Development of a multimeric index for assessing the condition of the Vasse-Wonnerup based on benthic invertebrate communities

S. Cronin-O'Reilly, A. Cottingham & J. Tweedley

Centre for Sustainable Aquatic Ecosystems, Harry Butler Institute,
Murdoch University
Environmental & Conservation Sciences, College of Science, Health, Engineering &
Education, Murdoch University

March 2023

Final Report

Prepared for the Department of Water and Environmental Regulation

Contents

Acknowledgements

The authors gratefully acknowledge the support of the Science Advisory Group for the Vasse Taskforce, in particular Dr Kath Lynch (Department of Water and Environmental Regulation). Gratitude is expressed to Dr Linda Kalnejais, Brian Poh, Theo Campbell, Angus Lawrie, Kurt Krispyn, Charles Maus and Dillon Newman for helping conduct the fieldwork and to Dr Emily Hugues dit Ciles, Dr Joanna Browne and Jane Wilshaw and helping to coordinate the sampling program. Dr Linda Kalnejais is thanked for providing data on the status of the sand bar and water level and quality throughout the system as are Drs Kath Lynch, Dr Linda Kalnejais and Dr Emily Hugues Dit Ciles for reviewing the draft final report. This research was commissioned by the Department of Water and Environmental Regulation on behalf of the Vasse Wonnerup wetlands Partnership with funding through the Revitalising Geographe Waterways Program.

Summary

This project is part of the Revitalising Geographe Waterways' Integrated Ecological Monitoring Study (IEM), which aims to better understand the relationships between water regime, food sources and abundance of benthic macroinvertebrates (> 500 µm), fish and birds utilising the range of habitats (regions) present in the Vasse-Wonnerup (see https://rgw.dwer.wa.gov.au/applying-science/vasse-wonnerup-science/). The benthic macroinvertebrate index component aims to:

- Determine the characteristics of the subtidal benthic macroinvertebrate community (*i.e.* community metrics) that best reflect differences in the environmental condition of the Vasse-Wonnerup, using all available macroinvertebrate samples collected during previous years (2017-2021) of the IEM monitoring.
- Combine these responsive metrics into a new index to assess benthic ecological condition of the system and calculate season and region-specific reference conditions.
- Evaluate how index values change through time and in response to water quality.
- Create a report card (Table 1 & 2) that summarises index scores as letter grades of health (A-E) to assist future monitoring of the Vasse-Wonnerup.

Five metrics were found to reflect differences in environmental condition in regions of the Vasse-Wonnerup above the surge barriers and were included in the Wetland Benthic Community Index (WBCI). These metrics were Simpson's index, qualitative taxonomic distinctness, community dominance, crustacean richness and SIGNAL2. Trends of WBCI scores were promising and showed expected rises and falls in scores with seasonal stressors. Healthier macroinvertebrate communities were found to occur when the Vasse and Wonnerup estuaries contained water with a salinity < 25 ppt, that was relatively deep, and had high levels of dissolved oxygen and pH.

For the Wonnerup Inlet, which is located below the surge barriers and harbours a distinct and more estuarine invertebrate community, the Estuarine Benthic Community Index (EBCI) was calculated, which was developed for the nearby Peel-Harvey Estuary. This biotic multimetric index combines scores of the number of species, Shannon-Wiener diversity and community dominance, and showed large variation within seasons, with no relationships evident with water quality parameters. Specific metric selection is likely required for this region, in order to capture its environmental conditions more conclusively.

Future work proposed includes validation of the WBCI, potential metric selection of the Wonnerup Inlet, and the proposal of a monitoring regime where the developed indices are used in combination with other monitoring tools.

Table 1. Summary report card of health grades, score thresholds, and community characteristics seen for the Wetland Benthic Community Index.

Index Grade	Score thresholds	Community characteristics
A	78-100	High levels of diversity Distinct taxonomic groups present High levels of community dominance Some crustacean species present Some sensitive and tolerant insects present
B	66-77	High levels of diversity Distinct taxonomic groups present High levels of community dominance A few crustacean species present Sensitive insects less abundant while tolerant insects present
C	54-65	Moderate levels of diversity Distinct taxonomic groups present High levels of community dominance Very few or no crustacean species present Tolerant insects mostly present
D	29-54	Some diversity remaining Distinct taxonomic groups present High levels of community dominance Crustacean species rarely present Only tolerant insects persist
E	0-29	No species diversity Distinct taxonomic groups absent No community dominance as generally only one species present Crustaceans absent Only tolerant insects persist

Table 2. Summary report card of health grades, score thresholds, and community characteristics seen for the Estuarine Benthic Community Index.

Index Grade	Score thresholds	Community characteristics
A	87-100	High numbers of species present High levels of diversity High levels of community dominance
B	66-87	Good number of species present Good levels of diversity High levels of community dominance
C	50-66	Moderate numbers of species present Moderate levels of diversity High to good levels of community dominance
D	30-50	Moderate to low numbers of species Low to moderate diversity High to good levels of community dominance
E	0-30	Few to no species present Low to moderate diversity Low community dominance

1. Introduction

The Vasse-Wonnerup is an intermittently-open estuary located near the town of Busselton, Western Australia. The system and surrounding land have been subjected to extensive anthropogenic modification, including land clearing, the creation of extensive drainage networks, including the diversion of several rivers that historically flowed into the system, and the construction of surge barriers, which can prevent seawater intrusion into the estuary (Wetland Research & Management 2007, Tweedley et al. 2017). Moreover, the large amounts of fertilizer applied to agricultural land, combined with animal waste discharged from pastures, has resulted in the Vasse-Wonnerup becoming “the most grossly enriched major wetland system known in Western Australia” (McAlpine et al. 1989). These anthropogenic pressures have had numerous impacts on the ecological health of the Vasse-Wonnerup including, in recent times, several large fish kills (Lane et al. 1997, Tweedley et al. 2014a).

Indicators of ecosystem health are particularly valuable in the temperate, microtidal systems of south-western Australia, as the low tidal amplitude, highly seasonal rainfall and ephemeral connectivity to the ocean results in the magnitude of any anthropogenic influences being amplified (Potter et al. 2015, Tweedley et al. 2016b, Hallett et al. 2018, Warwick et al. 2018). Analysis of the benthic macroinvertebrate community has shown that the health of several estuaries in this region has declined over time (Wildsmith et al. 2009, 2011, Tweedley et al. 2012). Moreover, as well as responding to long-term (*i.e.* inter-decadal) changes in environmental condition due to eutrophication and urbanisation, the composition of the benthic macroinvertebrate fauna can change markedly in response to, and subsequently recover from, hypoxic conditions (Tweedley et al. 2016a, Cronin-O'Reilly et al. 2022).

Given their recognised responses to local stressors, benthic macroinvertebrates may thus act as useful indicators of ecological health in these microtidal estuaries (Tweedley et al. 2012). Benthic macroinvertebrates are predominantly sessile (*i.e.* immotile), have relatively long life spans, and display differential tolerances to stress that enables spatial and temporal changes in ecosystem health to be detected (Tagliapietra et al. 2012, Tweedley et al. 2015). Multimetric biotic indices are tools used to assess the ecological health of aquatic ecosystems by combining various measures of a community’s richness, diversity, tolerance, and function (i.e. community metrics) in order to produce a health score and grade (Hering et al. 2006). Benthic indices, which assess these aspects in the macroinvertebrate community, are some of the most widely applied biotic indices globally, and are used in the monitoring of marine, estuarine and freshwater ecosystems (Borja et al. 2000, Diaz et al. 2004, Birk et al. 2012).

For microtidal estuaries in south-western Australia, benthic indices can act as a robust tool to assess the environmental condition of these systems. To date, two evaluations have assessed the suitability of benthic indices developed elsewhere for these estuaries and have shown them to be inadequate in measuring the health of their highly adaptive and tolerant macroinvertebrate communities (Tweedley et al. 2014b, Cronin-O’Reilly 2021). It is thus appropriate to develop new approaches, in order to reliably assess the benthic ecological health of these estuaries. As a result, the Estuarine Benthic Community Index (EBCI) has been developed for the Peel-Harvey Estuary, and combines scores of the community’s richness, diversity and dominance together to determine a health score (Cronin-O’Reilly 2021). However, the EBCI has yet to be applied in other estuaries in the region and it remains to be seen how it will respond to the geomorphological, physicochemical and ecological variability microtidal estuaries in south-western Australia display. For example, the Vasse-Wonnerup harbours two distinct macroinvertebrate communities, an estuarine and wetland community type (Tweedley et al. 2019a, 2021a), and may require its own benthic index that captures these distinct community differences while assessing the estuary’s environmental condition. Moreover, this wetland community is markedly different to the fauna present in other estuaries (Tweedley et al. 2011, 2020)

Given the usefulness of benthic macroinvertebrates as indicators of environmental health and their resulting benthic indices in long-term monitoring of ecosystem health, the aims of the current study were to:

1. Determine the characteristics of the benthic macroinvertebrate community (*i.e.* community metrics) that best reflect differences in the environmental condition of the Vasse-Wonnerup.
2. Combine these responsive metrics into a new index following determination of region-specific reference conditions in each season to assess the benthic ecological health of the Vasse-Wonnerup.
3. Evaluate how index trends change through time and in response to water quality.
4. Create a report card that summarises index scores as letter grades of health (A-E) to assist future monitoring of the Vasse-Wonnerup.

2. Materials and methods

2.1. Sampling regime

Benthic macroinvertebrates were sampled at four subtidal sites in each of the six regions of the Vasse-Wonnerup over a two-day period in each season between July 2017 and March 2020 (Fig. 1; Tweedley et al. 2021a). Additional samples were collected at the same suite of sites in March 2017, only with two replicate samples being collected at each site. Furthermore, sampling was also conducted in November 2020 and January 2021 at the same sites and at an extra four sites in each region giving a total of eight per region (trends in these data shown in Tweedley et al. 2021b). The regions extended from the 'marine' areas immediately upstream from the sand bar, which connects/disconnects the estuary to the coastal waters of Geographe Bay (*i.e.* Wonnerup Inlet), up the longitudinal axes of both the Vasse and Wonnerup estuaries (*i.e.* Lower and Upper Vasse, and Lower and Upper Wonnerup estuaries). Samples were also collected from the narrow area lying directly upstream from the Vasse surge barrier and running parallel to the coast (*i.e.* Vasse Exit Channel).

During the November 2020 and January 2021 sampling occasions, a rapid assessment protocol (RAP; Hallett et al. 2019c) was carried out to determine the sediment condition at each site. The RAP was developed in the nearby Peel-Harvey Estuary and assesses the colour, texture and odour of the sediment in-field, classifying them into a certain groups, to determine an overall sediment condition class, *i.e.* good, fair, and poor (Hallett et al. 2019c). For this purpose, a 60 mm x 600 mm Plexiglas core of sediment was sampled to assess the sediment's colour, texture, and odour.

On all sampling occasions, a randomly located sample(s) of sediment was collected at each site within a region from subtidal waters (*i.e.* 0.5 – 2 m deep) using an Ekman grab (Wildco, Yulee, Florida, USA) that collected substrata from an area of 225 cm^2 and sampled to a depth of 15 cm. An Ekman grab was preferred to a sediment corer as it enabled samples to be collected from waters deeper than 1.5 m (*i.e.* those in the Vasse Exit Channel). Furthermore, the grab samples a larger surface area of sediment than the corer allowing for the collection of larger, less abundant taxa such as the bivalve *Hiatula biradiata* (see Tweedley et al. 2019a for full rationale).

Each sediment sample was wet sieved in the field and preserved in a 5% formalin mixture buffered in estuary water and, after at least one week, subsequently wet-sieved though a 500 µm mesh and stored in 70% ethanol. Using a dissecting microscope, any invertebrates found in a sample were removed from the sediment retained on the mesh and identified to the lowest practical taxonomic level. Note that many of the keys for members of the Insecta only classify larval forms and/or only allow identification to a relatively high taxonomic level. Species names were adjusted accordingly, and the taxonomic classification of each taxon was determined using the World Register of Marine Species (https://www.marinespecies.org/). Taxa found above the surge barriers were categorized into one of the following feeding groups: gathering collectors (*i.e.* taxa that feed on deposited materials on the sediment), filtering

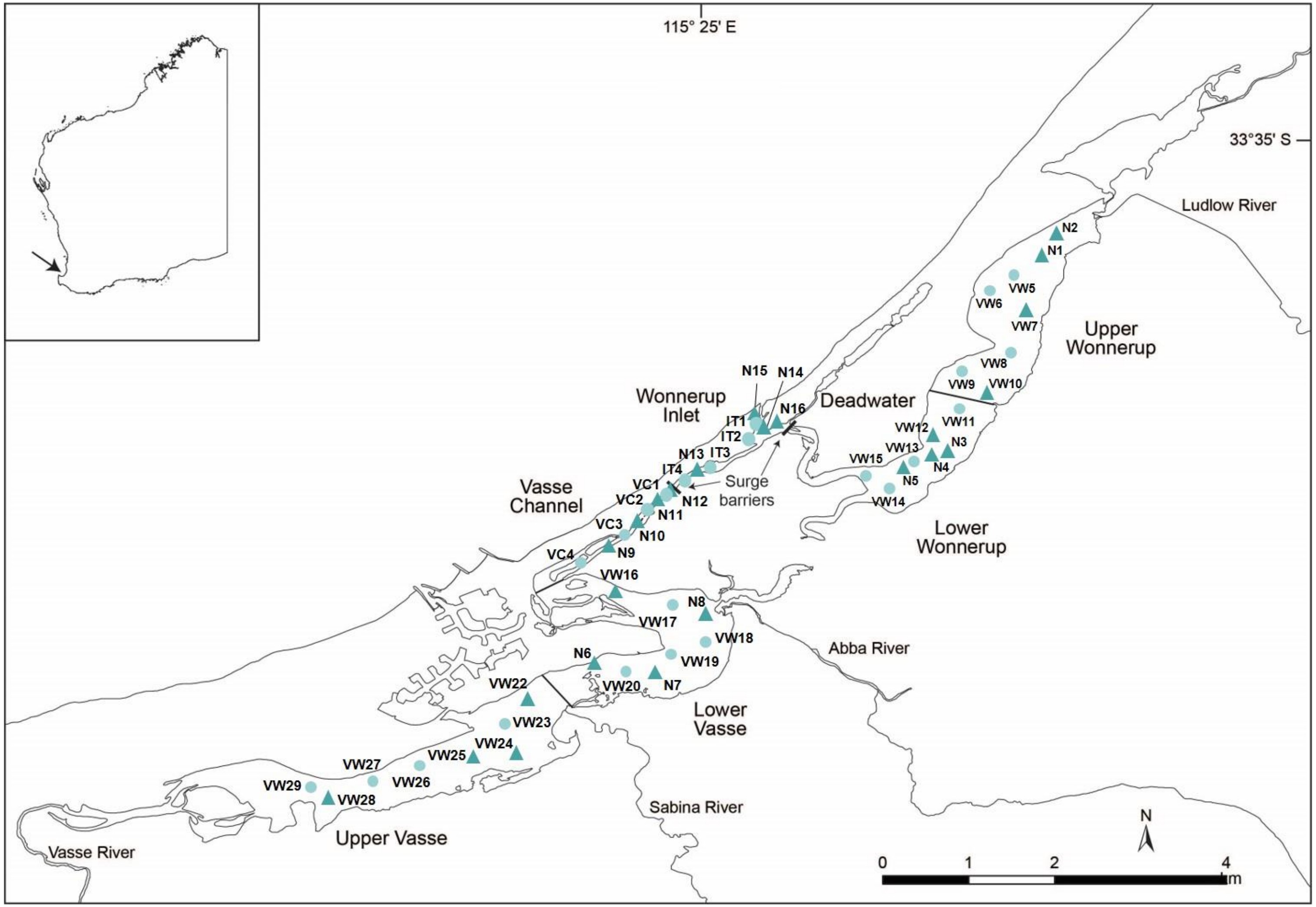


Fig. 1. Map of the Vasse-Wonnerup showing the four sites (●) sampled in each of the six regions at various times between March 2017 and January 2021. The location of additional sites sampled during November 2020 and January 2021 are also shown (▲) and are coded N.

collectors (*i.e.* taxa that feed on materials from water column), shredders (*i.e.* taxa that break down larger pieces of organic material), grazers (*i.e.* taxa that feed on living micro- and macroalgae), predators (*i.e.* taxa that consume other taxa), scavengers (*i.e.* taxa that eat dead or decaying plant or animal material), and groups that use two of these feeding modes or are generalists (*i.e.* use three or more feeding modes).

