (exemplified here by algal blooms), yet which is also sufficiently robust to the effects of natural variability. Moreover, the resulting alphanumeric outputs are consistent with the requirements of proposed estuarine report cards for this system and are likely to be easily understood by the public and a broad range of stakeholders. These advantages have led the Swan River Trust, the statutory body with primary responsibility for managing the Swan-Canning Estuary, to adopt and implement the quantile-based FCIs as an ecological monitoring tool for this system. Ongoing monitoring is revealing a slight, yet consistent, improvement in the condition of the estuary since 2005 (Hallett, unpublished data), demonstrating the ability of the FCIs to reflect changes in chronic stressors affecting the system and the possible effects of management actions.

These indices are also broadly applicable to other estuaries across WA and beyond. Work is currently underway to evaluate the degree to which the FCIs developed for the Swan-Canning Estuary are directly applicable to other permanently open estuaries across the state, or otherwise how they must be modified to best suit these other systems. In the short-term, the reference conditions and quantile-based scoring thresholds derived for the Swan-Canning Estuary may be applied to other permanently open systems until a more comprehensive and consistent data set has been compiled from multiple estuaries across WA. The latter will enable the development of FCIs based on shared reference conditions and scoring thresholds, which are generally applicable across all open systems and will allow direct comparisons of ecological health among such estuaries. Moreover, the approaches used to develop these FCIs might be modified further (e.g. by defining type-specific reference conditions and refining component metrics), for application to other estuaries that experience protracted periods of isolation from the ocean due to sand bar formation.

Acknowledgements

Thanks to all those from the Centre for Fish and Fisheries Research who have collected fish community data over several decades; to the Swan River Trust for funding this work; to Kerry Trayler and Shaun Meredith for their enthusiasm and constructive criticism during

index development, and to Fiona Valesini, James Tweedley and two anonymous reviewers for their helpful comments on the manuscript.

References

Birk, S., Bonne, W., Borja, A., Brucet, S., Courrat, A., Poikane, S., Solimini, A., van de Bund, W., Zampoukas, N., Hering, D. 2012. Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. Ecological Indicators 18, 31-41.

Hallett, C.S., Hall, N.G. 2012. Equivalence factors for standardising catch data across multiple beach seine nets to account for differences in relative bias. Estuarine, Coastal and Shelf Science 104-105, 114-122.

Hallett, C.S., Valesini, F.J., Clarke, K.R. 2012a. A method for selecting health index metrics in the absence of independent measures of ecological condition. Ecological Indicators 19, 240-252.

Hallett, C.S., Valesini, F.J., Clarke, K.R., Hesp, S.A., Hoeksema, S.D. 2012b. Development and validation of a fish-based, multimetric index for assessing the ecological health of Western Australian estuaries. Estuarine, Coastal and Shelf Science 104-105, 102-113.

Kurtz, J.C., Jackson, L.E., Fisher, W.S. 2001. Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency's Office of Research and Development. Ecological Indicators 1, 49-60.

Longstaff, B.J., Carruthers, T.J.B., Dennison, W.C., Lookingbill, T.R., Hawkey, J.M., Thomas, J.E., Wicks, E.C., Woerner, J. (eds). 2010. Integrating and applying science: A handbook for effective coastal ecosystem assessment. IAN Press, Cambridge, Maryland, 244 pp.

NLWRA. 2002. Australian Catchment, River and Estuary Assessment 2002, volume 1. National Land and Water Resources Audit, Canberra.

Rapport, D.J., Hildén, M. 2013. An evolving role for ecological indicators: From documenting ecological conditions to monitoring drivers and policy responses. Ecological Indicators 28, 10-15.

Figure captions

Figure 1. Frequency distributions of nearshore (a, c) and offshore (b, d) Fish Community Index scores from all samples collected between 1977 and 2009 in the Swan-Canning Estuary. Dashed lines indicate the scoring boundaries for the provisional condition classifications (a, b) and the quantile-based condition grades (c, d).