2.2. Index development

For benthic environments of estuaries in south-western Australia, a multimetric biotic index, *i.e.* the EBCI, has already been developed to assess benthic health in the Peel-Harvey Estuary (Cronin-O'Reilly 2021) and is designed to be relevant for estuaries in south-western Australia, like its counterpart for fish (Hallett et al. 2019b, Tweedley et al. 2021c, 2021d). As the macroinvertebrate communities in Wonnerup Inlet (below the surge barriers) are similar to those of a 'typical' estuary in the region (Tweedley et al. 2019a, 2021a), the EBCI can be calculated for this region (Section 2.8). However, as the macroinvertebrate community upstream of the surge barriers contains a different suite of species and those more typically found in wetlands (Tweedley et al. 2019a), the EBCI would not be suitable. Thus, the Wetland Benthic Community Index (WBCI) was developed specifically for the regions of the Vasse-Wonnerup upstream of the surge barriers. The main steps taken to develop the WBCI are summarised in Figure 2, with further detail provided in the subsequent sections.

All statistical analyses were performed using the PRIMER v7 multivariate statistics software package (Clarke & Gorley 2015) with the PERMANOVA+ add-on (Anderson et al. 2008) and R (R Core Team 2019).

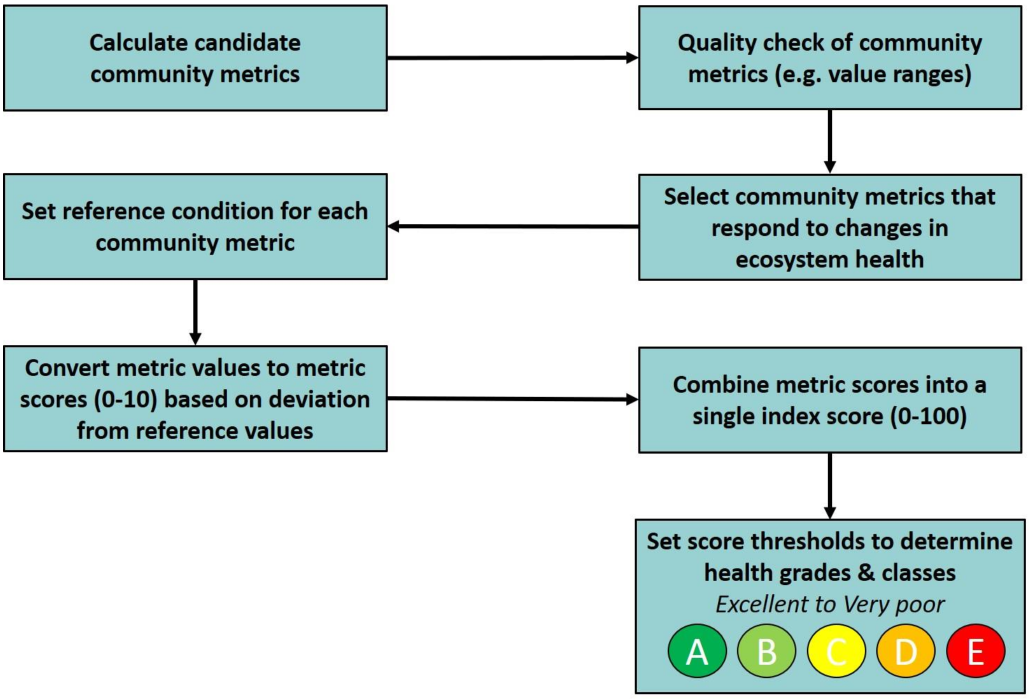


Fig. 2. A flow chart showing the main stages taken during development of the Wetland Benthic Community Index.

2.3. Metric calculation

The untreated, raw abundance data of benthic macroinvertebrate taxa collected during each of the 15 surveys (*i.e.* March 2017, seasonally between July 2017 and March 2020, and in November 2020 and January 2021; N = 360) were used to calculate a suite of univariate metrics of the community structure. The DIVERSE routine was employed to calculate the number of species (*S*), number of invertebrates (*N*; individuals 225 cm^{-2}), Margalef's richness (d), Shannon-Wiener diversity (*H'*), Pielou's Evenness (*J'*), Simpson's (Diversity) index (1-λ'; Somerfield et al. 2008), expected number of species in 50 individuals (ES[50]), quantitative taxonomic distinctness (Δ* or Delta*; Clarke & Warwick 1998), qualitative taxonomic distinctness (Δ^+; Clarke and Warwick, 1998) and variation in taxonomic distinctness (λ^+). Community dominance (D_c) was calculated following the methods described in Ma and Ellison (2018). Based on the taxonomic classification of species, the number and proportion (percentage contribution) of the following taxonomic groups were calculated: annelids, molluscs, arthropods, crustaceans, insects, and amphipods. The abundance and the proportion of the tolerant polychaete *Capitella capitata* (species complex) was also calculated, as well as the richness of annelids, crustaceans, insects and molluscs.

The richness and abundance of insects belonging to the sensitive orders of the Plecoptera (Stoneflies), Ephemeroptera (Mayflies) and Trichoptera (Caddisflies), collectively referred to as PET, were calculated, along with the Stream Invertebrate Grade Number Average Level 2 Index (SIGNAL2; Chessman 2003). For SIGNAL2, weights were applied to each taxon according to their abundance, with these weights multiplied by each taxon's sensitivity grade (sensitive taxa = higher grade). Total SIGNAL2 scores (0-10) for each site were then calculated as the sum of the weight multiplied by SIGNAL2 grades divided by the total sum of weights (Chessman 2003). The abundance and percentage contribution of the following feeding groups were also calculated: gathering collectors, filtering collectors, predators, scavengers, shredders, and grazers (definitions provided earlier). Abundances and proportions of taxa recognised to use two of these groups together, as well as use more than three feeding modes (i.e. generalist taxa) were calculated. A total of 56 community metrics were calculated from the raw abundance data covering richness/diversity, composition, tolerance, and functional aspects of the macroinvertebrate community (Appendix 1).

A quality check was then carried out to remove non-informative, numerically unsuitable metrics (Hering et al. 2006). The range of community metrics were checked using box-and-whisker plots, with those that displayed small ranges, mostly zero values, or had a high number of outliers excluded (Appendix 1). To remove those metrics that were autocorrelated the values for each pairwise combination of metrics was subjected to Pearson correlations (Appendix 2). In cases where the resultant correlation was > 0.9 only one of the two metrics were retained. A refined list of the 12 information-rich, independent community metrics that were considered for metric selection are presented in Table 3.

Table 3. Refined list of candidate community metrics encapsulating richness/diversity, composition, tolerance, and functional aspects of the benthic macroinvertebrate community tested for inclusion into the WBCI.

Community metric	Response to stress	Metric description
Richness/Diversity		
Number of species (*S*)	−	Number of taxa found at each site
Simpson’s (Diversity) index ($1-\lambda'$)	−	The spread of taxon abundances (evenness) among species
Qualitative taxonomic distinctness (Δ^{+})	−	The taxonomic spread of species found at each site
Community dominance (D_c)	+	The level of species dominance present
Crustacean richness	−	Number of crustacean taxa present
Insect richness	−	Number of insect taxa present
Composition		
Proportion of arthropods	−	Percentage contribution of arthropods to the community
Proportion of molluscs	−	Percentage contribution of molluscs to the community
Proportion of insects	−	Percentage contribution of insects to the community
Tolerance		
Proportion of *Capitella capitata*	+	Percentage contribution of the pollution tolerant annelid *C. Capitata* to the community
SIGNAL2 score	−	A tolerance-based index based on the recognised sensitivity and tolerance of various insect groups
Function		
Proportion of gathering collectors	+	Percentage contribution of taxa that feed on deposited materials on the sediment

2.4. Metric selection

Prior to metric selection, replicate samples for sites collected during March 2017 were averaged, so that all sites had one measure for each community metric. Each metric was also assessed to determine whether it exhibited a normal distribution. As a result, Simpson’s (Diversity) Index, the proportion of molluscs and SIGNAL2 were square-root transformed and the proportions of arthropods and insects were fourth-root transformed.

Metric selection requires an independent measure of anthropogenic stress against which the candidate community metrics can be tested. Given that data on sediment condition were only available for two of the 15 surveys and then showed only little discrimination (Appendix 3), a novel approach to select metrics in the absence of a sediment-based stress measure was taken based on the method developed by Hallett et al. (2012). The Vasse Exit Channel is a particularly degraded region of the system, which is independently known to become hypoxic, particularly during darkness and around dawn (Lane et al. 1997). In response, a small temporary artificial oxygenation plant was constructed by the Department of Water and Environmental Regulation to help prevent the occurrence of, or reduce the severity of, any hypoxia during summer and sites in this region. However, in March 2017 the benthic fauna of this region was shown to be highly degraded as a result of marked hypoxia in that occurred in the January of that year (Tweedley et al. 2019a). Numerous fish kills have occurred in this region (Beatty et al. 2018), the most recent of which took place in June 2021 resulting in the death of an estimated 10,000 fish, mainly Black Bream (Cottingham et al. 2021).

On this basis, the Vasse Exit Channel can be used as the most degraded region in which to select metrics that differ between this region and all other regions. To select these metrics, a Euclidean-distance model matrix (0-1) that captured these recognised regional differences in degradation was created. Distance-based linear modelling (DISTLM; Legendre & Anderson 1999, McArdle & Anderson 2001) and Biota-Environment matching routines (BEST; Clarke et al. 2008) were applied to select metrics that explain a significant amount of variation in the model matrix, selecting those that differ mostly between the Vasse Exit Channel and all other regions while remaining relatively consistent within the Vasse Exit Channel and between all other regions. For DISTLM, both marginal (testing metrics separately) and sequential (testing metrics collectively) tests were run, with stepwise used as the selection procedure and R^2 set as the selection criterion. Metrics were considered further if they displayed a significant response ($P \leq 0.05$) collectively in the sequential DISTLM and BEST tests.

It is possible that the metrics selected above are also detecting spatial changes due to naturally occurring regional differences (*e.g.* physicochemical variation, habitat differences). To assess if this was the case, each of the metrics selected from the above procedures were individually subjected to a Permutational Analysis of Variance (PERMANOVA; Anderson 2001) to test for differences in metric values among regions, with a pairwise comparison used to assess the relative magnitude of these regional differences. Ideally, the best metrics for index inclusion would show relatively small to no differences among other regions (i.e. upper and lower Vasse, upper and lower Wonnerup) while differing largely from the Vasse Exit Channel. To complete these tests, a Euclidean distance matrix was created for each metric and subjected to a main and pairwise PERMANOVA test with unrestricted permutations of the data and Type 3 Sums of Squares. The relative magnitude of differences in metric values between regions were assessed using t-values. Means plots with 95% confidence limits of metric values across regions were produced to assist investigating the causes of any differences found.

2.5. Establishing reference conditions

Reference conditions are a benchmark against which current and future trends in health measures are compared to. They can be established using expert judgement, comparing available data between impacted and unimpacted systems, using historical data or through modelling (Muxika et al. 2007). In the current study, reference conditions are set as the 95th percentile of observed metric values in the Vasse-Wonnerup since 2017, which are the best achievable conditions seen in the system. In order to account for spatial and temporal variation, separate reference conditions were determined for each region in each season (Summer; January; Autumn, March; Winter, July; Spring, October-November), pooling data from across the four years.

2.6. Conversion of metric values to scores

Raw values for each community metric were converted to metric scores set to scale between 0 and 10 by assessing their deviation from the reference condition. For this purpose, reference conditions are referred to as the upper anchor (UA) and the lower anchor (LA) is calculated as the 5th percentile of observed values for each community metric.

Negatively responding metrics are then calculated using the following formula:

$$Metric\ score = \frac{Metric\ value\ -\ LA}{UA\ -\ LA} \times 10$$

If metric scores fell outside of the set range (0-10), negative metric scores were manually bounded to 0 and scores greater than 10 were manually bounded to 10.

2.7. Index calculation and grade thresholds

Total index scores are then calculated by summing all metric scores together, dividing by the maximum metric score (e.g. 50 if five metrics were included), and multiplying by 100. The resultant index scores scale between 0 and 100, with higher scores indicative of better benthic ecological health.

The distribution of index scores is then assessed and divided into five health grades and associated classes. The five grades determined were grade A (Excellent), B (Good), C (Moderate), D (Poor), and E (Very poor), with grade thresholds set at unequal quintiles of the data (12.5%, 37.5%, 62.5%, 87.5%, Fig. 3).

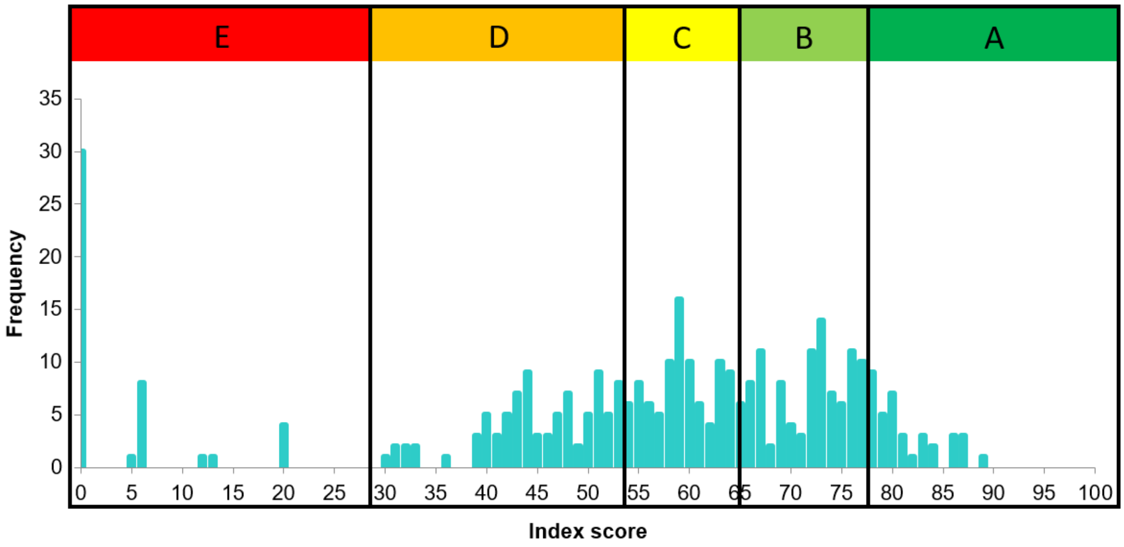


Fig. 3. The histogram distribution of index scores over the full monitoring period (March 2017-Jan 2021), with health grade thresholds set at unequal quintiles of the data (12.5%, 37.5%, 62.5%, 87.5%).

2.8. Index adjustment

The EBCI was developed to assess the health of estuarine macroinvertebrate communities in the Peel-Harvey Estuary, following rigorous testing of the response of community metrics to a quantitative gradient of sediment condition, *i.e.* mud content, organic enrichment, oxygenation and H_2S presence (Cronin-O'Reilly 2021). Given the similarities in the invertebrate species present in Wonnerup Inlet and the Peel-Harvey Estuary, the EBCI was adjusted and calculated for this region. The metrics the EBCI incorporates are species richness (*S*), Shannon-Wiener diversity (*H'*) and community dominance (D_c), all of which decrease with declining sediment condition. To capture natural temporal and spatial shifts in community metrics and potential metric range differences between estuaries, region-specific reference conditions were set for the Wonnerup Inlet by determining the 95th percentile of observed values in each season, i.e. Summer (January), Autumn (March), Winter (July) and Spring (October/November). Using the same approach as the WBCI above, metric values were converted to scores, total index scores were calculated, and health grade thresholds were set at slightly different unequal quintiles due to the spread of index scores, i.e. 10%, 37%, 63%, 90%.

2.9. Index trends with water quality

Seasonal trends of index scores for both the WBCI and EBCI were assessed using means plots over the full monitoring period for the two broad regions (above and below surge barriers) and for each region separately.

The main drivers of the seasonal trends in WBCI and EBCI for three years of the monitoring program data (March 2017-March 2020; N = 242) were described separately using a generalised additive model (GAM) which assumed a Tweedie distribution to allow for the presence of zero scores (Williams et al. 2020). The models were fitted using the 'gam' function in the 'mgcv' package (version 1.8–28) for R (R Core Team 2019). A thin-plate regression spline was employed as the smoothing basis for each of the environmental covariates, i.e. salinity (ppt), temperature (°C), dissolved oxygen (mg L^{-1}), pH, and turbidity (NTU) for both indices, with water level (m AHD), water level change (cm day^{-1}), and shallow area (< 0.1 m2; m^2) also investigated for the WBCI. In additional to a *P*-value, which were considered significant if ≤ 0.05, the proportion of the deviance explained (DE) by the model was also calculated. After this investigation, changes in metric scores for a select 'good' and 'poor' period were assessed in each region using radar plots. For regions above the surge barriers, October 2017 and March 2019 were selected as good and poor periods, respectively, while July 2019 and March 2018 were respectively selected for the Wonnerup Inlet.

3. Results

3.1. Metric calculation

Of the initial 56 community metrics calculated for the WBCI, 38 were found to have an unsuitable range of values and number of outliers (Appendix 1). Most of the excluded metrics were abundance-based metrics (*e.g.* number of individuals, abundance of annelids), which had extreme outliers, making them unsuitable as extreme values may unduly influence statistical analyses. Other metrics excluded at this stage were the proportions and abundances of most feeding groups, whose abundances were too scarce to produce a viable range of values. Of the 18 metrics that remained, correlation coefficients of > 0.9 were found between most of the richness indices (Appendix 2). The expected number of species (ES[50]) was highly correlated with number of species, Margalef's richness, and Shannon-Wiener diversity. The last index was also correlated with Margalef's richness, while Pielou's evenness was correlated with Simpson's index. The number of species and Simpson's index were retained for metric selection. Quantitative taxonomic distinctness and qualitative taxonomic distinctness were also highly correlated, with the latter metric retained. Proportions of annelids, tolerant annelids, and *C. capitata* were also highly correlated, with the proportion of *C. capitata* retained, given this species is the most abundant annelid in the system and is a well-recognised, pollution indicator species (Pearson & Rosenberg 1978).

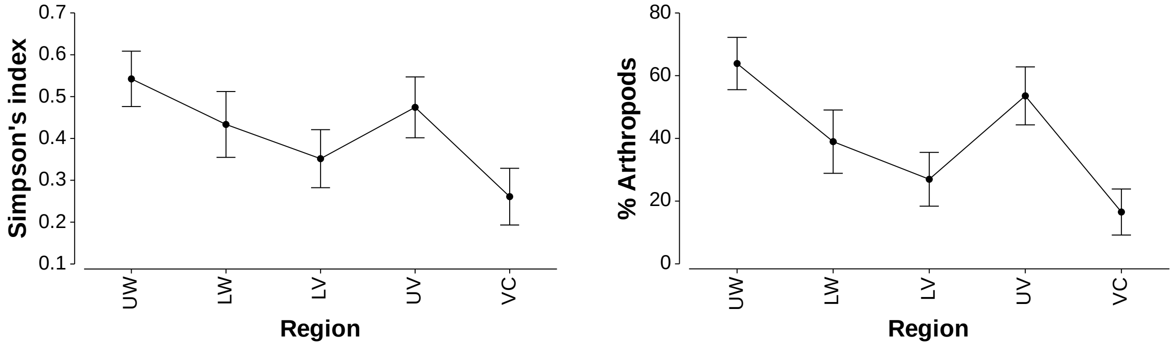
3.2. Selected metrics

Marginal DISTLM tests, which tested each metric's responsiveness separately, showed that nine of the 12 metrics explained a significant proportion ($P \leq 0.05$) in the model matrix capturing degradation differences between the Vasse Exit Channel and all other regions (Table 4). These metrics were number of species, Simpson's index, qualitative taxonomic distinctness, community dominance, SIGNAL2, richness of crustaceans and insects, and proportions of arthropods and insects. Sequential DISTLM tests, which investigated the responsiveness of metrics collectively, found that five metrics explained a significant proportion of the variation (23%) in the model matrix (Table 4). In descending order of importance, these metrics were proportion of insects, qualitative taxonomic distinctness, crustacean richness, Simpson's index, and community dominance. For the BEST test, community dominance, proportion of arthropods and SIGNAL2 was the best combination of metrics found to distinguish the Vasse Exit Channel from all other regions (Rho = 0.244, $P = 0.01$). The metrics that went on for further testing were Simpson's index, qualitative taxonomic distinctness, community dominance, crustacean richness, proportions of arthropods and insects, and SIGNAL2.

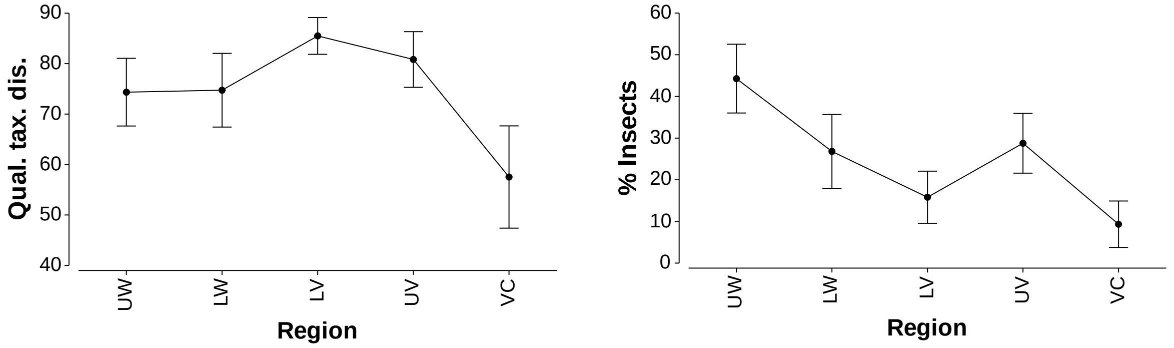
PERMANOVA test results showed that all seven metrics differed significantly between regions for main effect tests ($P \leq 0.002$, Appendix 4). Of the seven metrics tested, only crustacean richness differed solely between the Vasse Exit Channel and all other regions (Fig. 4d). Significant differences were found between multiple regions for proportions of arthropods and insects, making them unsuitable for index inclusion as there is a large amount of natural variation in their values between regions. However, the regional differences found in the remaining metrics were driven by higher values in one particular region. For Simpson's index and SIGNAL2, highest values were found in the Upper Wonnerup, which generally differed with all other regions (Fig. 4a, g). For qualitative taxonomic distinctness and community dominance, highest values were recorded in the Lower Wonnerup, which also generally differed with all other regions (Fig. 4b, c). Despite these naturally driven differences, these four metrics displayed consistently larger differences between the Vasse Exit Channel and all other regions. Given that differences in high values of metrics can be counteracted by using region-specific reference values, these four metrics along with crustacean richness were selected for index inclusion. To summarise, the metrics selected for index inclusion were Simpson's index, qualitative taxonomic distinctness, community dominance, crustacean richness and SIGNAL2.

Table 4. DISTLM results displaying the Pseudo-*F* statistic, significance value (*P*), and the variation explained separately and cumulatively by community metrics. Significant responses (*P* ≤ 0.05) are presented in bold text, with an asterisk denoting those that followed an insignificant response in the sequential tests, which can be unreliable.

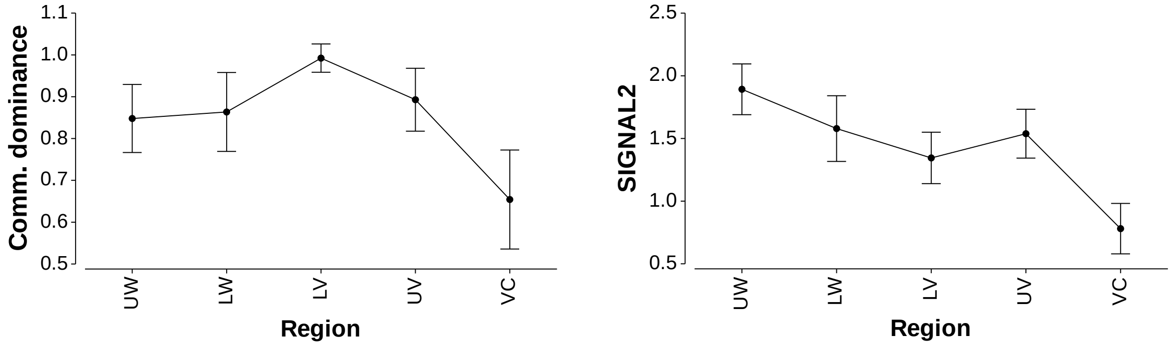
Community metric(s)	Pseudo-F statistic	*P*	Variation explained (%)	Cumulative variation (%)
DISTLM – Marginal tests				
Number of species	25.08	**0.001**	6.9	
Simpson's (Diversity) Index	25.21	**0.001**	6.9	
Qual. taxonomic distinctness	29.26	**0.001**	8.0	
Community dominance	26.01	**0.001**	7.1	
Crustacean richness	16.59	**0.001**	4.4	
Insect richness	42.13	**0.001**	11.1	
Prop. of molluscs	0.24	0.632	0.1	
Prop. of arthropods	64.32	**0.001**	16.0	
Prop. of insects	65.37	**0.001**	16.2	
Prop. of *C. capitata*	3.55	0.055	1.0	
SIGNAL2	58.05	**0.001**	14.7	
Prop. of gathering collectors	1.12	0.301	0.3	
DISTLM – Sequential test				
+ Prop. of insects	65.37	**0.001**	16.2	16.2
+ Qual. taxonomic distinctness	12.33	**0.001**	3.0	19.2
+ Crustacean richness	4.43	**0.040**	1.1	20.2
+ Simpson's (Diversity) Index	7.18	**0.011**	1.7	21.9
+ Community dominance	5.05	**0.029**	1.2	23.0
+ Prop. of arthropods	3.41	0.068	0.8	23.8
+ Insect richness	3.89	0.046*	0.9	24.7
+ Number of species	6.42	0.011*	1.4	26.1
+ Prop. of *C. capitata*	2.59	0.115	0.6	26.7
+ Prop. of molluscs	6.92	0.009*	1.5	28.2
+ Prop. of gathering collectors	0.89	0.356	0.2	28.4
+ SIGNAL2	0.49	0.481	0.1	28.5


(a) Simpson’s (Diversity) Index

(e) Proportion of arthropods


(b) Qualitative taxonomic distinctness

(f) Proportion of insects


(c) Community dominance

(g) SIGNAL2

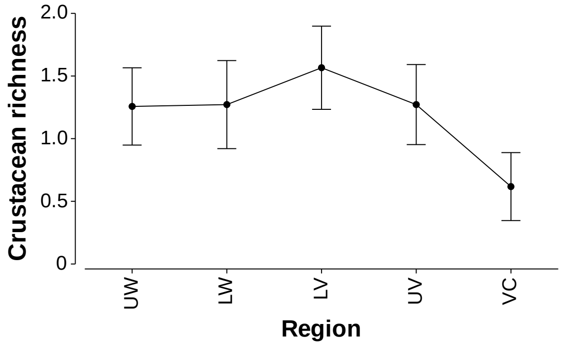

(d) Crustacean richness

Fig. 4. Mean values (± 95% confidence intervals) for (a) Simpson’s (Diversity) index, (b) qualitative taxonomic distinctness, (c) community dominance, (d) crustacean richness, (e) proportions of arthropods, (f) proportions of insects and (g) SIGNAL2 among regions (UW, Upper Wonnerup; LW, Lower Wonnerup; LV, Lower Vasse; UV, Upper Vasse; VC, Vasse Exit Channel) across all 15 sampled seasons.

3.3. Reference conditions

The reference conditions for the WBCI and EBCI, set as the 95th percentile of observed values, are presented in Table 5. Reference conditions for all metrics remained relatively stable among seasons, indicating that seasonal data of index scores are largely comparable. There are small differences in reference conditions among regions for the WBCI, with no notable differences in reference conditions between the Vasse Exit Channel and all other regions.