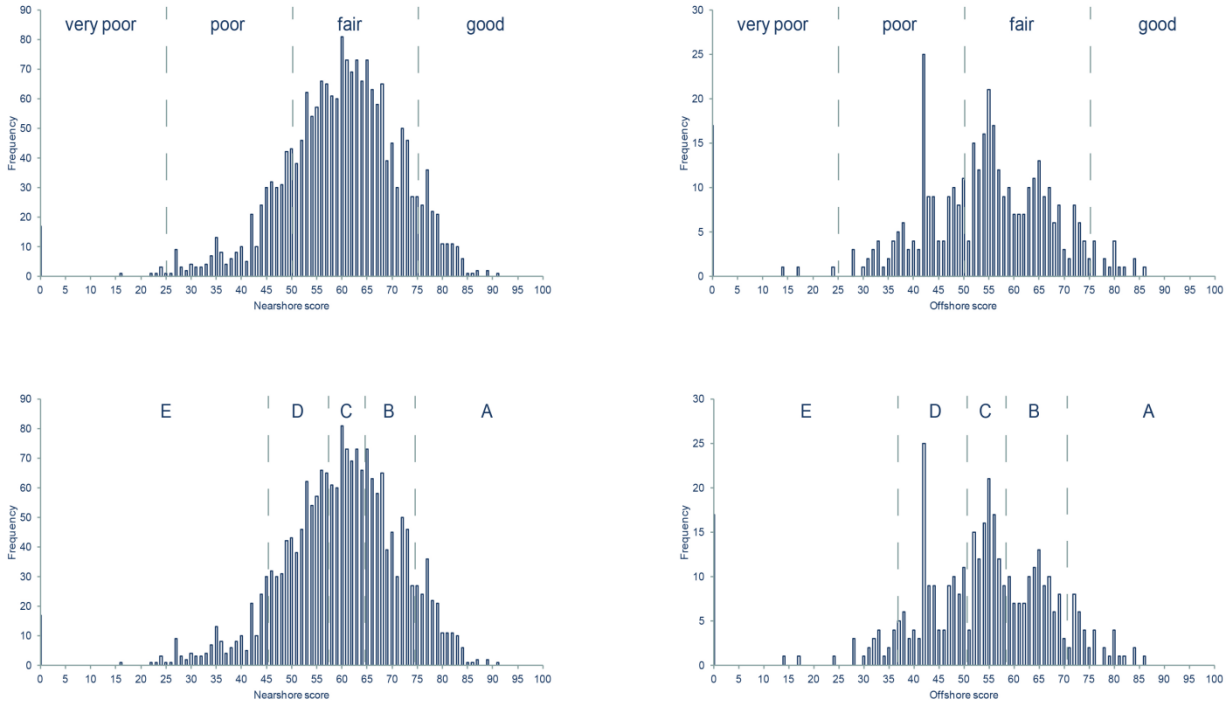
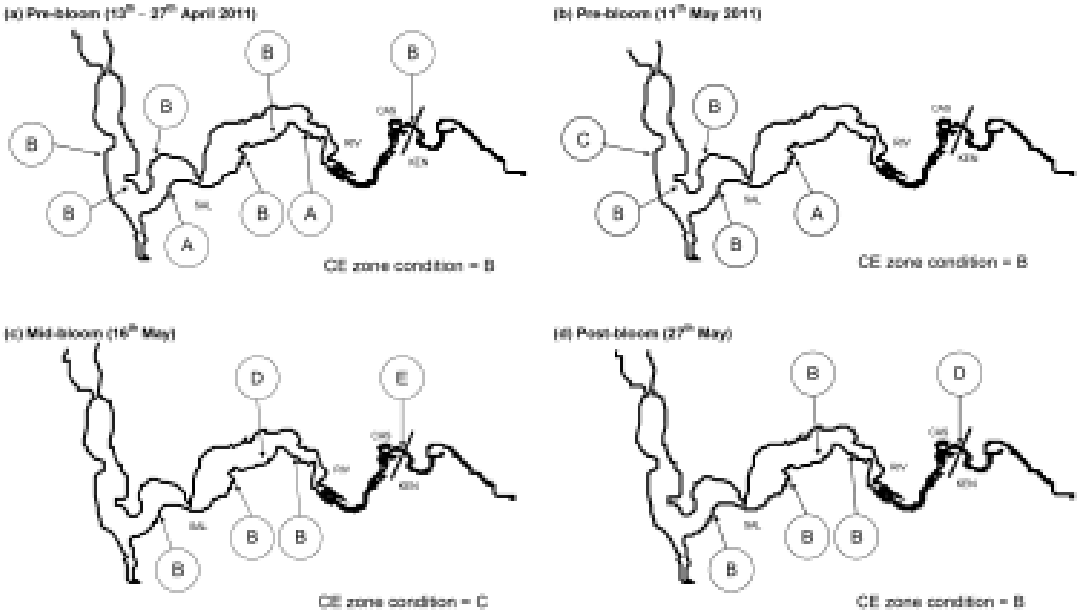


Figure 2. Maps of the Canning Estuary (CE) zone of the Swan-Canning Estuary, illustrating nearshore Fish Community Index condition grades (A-E) derived via the quantile-based grading system, for sites sampled before (a, b), during (c) and after (d) a *Karlodinium veneficum* bloom in May 2011. The overall condition grade for the CE zone, based on the mean index score across sites, is also shown.



Tables

Table 1. Nearshore and offshore Fish Community Index scores comprising each of the condition classes or grades under the provisional classification and quantile-based grading schemes. The percentage of samples in the 2011-12 validation data sets (nearshore, *n* = 190; offshore, *n* = 146) that were awarded each class or grade is also shown.

Provisional classification scheme				
	Nearshore		*Offshore*	
Class	Boundary scores	Validation %	Boundary scores	Validation %
Good	>75	13.7	>75	6.8
Fair	50-75	76.3	50-75	72.6
Poor	25-50	9.5	25-50	20.5
V. poor	<25	0.5	<25	0

Quantile-based grading scheme				
	Nearshore		*Offshore*	
Grade	Boundary scores	Validation %	Boundary scores	Validation %
A	>74.5	14.7	>70.7	13.0
B	64.6-74.5	37.9	58.4-70.7	41.1
C	57.1-64.6	22.6	50.6-58.4	21.2
D	45.5-57.1	18.9	36.8-50.6	23.3
E	<45.5	5.8	<36.8	1.4

Table 2. Nearshore and offshore condition grades (A-E), determined using the quantile-based grading system, for individual zones of the Swan-Canning Estuary and for the estuary as a whole on repeated sampling occasions during summer (Su) and autumn (Au) of 2011 and 2012.

Zone	*Nearshore*						*Offshore*					
	Su 2011		**Au 2011**		**Su 2012**	**Au 2012**	**Su 2011**		**Au 2011**		**Su 2012**	**Au 2012**
	Mth 1	Mth 2	Mth 1	Mth 2			Mth 1	Mth 2	Mth 1	Mth 2		
LSCE	B	C	B	C	B	C	B	B	B	C	B	B
CE	C	C	B	B	B	B	C	D/C	C	C	C/D	C
MSE	C	C	C	B	B	C	B	B	C/D	C/B	B	B
USE	B	B	C	D	D	B	B	B	B	C/D	B	C/B
Estuary	**C**	**C**	**B**	**C**	**B**	**B**	**B**	**B**	**C**	**C**	**B**	**B**