Table 5. Region-specific reference conditions determined for selected community metrics in each season for the WBCI and the EBCI.

WBCI						
Season	**N**	**Simpson's index $(1-\lambda')$**	**Qual. Tax. Dist. (Δ^+)**	**Community dominance (D_c)**	**Crustacean richness**	**SIGNAL2**
Upper Wonnerup (UW)						
Summer (January)	20	0.85	100.00	1.08	3.00	3.12
Autumn (March)	20	0.87	100.00	1.08	3.00	3.18
Winter (July)	12	0.87	100.00	1.08	3.05	2.93
Spring (Oct/Nov)	20	0.87	100.00	1.10	4.00	2.93
Lower Wonnerup (LW)						
Summer (January)	20	0.86	100.00	1.08	3.00	3.10
Autumn (March)	20	0.89	100.00	1.09	3.45	3.18
Winter (July)	12	0.82	100.00	1.08	4.00	3.22
Spring (Oct/Nov)	20	0.84	100.00	1.10	4.00	2.93
Upper Vasse (UV)						
Summer (January)	20	0.89	100.00	1.09	3.00	3.11
Autumn (March)	20	0.87	100.00	1.07	3.00	3.00
Winter (July)	12	0.89	100.00	1.08	4.00	2.92
Spring (Oct/Nov)	20	0.89	100.00	1.10	4.00	2.86
Lower Vasse (LV)						
Summer (January)	20	0.84	100.00	1.07	3.45	3.03
Autumn (March)	20	0.83	100.00	1.06	4.00	3.03
Winter (July)	12	0.77	100.00	1.06	4.00	2.86
Spring (Oct/Nov)	20	0.83	100.00	1.10	4.00	2.85
Vasse Exit Channel (VC)						
Summer (January)	20	0.85	100.00	1.08	3.00	3.03
Autumn (March)	20	0.89	100.00	1.07	3.00	2.93
Winter (July)	12	0.79	100.00	1.06	3.05	2.86
Spring (Oct/Nov)	20	0.82	100.00	1.10	4.00	2.83

EBCI				
Season	**N**	**Number of species (S)**	**Shannon diversity (H')**	**Community dominance (D_c)**
Summer (January)	20	10.15	1.72	1.05
Autumn (March)	20	9.00	1.31	1.07
Winter (July)	12	11.00	1.81	1.03
Spring (Oct/Nov)	20	12.05	1.80	1.07

3.4. Index trends with water quality

Scores of the WBCI exhibited seasonal trends with higher scores (grade B) measured during winter months (i.e. July and October) and lower scores during the summer months (i.e. January and March; Fig. 5, Appendix 7). Generalized linear models did confirm significant relationships between the WBCI and salinity (DE = 1.42%, $P \leq 0.001$), dissolved oxygen (DE = 1.06%, $P \leq 0.001$), pH (DE = 1.73%, $P \leq 0.001$), water level (DE = 0.76%, $P \leq 0.001$) and shallow water area (DE = 0.35%, $P \leq 0.002$), while no significant relationships were found with temperature (DE = 0.15%, $P = 0.054$), turbidity (DE = 0.06%, $P = 0.232$) or water level change (DE = 0.01%, $P = 0.72$; Fig. 6).

Relationships between the average WBCI scores and pH, dissolved oxygen and water level were generally positive, with some increase in index scores seen with a higher pH, more dissolved oxygen in the water and greater (deeper) water levels (Fig. 6). In comparison, the relationship between WBCI scores and salinity varied, with scores remaining high and relatively stable in fresher waters and declining at around 25 ppt before stabilising again at lower health scores (Fig. 6). Looking at each region separately, the Vasse Exit Channel has the poorest benthic health, with many seasons receiving low index scores (Grade E) and large variability in index scores evident within and between seasons. For the upper and lower Wonnerup, March of each year was the month generally receiving low index scores (grade D to E), while benthic health did not decline as severely in the upper and lower Vasse. Thus, the Vasse displays more stable environmental condition than the Wonnerup.

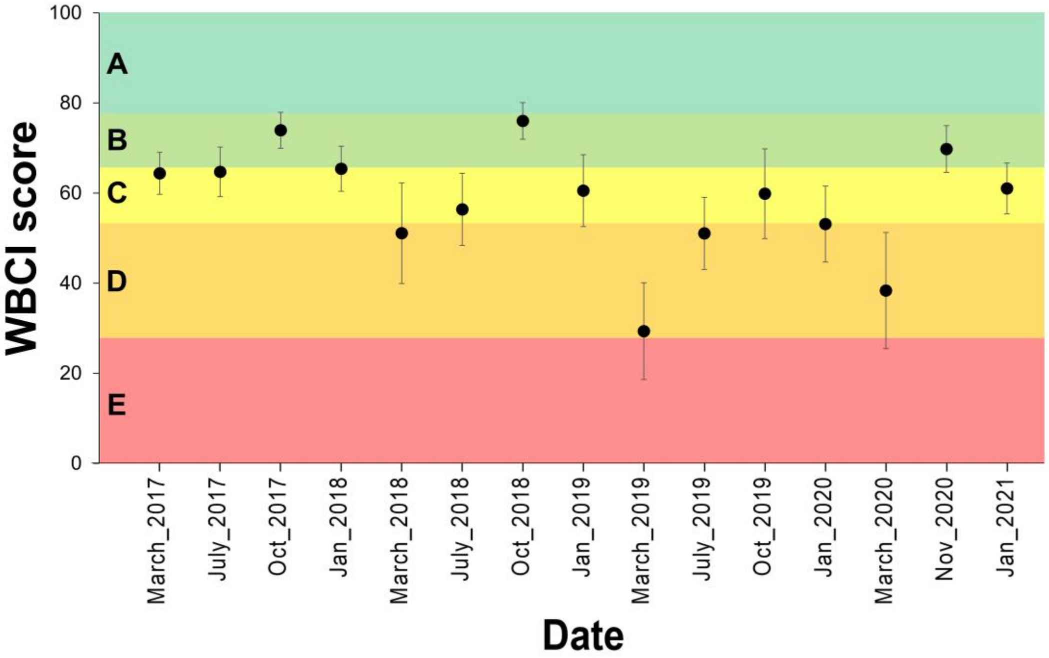


Fig. 5. Average WBCI scores (± 1 standard error) for regions above the surge barriers (excluding the Vasse Exit Channel) over the full monitoring period (March 2017-Jan 2021), with colour shading depicting the score thresholds for each health grade (A-E).

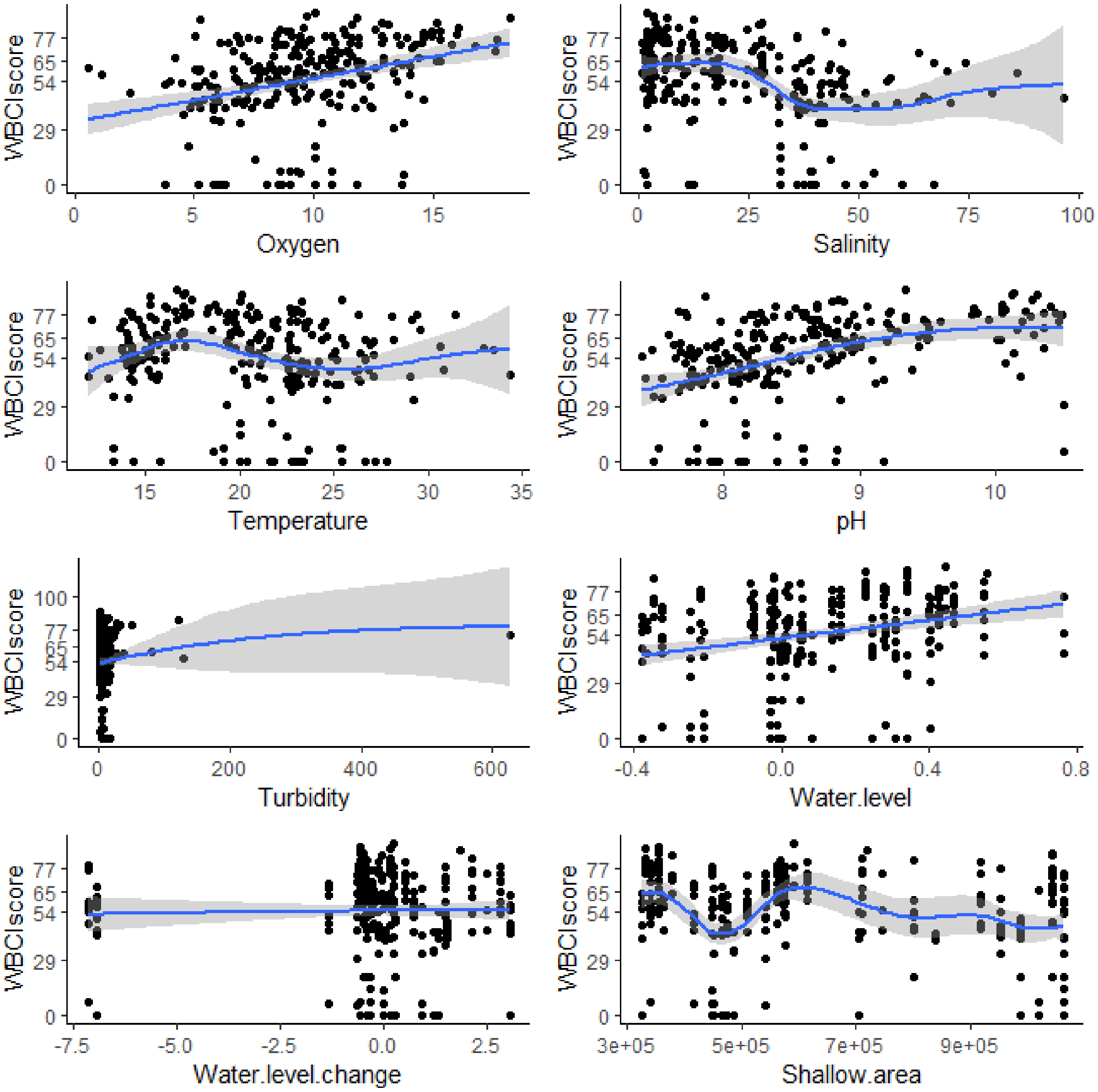
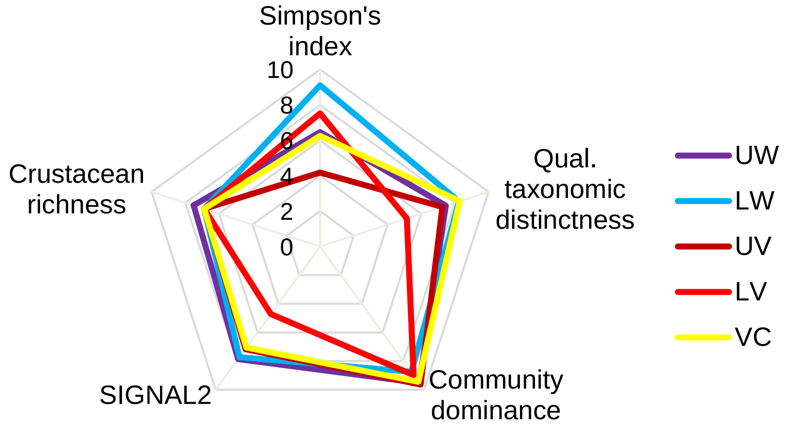


Fig. 6. Scatter plots derived from generalized linear models depicting the best-fit line for the relationships between WBCI scores and water quality parameters. *Y* axis divisions align with those set for grade thresholds, i.e. Grade A = 77-100, Grade B = 65-77, Grade C = 54-65, Grade D = 29-54, Grade E = 0-29.

In the good period investigated in more detail, namely October 2017, all regions above the surge barriers were seen to have moderate to high scores for crustacean richness, SIGNAL2 and qualitative taxonomic distinctness, while Simpson's index varied, and community dominance remained high (Fig. 7a). In the poor period, i.e. March 2019, the scores for Simpson's index and SIGNAL2 were low with crustaceans mostly absent, while qualitative taxonomic distinctness and community dominance varied largely (Fig. 7b).

(a) Good period (October 2017)



(b) Poor period (March 2019)

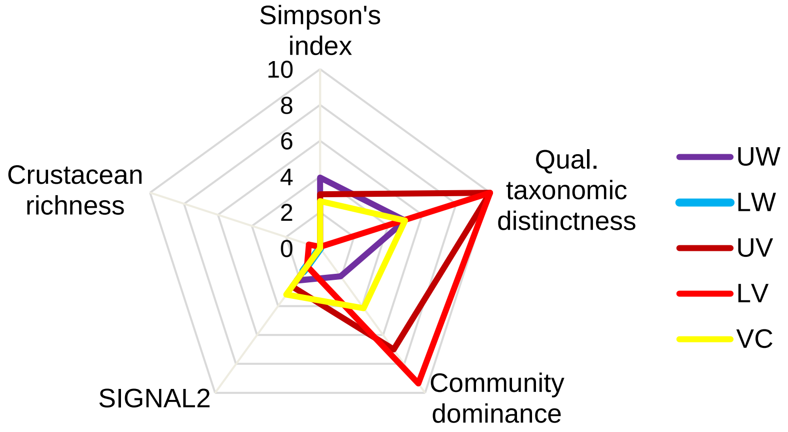


Fig. 7. Average metric scores (0-10) for each community metric measured in each region (UW, Upper Wonnerup; LW, Lower Wonnerup; UV, Upper Vasse; LV, Lower Vasse; VC, Vasse Exit Channel) in a (a) good (October 2017) and (b) poor (March 2019) period. The larger the area covered by the radar plot, the better the benthic ecological health of that region.

Scores of the EBCI in the Wonnerup Inlet were highly variable, with inconsistent patterns between seasons and years (Fig. 8, Appendix 8). For example, some of the best scores (grade B) were seen in July in 2017 and 2019, but measured lower scores (grade C) in 2018. Low EBCI scores (grade D) were seen in March 2018, which aligns with a drier period when more natural stress would be expected, but low scores (grade D) were also observed in the wet period (October 2019). Generalized linear modelling showed that the EBCI scores did not exhibit a significant relationship with any of the water quality parameters tested (DE = 0.03-0.8%, *P* = 0.17-0.77; Fig. 9). Greater metric scores of taxa richness, Shannon-Wiener diversity and community dominance were recorded in the good period (July 2019) while lower scores of all community metrics were measured in the poor period (March 2018, Fig. 10).

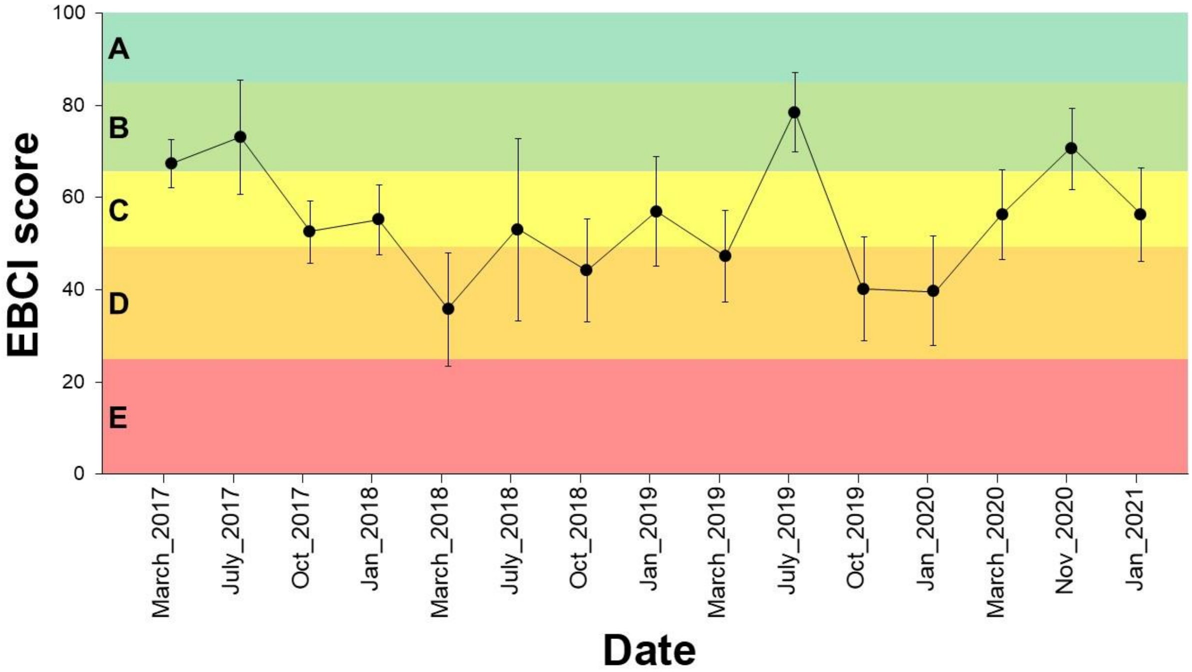


Fig. 8. Average EBCI scores (± 1 standard error) for the Wonnerup Inlet over the full monitoring period (March 2017-Jan 2021), with colour shading depicting the score thresholds for each health grade (A-E).

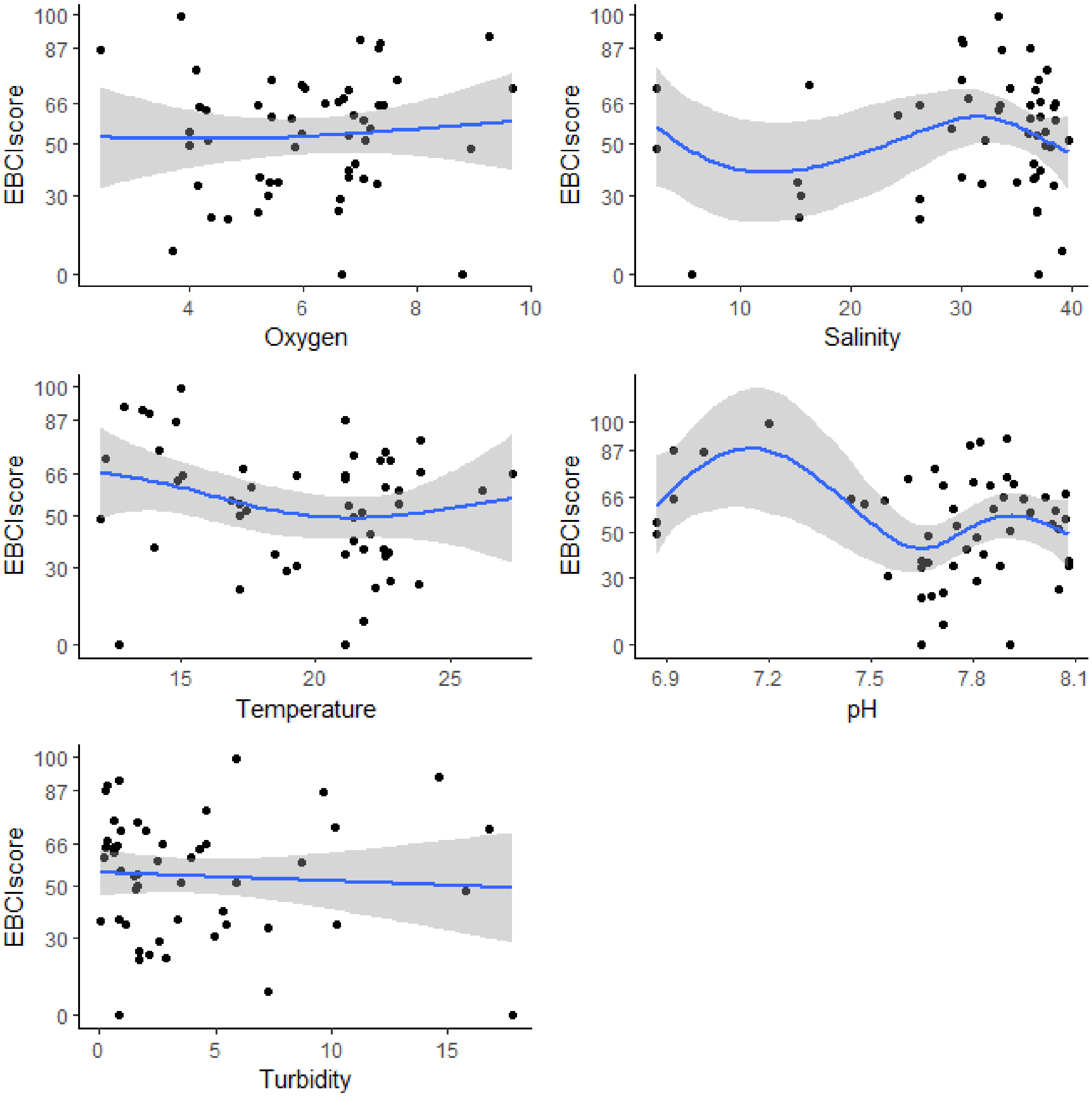


Fig. 9. Scatter plots derived from generalized linear models depicting the best-fit line for the relationships between EBCI scores and water quality parameters. Axis divisions align with those set for grade thresholds (Grade A = 87-100, Grade B = 66-87, Grade C = 50-66, Grade D = 30-50, Grade E = 0-30)

Fig. 10. Average metric scores (0-10) for each community metric measured in Wonnerup Inlet in a good (July 2019) and poor (March 2018) period. The larger the area covered by the radar plot, the better the benthic ecological health of that period.

4. Discussion

The aims of this study were to (i) determine the community metrics of the subtidal benthic macroinvertebrate community that best reflect differences in the environmental condition of the Vasse-Wonnerup, using all available macroinvertebrate data collected during previous years (2017-2021) of the IEM monitoring (Tweedley et al. 2021a, b), (ii) combine these responsive metrics into a new index to assess benthic ecological condition using season and region-specific reference conditions, (iii) evaluate how index values change through time and in response to water quality, and (iv) create a report card (Table 1 and 2) that summarises index scores as letter grades of health (A-E) to assist future monitoring of the Vasse-Wonnerup. Having a multimetric benthic index tailored for the Vasse-Wonnerup will assist in monitoring of the estuary and evaluate whether management objectives, such as maintaining healthy macroinvertebrate communities, are being met. Given the marked difference in the benthic macroinvertebrate species present above and below the surge barriers, the WBCI and EBCI were made for these broader areas, respectively. Their development and adjustment, as well as trends displayed in their scores are critically evaluated, with suggestions for implementation in a future monitoring plan.

4.1. Index trends in estuary health

After a comprehensive selection process involving 56 different metrics, the five that were found to be most appropriate and reliable and which now comprise the WBCI are Simpson's index, qualitative taxonomic distinctness, community dominance, crustacean richness and SIGNAL2. The advantage of having a multimetric index over an indicator species or single metric is that it enables multiple aspects of the community to be assessed, which increases the reliability of the index to detect environmental degradation that may affect differing aspects of the community unequally (Hering et al. 2006). Although the WBCI still requires validation to be considered fully developed, the preliminary trends it has captured are regarded as promising.

The WBCI measured the Vasse Exit Channel as the region in the Vasse-Wonnerup in the poorest condition, displaying the lowest index scores and with those scores exhibiting high variability both within and between months. This was anticipated given the range of deleterious environmental events, including hypoxia, toxic algal blooms and fish kills, that occurred in this region before, during and after the time-frame of this study (Lane et al. 1997, Beatty et al. 2018, Tweedley et al. 2019a, Cottingham et al. 2021). The high variability in metric scores (this study) and in faunal composition (Tweedley et al. 2021a) is considered a symptom of anthropogenic stress, where due to persistent disturbance/stress the fauna cannot reach a climax community (Warwick & Clarke 1993). This high variability was caused, in part, by the high proportion of azoic samples (i.e. those that contained no invertebrates, presumably as conditions were particularly poor), which results in an WBCI score of 0 and those containing only a single species. For example, the proportion of azoic samples in the Vasse Exit Channel was 2.5 to 12 times greater than those recorded in the other regions of the Vasse-

Wonnerup over the same three-year timeframe. Evidently, environmental degradation in the Vasse Exit Channel is exerting regular pressure on the macroinvertebrate community, preventing it from establishing relative to the other regions in the Vasse and Wonnerup estuaries.

According to the WBCI more broadly, the macroinvertebrate community in these regions were generally in better health during the wetter (*i.e.* October and November) than dry periods (*i.e.* January and March). Similar seasonal shifts in index scores and associated ecological health have also been observed for macroinvertebrates in the Peel-Harvey (Cronin-O'Reilly 2021) and fish in the Swan-Canning and Peel-Harvey estuaries (Hallett et al. 2019a, Tweedley et al. 2021c, 2021d). In these examples, the better ecological health in the 'wet' season was attributed to the winter rainfall and freshwater inflow reducing stressors such as hypersalinity, stagnation (high residence time) and low oxygen concentrations, which can negatively affect the fauna (Breitburg et al. 2009, Tweedley et al. 2016a, 2019b). For example, protracted marked hypersalinity in Beaufort Inlet on the south coast of Western Australia led to the invertebrate community being overwhelmingly dominated by the larvae of a single chironomid species, contributing to the loss of zoobenthic feeding taxa, e.g. fish (Krispyn et al. 2021).

Evidently, there was also a difference between the Vasse and Wonnerup regions according to the WBCI, with the Vasse tending to be less susceptible to severe declines in health than the Wonnerup. For example, benthic health in the Wonnerup sharply declined in March of 2019 and 2020, while that the Vasse exhibited only minor declines. The pronounced declines in the Wonnerup were driven by a virtual absence of crustaceans and typically low scores of all other community metrics. The community seen during these periods was dominated by the highly stress-tolerant polychaete *Capitella capitata*, accompanied by very low abundances of beetle larvae of Hydrophilidae, the gastropod *Potamopyrgus* sp., the polychaete *Simplisetia aequisetis*, and the amphipod *Perthia* sp. In comparison, the Vasse maintained high levels of community dominance and qualitative taxonomic distinctness, while similarly having low Simpson's index, SIGNAL2 and crustacea richness scores. The Vasse harboured a community similar to that of the Wonnerup, but with the addition of the amphipod *Australochiltonia subtenuis*, larvae of the non-biting midge *Procladius* sp. and Chironominae sp., the beetle *Haliplus* sp. and the ostracod *Mytilocypris ambiguosa*. The less marked decline in benthic health in the Vasse Estuary could be due to the opening of the fish gate of the Vasse surge barrier prior to and during these periods (Appendix 6). However, the exact cause-effect mechanism behind this is currently unknown and further investigation is required to say for certain whether this management action had a positive effect.

Significant relationships were detected between WBCI scores and a range of water quality parameters throughout three years (2017-2020) indicating that community health is to some extent influenced by water quality. The main outcome from these analyses is that healthier macroinvertebrate communities occur when the Vasse and Wonnerup estuaries contain water with a salinity < 25 ppt, that is relatively deep and with high levels of dissolved oxygenation and pH. Oligohaline and euhaline salinities would better suit the more wetland suite of taxa that characterise the Vasse and Wonnerup estuaries (Tweedley et al. 2019a, 2021a) and prevent the simplification of the fauna that occurs during periods of hypersalinity (Dittmann et al. 2015, Krispyn et al. 2021). Moreover, hypoxia is known to result in the death of invertebrate species, especially crustaceans, which are particularly sensitive to low oxygen concentrations (Vaquer-Sunyer & Duarte 2008, Poh et al. 2019). The positive relationship between WBCI score and pH is interesting, while this environmental variable has been shown to influence nematode communities of the Swan-Canning Estuary, its effect was not able to be separated from that of salinity (Warwick et al. 2021). In the case of the Vasse-Wonnerup it hypothesised that pH could be acting as a proxy for macroalgae and seagrass productivity in the wetter, winter periods. The higher pH values at this time may indicate that photosynthesis is converting water and carbon dioxide into algal and seagrass biomass and oxygen. This positive relationship with the WBCI could then be due to that fact that many macroinvertebrates feed on aquatic vegetation, both when alive and as detritus, and these macrophytes provide habitat for epibenthic crustaceans and insects, sheltering them from predators (Rose et al. 2019, 2020).

In comparison to the WBCI, the EBCI scores showed no relationship to water quality, suggesting that the health of the community below the surge barriers is independent of the range of environmental conditions that occurred in the region over the three-year sampling period. This mirrors similar analyses between the abundance of common taxa in Wonnerup Inlet and the same set of water quality variables (Tweedley et al. 2021a). It is notable that hypersaline conditions were not present in this region during the current study although they have occurred in the past (Lane et al. 2011, Tweedley et al. 2014a). Scores of the EBCI were also highly variable within each sampling occasion, making it difficult to determine the health of the community conclusively. As the community below the surge barrier is estuarine and dominated by stress-tolerant taxa such as *C. capitata*, *S. aequisetis* and *Arthritica semen*, the community is well adapted to dealing with variations in hydrology (Wells & Threlfall 1982, Kanandjembo et al. 2001). The similarity between the benthic macroinvertebrate communities in the Peel-Harvey and Wonnerup Inlet led the EBCI to being adjusted for the Wonnerup Inlet. The EBCI was originally developed to respond to changes in sediment quality as opposed to water quality in the Peel-Harvey (Cronin-O’Reilly 2021), which may further explain the lack of relationship between the scores and water quality.

4.2 Variation in index scores

There was considerable variation in the scores of the EBCI and, in some instances, scores of the WBCI at a region level. The standard error of these scores varied, depicting the range of average index scores a region may experience. High WBCI score variability within a region can impede detecting reliable changes in invertebrate health over time. The variability being seen is likely due to an insufficient number of samples used to characterise the invertebrate community at a regional scale, i.e. typically four samples per region. In other instances, the variability is likely a symptom of stress (Warwick & Clarke 1993), as is probably seen in health scores of the Vasse Exit Channel. For the WBCI, ideally eight grab samples are collected in each region as per the sampling undertaken in November 2021 and January 2022. In time, it may be appropriate to conduct a power analysis on WBCI data to discern if eight samples are sufficient to obtain a reliable health score for a region. For the EBCI, the variation present indicates spatial heterogeneity in health but may also be an artifact of the sampling and resulting low numbers of samples per sampling occasion. It thus is advisable to increase sample replication in this region in the future in order to characterise its health more conclusively, with the potential to further refine the metrics included, reference conditions and broader thresholds for index grades. Future work should focus on such tasks, prior to applying the EBCI index for management purposes.

There was also a notable amount of seasonal variation in WBCI scores among monitoring periods, with a significant negative relationship also evident with increasing salinity. The metric SIGNAL2 included in the WBCI is an index that is designed for freshwater riverine environments (Chessman 2003), which partly explains the restricted range of SIGNAL2 scores found in the Vasse-Wonnerup Estuary (0 to 4, when its full range is 0 to 10). The inclusion of SIGNAL2 may thus be driving the negative relationship with salinity and can aid the estuary in receiving better benthic health in the wetter, winter periods, when insects become more prevalent in the wetland (Fig. 1; Tweedley et al. 2021a). As the Vasse-Wonnerup Estuary becomes more saline in the drier months, natural declines in SIGNAL2 due to the absence of freshwater insects may thus make conditions appear poorer according to the WBCI. There are two potential solutions: (1) use the WBCI in the wet seasons (July and October) only, when SIGNAL2 is an appropriate metric to include, or (2) remove SIGNAL2 as a metric if the index is to be applied in another season. Given the marked temporal variation in environmental conditions that occur in estuaries in Mediterranean climates where rainfall is highly seasonal, it is already advised that indices in these areas are applied at the same time of year to reduce the potential for these confounding effects to influence index scores (Hallett 2010).

4.3. Index validation

The final stage for development of the WBCI is index validation, which assesses how the index performs when used on data from a sampling occasion independent of those used during its development (Engle et al. 1994, Engle & Summers 1999). It is thus necessary to carry out additional sampling, preferentially during the season(s) in which the index is likely to be applied in the future. Index validation will involve assessing whether the index can detect differences in invertebrate community health between the Vasse Exit Channel and all other regions. Thus, the recent sampling that occurred in the estuary to assist index development (*i.e.* eight sites per region; N = 40) should be sufficient to validate the index and further gauge the level of variability of index scores as discussed earlier. Given the purpose of these surveys is validation of the WBCI (and not the EBCI), it is not necessary to obtain samples form Wonnerup Inlet on this occasion. Index validation is scheduled for December 2024 when sufficient, independent monitoring data is available, and the WBCI may be subject to change based on the outcome.

4.4. Future monitoring

Following successful index validation, an appropriate monitoring plan can be devised. Monitoring plans can take many forms, with most plans devised with a particular management task in mind (Sparrow et al. 2020) . There are two overarching tasks where monitoring of the invertebrate communities could be used to inform and evaluated management actions in the Vasse-Wonnerup, which are the following:

1. Assess the health of the invertebrate communities as a food source for bird species that use the estuary prior to their arrival in summer.
2. Assess the success of management over the 'dry' season where fish kills, algal blooms and severe hypersalinity are more likely.

Both would require the invertebrate community to be sampled at different periods in the year and the task chosen will determine the timing of monitoring. For example, bird populations increase during summer months as migratory birds arrive and aim to obtain sufficient fat reserves to support their return flight (Lane et al. 2007). In this case, monitoring the health and abundance of the invertebrate communities prior to this period in spring (*i.e.* October) would be advisable. Alternatively, the management action of opening the sand bar and/or surge barriers to reduce hydrological stressors during drier periods in the estuary occurs in summer and determining the effects this management action has on the invertebrate communities would require monitoring to occur in late summer (*i.e.* March). There are advantages to both periods, with monitoring in spring aligning with the timing of macrophyte monitoring, allowing for a more holistic view of the condition of the system to be attained through multiple lenses. In contrast, monitoring in late summer may give a more realistic view of the health of the invertebrate community after the period when most deleterious events occur (hypersalinity, high temperatures, hypoxia and algal blooms), while monitoring in October may provide a more 'rose-tinted' or optimistic view, as it is generally the period when the healthiest communities are seen (Fig. 5), likely due to the input of freshwater over winter. It is important to note that the abundance of macroinvertebrates gives direct indication of the availability and amount of food present for birds (Kalejta & Hockey 1991, Gurney et al. 2017). It may thus be advisable to consider total abundances of macroinvertebrates in conjunction with the WBCI during those monitoring periods. Such a productivity metric can easily be calculated and compared to previous years.

Another important aspect to consider is the sampling resolution of a monitoring plan, which can reliably detect a signal due to anthropogenic stress while reducing any natural variation present (Hallett 2010). When viewing the benthic health of each region through time (Appendix 5), a considerable amount of variation in index scores can be present, particularly during seasons when four sites with no site-level replication have been sampled, with large amounts of variation in index scores for the Wonnerup Inlet and Vasse Exit Channel (although it is to be expected in the Vasse Exit Channel as variation can be a symptom of stress; Warwick & Clarke 1993). To reduce this variation, three alternative approaches that increase sampling resolution can be adopted:

1. Sample more sites within a region, without increasing site-level replication.
2. Increase site-level replication while maintaining the number of sites.
3. Increase both the number of sites and site-level replicates sampled.

All three approaches will likely reduce variation in index scores, with the first providing more spatial resolution, the second providing a more robust site-level assessment, and the third approach providing both outcomes. The third approach would be the costliest as it produces more samples per monitoring occasion whereas the first two approaches will have comparable costs as they provide the same number of samples per survey.

Finally, the most pivotal part of a management plan is to determine when a management response is triggered. For instance, management actions may be taken when the ecosystem is seen to be moving to an undesired state using a clear set of decision triggers that prompt a management intervention (Cook et al. 2016). For the Vasse-Wonnerup, the decision triggers established will vary depending on the aspect of health being considered, i.e. general community health, management responses or food provisioning for birds. However, they should consist of a bottom threshold that needs to be surpassed for a particular duration. For example, when assessing general health of the macroinvertebrate community, if any particular region of the estuary continuously scores a health grade of D or lower for a number of consecutive monitoring periods, management actions may be triggered. Alternatively, a threshold can be set for the number of macroinvertebrates necessary to provide resident and visiting birds with sufficient food, with management actions triggered if abundances fall below this point. Establishing such decision triggers will be imperative in maintaining the health of the estuary into the future, and thus should be discussed at length among managers and scientific advisors. Although further work is required to ensure appropriate monitoring and management of the macroinvertebrate communities in the Vasse-Wonnerup Estuary, the development and adjustment of two benthic indices to assess their health herein is an essential step taken towards achieving that goal.

5. References

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32-46

Anderson MJ, Gorley RN, Clarke KR (2008) PERMANOVA+ for PRIMER: guide to software and statistical methods. PRIMER-E, Plymouth, UK

Beatty SJ, Tweedley JR, Cottingham A, Ryan T, Williams J, Lynch K, Morgan DL (2018) Entrapment of an estuarine fish associated with a coastal surge barrier can increase the risk of mass mortalities. Ecological Engineering 122:229-240

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

Borja Á, Franco J, Pérez V (2000) A marine biotic index to establish the ecological quality of soft-bottom benthos within european estuarine and coastal environments. Marine Pollution Bulletin 40:1100-1114

Breitburg DL, Hondorp DW, Davias LA, Diaz RJ (2009) Hypoxia, Nitrogen, and Fisheries: Integrating Effects Across Local and Global Landscapes. Annual Review of Marine Science 1:329-349

Chessman BC (2003) SIGNAL 2 – A scoring system for macro-invertebrates (‘Water Bugs’) in Australian rivers. Commonwealth of Australia, Canberra, Monitoring River Heath Initiative Technical Report 31

Clarke KR, Gorley RN (2015) PRIMER v7: User Manual/Tutorial. PRIMER-E, Plymouth

Clarke KR, Somerfield PJ, Gorley RN (2008) Testing of null hypotheses in exploratory community analyses: Similarity profiles and biota-environment linkage. Journal of Experimental Marine Biology and Ecology 366:56-69

Clarke KR, Warwick RM (1998) A taxonomic distinctness index and its statistical properties. Journal of Applied Ecology 35:523-531

Cook CN, de Bie K, Keith DA, Addison PFE (2016) Decision triggers are a critical part of evidence-based conservation. Biological Conservation 195:46-51

Cottingham A, Beatty SJ, Tweedley JR (2021) Development of a report card for assessing Black Bream heath in the Vasse-Wonnerup and impact of a large fish kill. Murdoch University, Perth, Western Australia, Report for the Department of Water and Environmental Regulation

Cronin-O'Reilly S, Wells N, McCallum R, Hallett C, Tweedley J, Valesini F, Eyre B (2022) Defaunation by deoxygenation: efficacy and divergent responses of estuarine macroinvertebrates. Marine Ecology Progress Series 701:17-24

Cronin-O’Reilly S (2021) Benthic community structure, health and function of a microtidal estuary in south-western Australia. PhD, Murdoch, Perth, Western Australia

Diaz RJ, Solan M, Valente RM (2004) A review of approaches for classifying benthic habitats and evaluating habitat quality. Journal of Environmental Management 73:165-181

Dittmann S, Baring R, Baggalley S, Cantin A, Earl J, Gannon R, Keuning J, Mayo A, Navong N, Nelson M, Noble W, Ramsdale T (2015) Drought and flood effects on macrobenthic communities in the estuary of Australia's largest river system. Estuarine, Coastal and Shelf Science 165:36-51

Engle VD, Summers JK (1999) Refinement, validation, and application of a benthic condition index for Northern Gulf of Mexico estuaries. Estuaries 22:624

Engle VD, Summers JK, Gaston GR (1994) A benthic index of environmental condition of Gulf of Mexico estuaries. Estuaries 17:372-384

Gurney KEB, Clark RG, Slattery SM, Ross LCM (2017) Connecting the trophic dots: responses of an aquatic bird species to variable abundance of macroinvertebrates in northern boreal wetlands. Hydrobiologia 785:1-17

Hallett C, Valesini FJ, Yeoh D (2019a) Assessing the health of the Peel-Harvey Estuary through its fish communities. Murdoch University, Perth, Western Australia

Hallett CS (2010) The development and validation of an estuarine health index using fish community characteristics. PhD PhD, Murdoch University, Australia, Perth

Hallett CS, Hobday AJ, Tweedley JR, Thompson PA, McMahon K, Valesini FJ (2018) Observed and predicted impacts of climate change on the estuaries of south-western Australia, a Mediterranean climate region. Regional Environmental Change 18:1357-1373

Hallett CS, Trayler KM, Valesini FJ (2019b) The Fish Community Index: A practical management tool for monitoring and reporting estuarine ecological condition. Integrated Environmental Assessment and Management 15:726-738

Hallett CS, Valesini FJ, Clarke KR (2012) A method for selecting health index metrics in the absence of independent measures of ecological condition. Ecological Indicators 19:240-252

Hallett CS, Valesini FJ, Kilminster K, Wells NS, Eyre BD (2019c) A rapid protocol for assessing sediment condition in eutrophic estuaries. Environmental Science: Processes & Impacts 21:1021-1037

Hering D, Feld CK, Moog O, Ofenböck T (2006) Cook book for the development of a Multimetric Index for biological condition of aquatic ecosystems: experiences from the European AQEM and STAR projects and related initiatives. In: Furse MT, Hering D, Brabec K, Buffagni A, Sandin L, Verdonschot P (eds) The ecological status of European rivers: Evaluation and intercalibration of assessment methods. Springer, Dordrecht

Kalejta B, Hockey PAR (1991) Distribution, abundance and productivity of benthic invertebrates at the Berg River Estuary, South Africa. Estuarine, Coastal and Shelf Science 33:175-191

Kanandjembo AN, Platell ME, Potter IC (2001) The benthic macroinvertebrate community of the upper reaches of an Australian estuary that undergoes marked seasonal changes in hydrology. Hydrological Processes 15:2481-2501

Krispyn KN, Loneragan NR, Whitfield AK, Tweedley JR (2021) Salted mullet: protracted occurrence of *Mugil cephalus* under extreme hypersaline conditions. Estuarine, Coastal and Shelf Science 261:107533

Lane JA, Hardcastle KA, Tregonning RJ, Holtfreter S (1997) Management of the Vasse-Wonnerup wetland system in relation to sudden, mass fish deaths. Vasse Estuary Technical Working Group, Busselton, Australia

Lane JAK, Clarke AG, Pearson GB (2007) Waterbirds of the Vasse-Wonnerup Wetlands in 1998-2000 and some comparisons with earlier data. Western Australian Department of Environment and Conservation., Busselton, Australia

Lane JAK, Clarke AG, Winchcombe YC (2011) Depth, salinity and temperature profiling of Vasse-Wonnerup wetlands in 1998-2000. Western Australian Department of Environment and Conservation, Busselton, Australia

Legendre P, Anderson MJ (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecological Monographs 69:1-24

Ma Z, Ellison AM (2018) A unified concept of dominance applicable at both community and species scales. Ecosphere 9:e02477

McAlpine KW, Spice JF, Humphries R (1989) The environmental condition of the Vasse-Wonnerup wetland system and a discussion of management options. Western Australian Environmental Protection Authority Technical Series 31:1-35

McArdle BH, Anderson MJ (2001) Fitting multivariate models to community data: A comment on distance-based redundancy analysis. Ecology 82:290-297

Muxika I, Borja Á, Bald J (2007) Using historical data, expert judgement and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Water Framework Directive. Marine Pollution Bulletin 55:16-29

Pearson TH, Rosenberg R (1978) Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanography and Marine Biology: An Annual Review 16:229-311

Poh B, Tweedley JR, Chaplin JA, Trayler KM, Crisp JA, Loneragan NR (2019) Influence of physico-chemical and biotic factors on the distribution of a penaeid in a temperate estuary. Estuarine, Coastal and Shelf Science 218:70-85

Potter IC, Warwick RM, Hall NG, Tweedley JR (2015) The physico-chemical characteristics, biota and fisheries of estuaries. In: Craig J (ed) Freshwater Fisheries Ecology. Wiley-Blackwell

R Core Team (2019) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.

Rose TH, Tweedley JR, Warwick RM, Potter IC (2019) Zooplankton dynamics in a highly eutrophic microtidal estuary. Marine Pollution Bulletin 142:433-451

Rose TH, Tweedley JR, Warwick RM, Potter IC (2020) Influences of microtidal regime and eutrophication on estuarine zooplankton. Estuarine, Coastal and Shelf Science 238:106689

Somerfield PJ, Clarke KR, Warwick RM (2008) Simpson Index. In: Jørgensen SE, Fath BD (eds) Encyclopaedia of Ecology Vol 4. Elsevier, Oxford

Sparrow BD, Edwards W, Munroe SEM, Wardle GM, Guerin GR, Bastin J-F, Morris B, Christensen R, Phinn S, Lowe AJ (2020) Effective ecosystem monitoring requires a multi-scaled approach. Biological Reviews 95:1706-1719

Tagliapietra D, Sigovini M, Magni P (2012) Saprobity: a unified view of benthic succession models for coastal lagoons. Hydrobiologia 686:15-28

Tweedley JR, Chambers JM, Potter IC (2011) The benthic macroinvertebrate faunas of the Vasse-Wonnerup Estuary as indicators of environmental degradation. Report for the South West Catchments Council. Murdoch University, Perth, Western Australia

Tweedley JR, Cottingham A, Beatty SJ (2019a) Benthic macroinvertebrate monitoring in the Vasse-Wonnerup Wetlands: March 2017. Murdoch University, Perth, Western Australia, Report for the Department of Water and Environmental Regulation, Western Australia

Tweedley JR, Cottingham A, Strachan S, Beatty SJ (2020) Vasse-Wonnerup Integrated Monitoring 2018-19: Benthic macroinvertebrate component. Murdoch University, Perth, Western Australia, Final report to the Department of Water and Environmental Regulation, Western Australia

Tweedley JR, Cronin-O'Reilly S, Cottingham A, Beatty SJ (2021a) Vasse-Wonnerup Integrated Monitoring Review of 2017-20: Benthic macroinvertebrate component. Murdoch University, Perth, Western Australia, Report for the Department of Water and Environmental Regulation

Tweedley JR, Cronin-O'Reilly S, Cottingham A, Beatty SJ (2021b) Vasse-Wonnerup Integrated Monitoring Summary 2017-21: Benthic macroinvertebrate component. Murdoch University, Perth, Western Australia, Report for the Department of Water and Environmental Regulation

Tweedley JR, Dittmann SR, Whitfield AK, Withers K, Hoeksema SD, Potter IC (2019b) Hypersalinity: Global Distribution, Causes, and Present and Future Effects on the Biota of Estuaries and Lagoons. In: Wolanski E, Day JW, Elliott M, Ramachandran R (eds) Coasts and Estuaries. Elsevier

Tweedley JR, Hallett CS, Beatty SJ (2017) Baseline survey of the fish fauna of a highly eutrophic estuary and evidence for its colonisation by Goldfish (*Carassius auratus*). Int Aquat Res 9:259-270

Tweedley JR, Hallett CS, Warwick RM, Clarke KR, Potter IC (2016a) The hypoxia that developed in a microtidal estuary following an extreme storm produced dramatic changes in the benthos. Marine and Freshwater Research 67:327-341

Tweedley JR, Keleher J, Cottingham A, Beatty SJ, Lymbery AJ (2014a) The fish fauna of the Vasse-Wonnerup and the impact of a substantial fish kill event. Murdoch University, Perth, Australia

Tweedley JR, Krispyn KN, Hallett CS (2021c) Swan Canning Estuary condition assessment based on fish communities - 2020. Murdoch University, Perth, Western Australia, Final report to the Department of Biodiversity, Conservation and Attractions

Tweedley JR, Krispyn KN, Maus C, Cottingham A (2021d) Peel-Harvey Estuary condition assessment based on fish communities – 2020/21. Murdoch University, Perth, Western Australia, Final report to the Peel Harvey Catchment Council

Tweedley JR, Warwick RM, Clarke KR, Potter IC (2014b) Family-level AMBI is valid for use in the north-eastern Atlantic but not for assessing the health of microtidal Australian estuaries. Estuarine, Coastal and Shelf Science 141:85-96

Tweedley JR, Warwick RM, Potter IC (2015) Can biotic indicators distinguish between natural and anthropogenic environmental stress in estuaries? Journal of Sea Research 102:10-21

Tweedley JR, Warwick RM, Potter IC (2016b) The contrasting ecology of temperate macrotidal and microtidal estuaries. Oceanography and Marine Biology: An Annual Review 54:73-172

Tweedley JR, Warwick RM, Valesini FJ, Platell ME, Potter IC (2012) The use of benthic macroinvertebrates to establish a benchmark for evaluating the environmental quality of microtidal, temperate southern hemisphere estuaries. Marine Pollution Bulletin 64:1210-1221

Vaquer-Sunyer R, Duarte CM (2008) Thresholds of hypoxia for marine biodiversity. Proceedings of the National Academy of Sciences 105:15452-15457

Warwick RM, Clarke KR (1993) Increased variability as a symptom of stress in marine communities. Journal of experimental Marine Biology and Ecology 172:215-226

Warwick RM, Hourston M, Tweedley JR, Potter IC (2021) The contrasting ecology of free-living nematodes in macrotidal and microtidal estuaries. Estuaries and Coasts 44:214-228

Warwick RM, Tweedley JR, Potter IC (2018) Microtidal estuaries warrant special management measures that recognise their critical vulnerability to pollution and climate change. Marine Pollution Bulletin 135:41-46

Wells FE, Threlfall TJ (1982) Reproductive strategies of *Hydrococcus brazeri* (Tenison Woods, 1876) and *Arthritica semen* (Menke, 1843) in Peel Inlet Western Australia. Journal of the Malacological Society of Australia 5:157-166

Wetland Research & Management (2007) Ecological character description: Vasse-Wonnerup wetlands Ramsar site south-west Western Australia. Wetland Research & Management, Perth, Australia

Wildsmith MD, Rose TH, Potter IC, Warwick RM, Clarke KR (2011) Benthic macroinvertebrates as indicators of environmental deterioration in a large microtidal estuary. Marine Pollution Bulletin 62:525-538

Wildsmith MD, Rose TH, Potter IC, Warwick RM, Clarke KR, Valesini FJ (2009) Changes in the benthic macroinvertebrate fauna of a large microtidal estuary following extreme modifications aimed at reducing eutrophication. Marine Pollution Bulletin 58:1250-1262

Williams J, Cottingham A, Denham A, Hall NG, Potter IC (2020) Relationship between spawning and egg and larval stages of a unique estuarine-resident species and environmental variables and prey. Estuarine, Coastal and Shelf Science 246:107039

6. Appendices

Appendix 1. Full list of the 56 metrics spanning aspects of the community's richness/diversity, composition, tolerance, and function investigated for potential index inclusion before metric selection, with the results of the first quality check for extreme outliers and restricted ranges given (pass/fail).

Community metric	Response to stress	Metric description	Quality check
Richness/Diversity			
Number of species (*S*)	−	Number of taxa found at each site	Pass
Total abundance (N)	−/+	Number of individuals at each site	Fail
Margalef's richness (d)	−	Richness index	Pass
Simpson's Index ($1-\lambda'$)	−	The spread of taxon abundances (evenness) among species	Pass
Shannon-Wiener diversity (*H'*)	−	A richness index that assesses the evenness among species numbers	Pass
Pielou's Evenness (*J'*)	−	The spread of taxon abundances (evenness) among species	Pass
Expected number of species (ES[50])	−	The number of species expected if 50 individuals are present in a sample	Pass
Quantitative taxonomic distinctness (Δ^*)	−	The taxonomic spread of species found at each site based on their abundances	Pass
Qualitative taxonomic distinctness (Δ^+)	−	The taxonomic spread of species found at each site based on their presence or absence	Pass
Variation in taxonomic distinctness (λ^+)	+	A measure of how much the taxonomic spread of species varies among samples	Fail
Community dominance (D_c)	+	The level of species dominance present	Pass
Annelid richness	+	Number of annelid richness present	Fail
Crustacean richness	−	Number of crustacean taxa present	Pass
Insect richness	−	Number of insect taxa present	Pass
Mollusc richness	−	Number of mollusc taxa present	Fail
Composition			
Proportion of annelids	+	Percentage contribution of annelids to the community	Pass
Proportion of arthropods	−/+	Percentage contribution of arthropods to the community	Pass
Proportion of molluscs	−	Percentage contribution of molluscs to the community	Pass
Proportion of crustaceans	−	Percentage contribution of crustaceans to the community	Fail
Proportion of insects	−	Percentage contribution of insects to the community	Pass
Annelid abundance	+	Number of annelids at each site	Fail
Arthropod abundance	−/+	Number of arthropods at each site	Fail
Mollusc abundance	−	Number of molluscs at each site	Fail
Crustacean abundance	−	Numbers of crustaceans at each site	Fail
Insect abundance	−	Numbers of insects at each site	Fail

Appendix 1. Cont.

Community metric	Response to stress	Metric description	Quality check
Tolerance			
Plecoptera, Ephemeroptera, Trichoptera (PET) richness	−	Number of sensitive PET taxa present	Fail
Plecoptera, Ephemeroptera, Trichoptera (PET) abundance	−	Number of sensitive PET individuals present	Fail
Proportion of amphipods	−	Percentage contribution of amphipods that are thought to be environmentally sensitive	Fail
Proportion of *Capitella capitata*	+	Percentage contribution of the pollution tolerant annelid *C. Capitata* to the community	Pass
Proportion of tolerant annelids	+	Percentage contribution of annelids belonging to the Oligochaeta, Cirratulidae, Capitellidae, *Prionospio*, *Pseudopolydora* and *Polydora.*	Pass
Amphipod abundance	−	Number of amphipods at a site	Fail
Capitella capitata abundance	+	Number of *C. capitata* at a site	Fail
Tolerant annelids abundance	+	Number of annelids belonging to the Oligochaeta, Cirratulidae, Capitellidae, *Prionospio*, *Pseudopolydora* and *Polydora.*	Fail
SIGNAL2 score	−	A tolerance-based index based on the recognised sensitivity and tolerance of various insect groups	Pass
Function			
Proportion of gathering collectors	+	Percentage contribution of taxa that feed on deposited materials on the sediment	Pass
Proportion of grazers	−	Percentage contribution of taxa that feed on living micro- and macroalgae	Fai
Proportion of filtering collectors	−	Percentage contribution of taxa that feed on materials from water column	Fail
Proportion of predators	−	Percentage contribution of taxa that consume other taxa	Fail
Proportion of scavengers	+	Percentage contrition of taxa that eat dead or decaying plant or animal material	Fail
Proportion of shredders	−	Percentage contribution of taxa that break down larger pieces of organic material	Fail
Proportion of gathering/filtering collectors	+	Percentage contribution of taxa that deposit and filter feed (i.e. are more generalist)	Fail
Proportion of gathering collectors/predators	+	Percentage contribution of taxa that deposit feed and consume other taxa (i.e. are more generalist)	Fail
Proportion of predators/scavengers	+	Percentage contribution of taxa that consume other taxa and eat decaying or dead plants or animals (i.e. are more generalist)	Fail
Proportion of predators/shredders	+	Percentage contribution of taxa that deposit feed and consume other taxa (i.e. are more generalist)	Fail

Appendix 1. Cont.

Community metric	Response to stress	Metric description	Quality check
Function			
Proportion of generalists	+	Number of taxa that are recognised to use three feeding modes or more (i.e. are generalists)	Fail
Abundance of gathering collectors	+	Number of taxa that feed on deposited materials on the sediment	Fail
Abundance of grazers	−	Number of taxa that feed on living micro- and macroalgae	Fail
Abundance of filtering collectors	−	Number of taxa that feed on materials from water column	Fail
Abundance of predators	−	Number of taxa that consume other taxa	Fail
Abundance of scavengers	+	Number of taxa that eat dead or decaying plant or animal material	Fail
Abundance of shredders	−	Number of taxa that break down larger pieces of organic material	Fail
Abundance of gathering/filtering collectors	+	Number of taxa that deposit and filter feed (i.e. are more generalist)	Fail
Abundance of gathering collectors/predators	+	Number of taxa that deposit feed and consume other taxa (i.e. are more generalist)	Fail
Abundance of predators/scavengers	+	Number of taxa that consume other taxa and eat decaying or dead plants or animals (i.e. are more generalist)	Fail
Abundance of predators/shredders	+	Number of taxa that deposit feed and consume other taxa (i.e. are more generalist)	Fail
Abundance of generalists	+	Number of taxa that are recognised to use three feeding modes or more (i.e. are generalists)	Fail

Appendix 2. Pearson correlations between community metrics that made it past the first quality check of their value distributions, with only one metric retained from metric pairs that had correlations greater than 0.9 (shaded grey). Metrics included were number of species (*S*), Margalef's richness (d), Pielou's eveness (*J'*), expected number of species (ES[50]), Shannon-Wiener diversity (*H'*), Simpson's (div.) Index (1−λ'), quantitative taxonomic distinctness (Δ*), qualitative taxonomic distinctness (Δ^+), community dominance (D^c), proportions of annelids (%Ann.), molluscs (%Moll.), arthropods (%Arth.), insects (%In.), tolerant annelids (%Tol. Ann.), *Capitella capitata* (%*C. cap*.), SIGNAL, crustacean (C. rich.), insect richness (I. rich), and the proportion of gathering collectors (%GC).

	S	**d**	***J'***	**ES[50]**	***H'***	**1−λ'**	**Δ***	**Δ^+**	**D_c**	**%Ann.**	**%Moll.**	**%Arth.**	**%In.**	**%Tol. Ann.**	**%*C. cap.***	**SIGNAL 2**	**C. rich**	**I. rich**
d	0.85																	
J'	0.42	0.73																
ES[50]	0.91	0.92	0.62															
H'	0.76	0.91	0.85	0.91														
1−λ'	0.56	0.83	0.96	0.73	0.92													
Δ*	0.37	0.36	0.40	0.34	0.33	0.32												
Δ^+	0.37	0.37	0.43	0.36	0.35	0.35	0.96											
D_c	0.57	0.42	0.32	0.54	0.42	0.27	0.81	0.84										
%Ann.	-0.29	-0.44	-0.43	-0.39	-0.47	-0.47	0.19	0.15	0.05									
%Moll.	-0.05	-0.02	0.03	-0.01	-0.01	0.01	0.08	0.03	-0.04	-0.36								
%Arth.	0.48	0.60	0.57	0.57	0.63	0.61	-0.03	0.06	0.21	-0.68	-0.28							
%In.	0.25	0.38	0.47	0.36	0.44	0.45	-0.05	0.06	0.18	-0.50	-0.22	0.78						
%Tol.Ann.	-0.28	-0.44	-0.45	-0.39	-0.48	-0.49	0.20	0.15	0.06	0.98	-0.35	-0.66	-0.49					
%*C. cap.*	-0.30	-0.45	-0.45	-0.41	-0.49	-0.49	0.19	0.15	0.05	0.97	-0.35	-0.66	-0.49	0.99				
SIGNAL2	0.51	0.55	0.49	0.56	0.56	0.51	0.20	0.24	0.38	-0.53	0.00	0.70	0.75	-0.51	-0.51			
C. rich	0.76	0.65	0.30	0.66	0.57	0.43	0.22	0.20	0.35	-0.35	-0.04	0.46	0.07	-0.33	-0.34	0.31		
I. rich	0.82	0.74	0.41	0.79	0.69	0.52	0.20	0.19	0.43	-0.40	-0.14	0.62	0.52	-0.38	-0.39	0.67	0.50	
%GC	-0.17	-0.33	-0.38	-0.30	-0.39	-0.41	0.27	0.24	0.18	0.79	-0.26	-0.55	-0.42	0.80	0.81	-0.39	-0.24	-0.29

Appendix 3. Rapid Assessment overview of sediment condition, showing the average scores of colour, texture, and odour measured in each region and season, with the number of sites classified into good, fair and poor sediment conditions based on these scores provided. The proportion of sites assigned to sediment condition classes are also presented. Higher colour, texture and odour scores indicate better sediment condition. Note that the RAP classification was developed for the Peel-Harvey Estuary (Hallett et al. 2019c) and is not validated for the Vasse-Wonnerup, thus these condition classes act as provisional classifications only. Regions: UW, Upper Wonnerup; LW, Lower Wonnerup; LV, Lower Vasse; UV, Upper Vasse; VC, Vasse Exit Channel.

	UW	LW	UV	LV	VC	%	IT	%
November 2020								
Colour	0.4	0.8	1.9	1.6	2.0		3.5	
Texture	3.3	3.3	3.8	3.3	1.8		3.4	
Odour	3.6	2.5	3.9	2.6	1.6		2.9	
Good	1	2	5	3	2	33	4	50
Fair	0	0	0	0	1	3	0	0
Poor	7	6	3	5	5	65	4	50
January 2021								
Colour	0.3	0.0	1.0	0.0	0.3		3.3	
Texture	2.5	1.5	2.0	2.8	2.0		4.0	
Odour	3.5	3.8	1.4	3.4	1.0		4.6	
Good	1	0	2	0	1	10	8	100
Fair	0	0	0	0	0	0	0	0
Poor	7	8	6	8	7	90	0	0

Appendix 4. T-values for PERMANOVA pairwise comparisons between regions for community metrics selected following DISTLM and BEST tests. Higher t-statistic values indicate greater regional differences in metric values. Significance levels for the main test of region are presented beside the community metric name, with significant pairwise comparisons highlighted grey, denoted in bold text, with an asterisk (*) to signify the level of significant difference: * for pairwise comparisons with *P* ≤ 0.05, ** for *P* ≤ 0.01, and *** for *P* ≤ 0.001. Regions: UW, Upper Wonnerup; LW, Lower Wonnerup; LV, Lower Vasse; UV, Upper Vasse; VC, Vasse Exit Channel.

(a) Simpson's (diversity) Index *P* = 0.001

	UW	LW	UV	LV
LW	**2.05***			
UV	1.07	1.05		
LV	**3.37****	1.15	2.33*	
VC	**5.51*****	**3.29****	**4.54*****	**2.28***

(b) Qual. taxonomic distinctness *P* = 0.001

	UW	LW	UV	LV
LW	0.08			
UV	1.49	1.33		
LV	**2.92****	**2.63****	1.41	
VC	**2.76****	**2.75***	**4.03*****	**5.18*****

(c) Community dominance *P* = 0.001

	UW	LW	UV	LV
LW	0.25			
UV	0.81	0.49		
LV	**3.27*****	**2.56***	**2.41***	
VC	**2.69***	**2.76***	**3.40****	**5.48*****

(d) Crustacea richness *P* = 0.002

	UW	LW	UV	LV
LW	0.06			
UV	0.07	N		
LV	1.36	1.21	1.27	
VC	**3.12****	**2.94****	**3.12****	**4.42*****

(e) Prop. of arthropods *P* = 0.001

	UW	LW	UV	LV
LW	**3.25****			
UV	0.74	**2.74****		
LV	**4.76*****	1.24	**4.33*****	
VC	**8.23*****	**4.87*****	**7.98*****	**3.92*****

(f) Prop. of insects *P* = 0.001

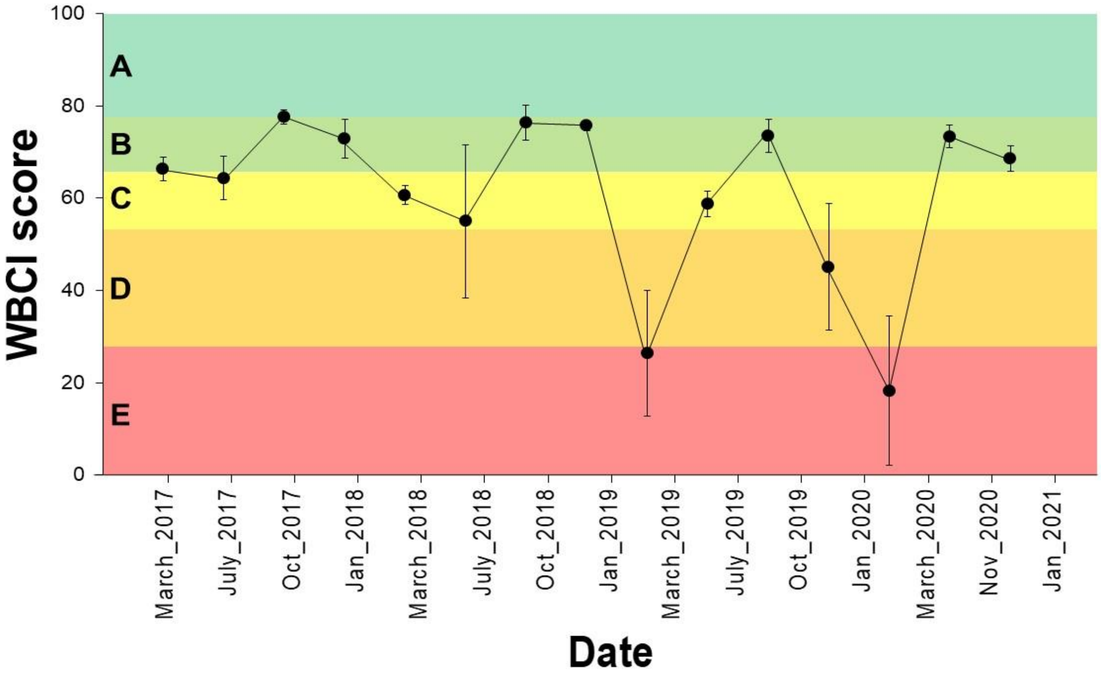
	UW	LW	UV	LV
LW	**3.20*****			
UV	1.56	1.93		
LV	**5.16*****	1.65	**3.94****	
VC	**9.05*****	**5.38*****	**8.16*****	**4.02*****

(g) SIGNAL2 *P* = 0.001

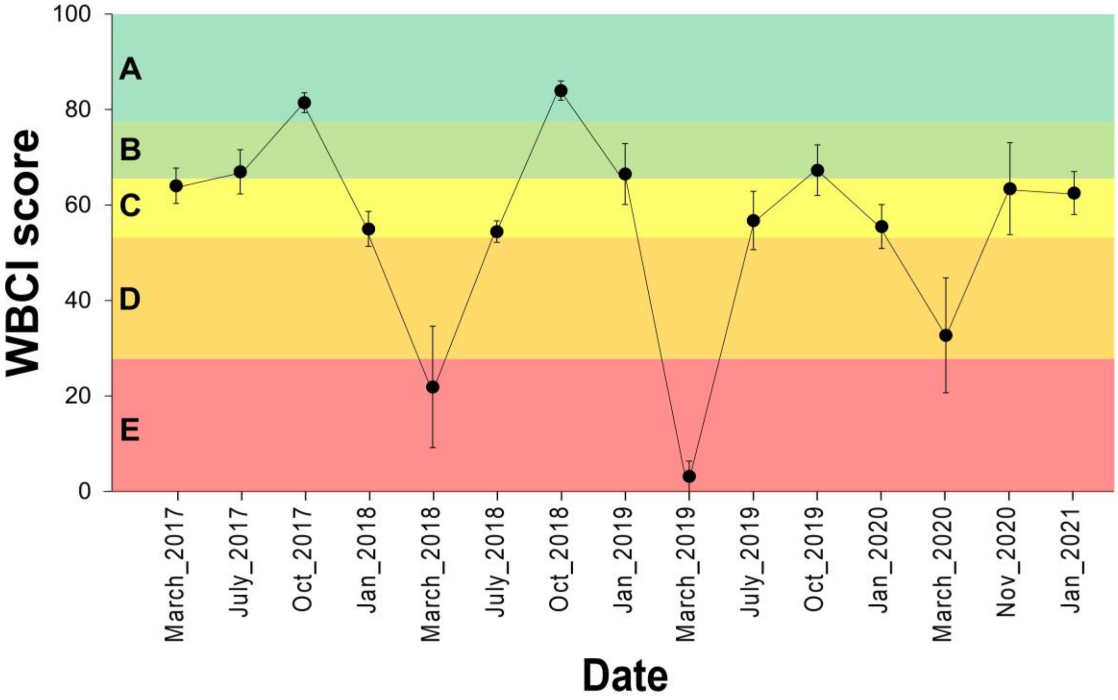
	UW	LW	UV	LV
LW	**2.27***			
UV	**2.13***	0.47		
LV	**3.38****	0.59	1.248	
VC	**7.60*****	**4.71*****	**5.77*****	**4.70*****

Appendix 5. Average WBCI scores (± 1 standard error) for each region (a-e) above the surge barriers over the full monitoring period (March 2017-Jan 2021), with colour shading depicting the score thresholds for each health grade (A-E). The total number of samples for most periods was four per region, except for March 2017, November 2020 and January 2021, when eight samples were collected in each region.

(a) Upper Wonnerup

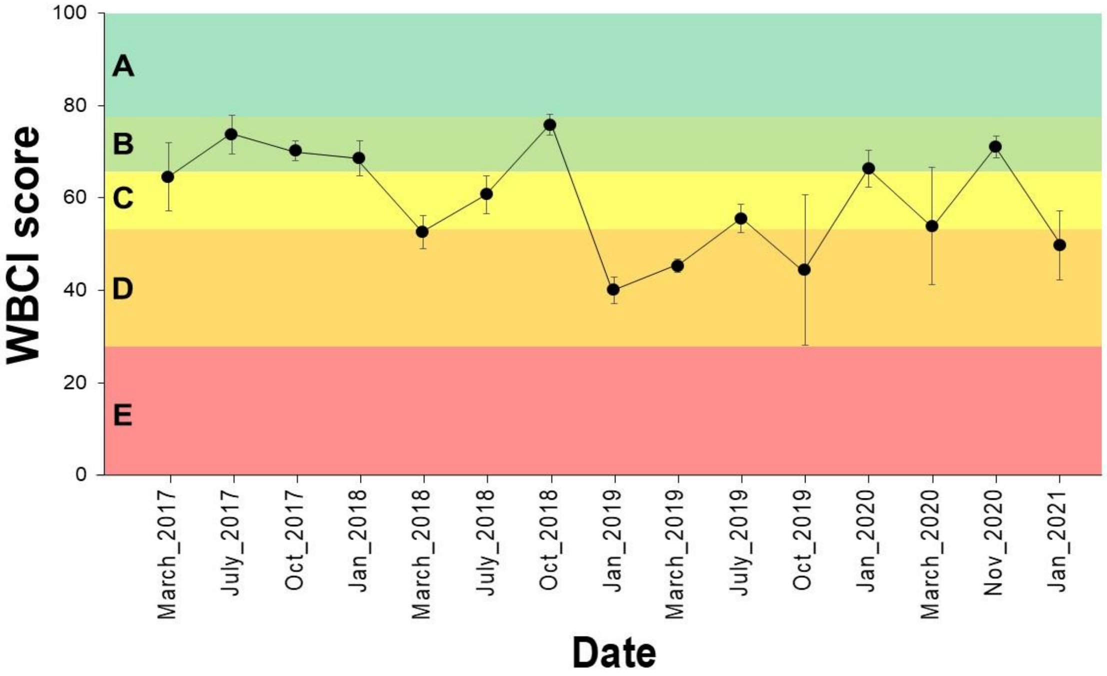


(b) Lower Wonnerup

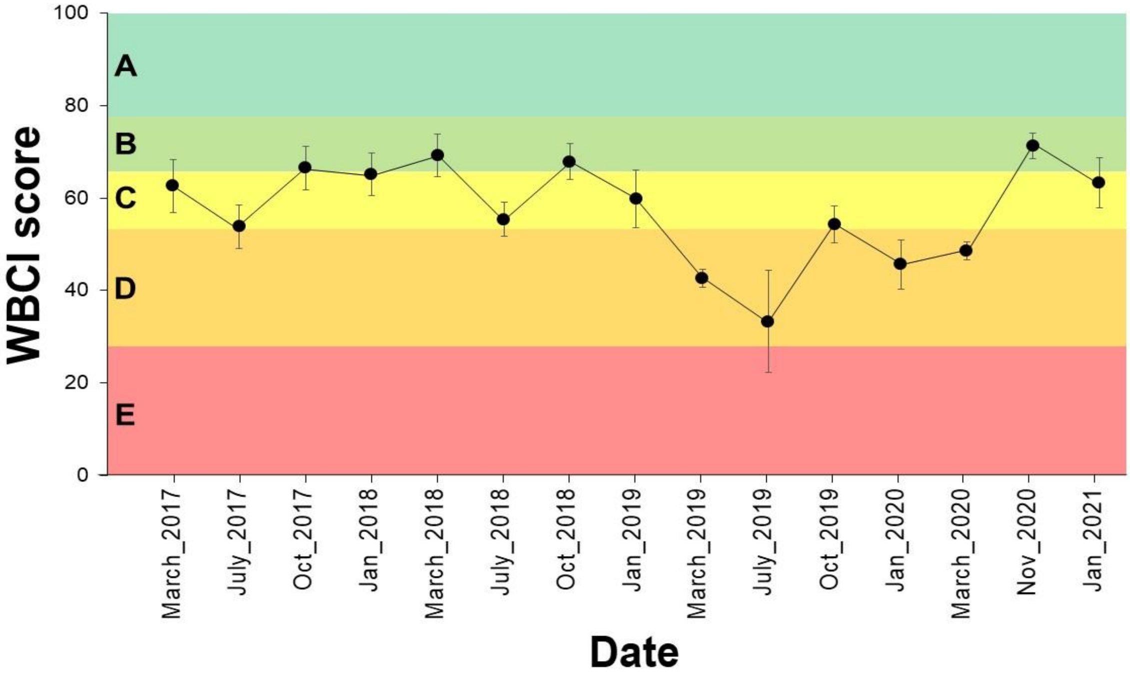


Appendix 5. cont.

(c) Upper Vasse

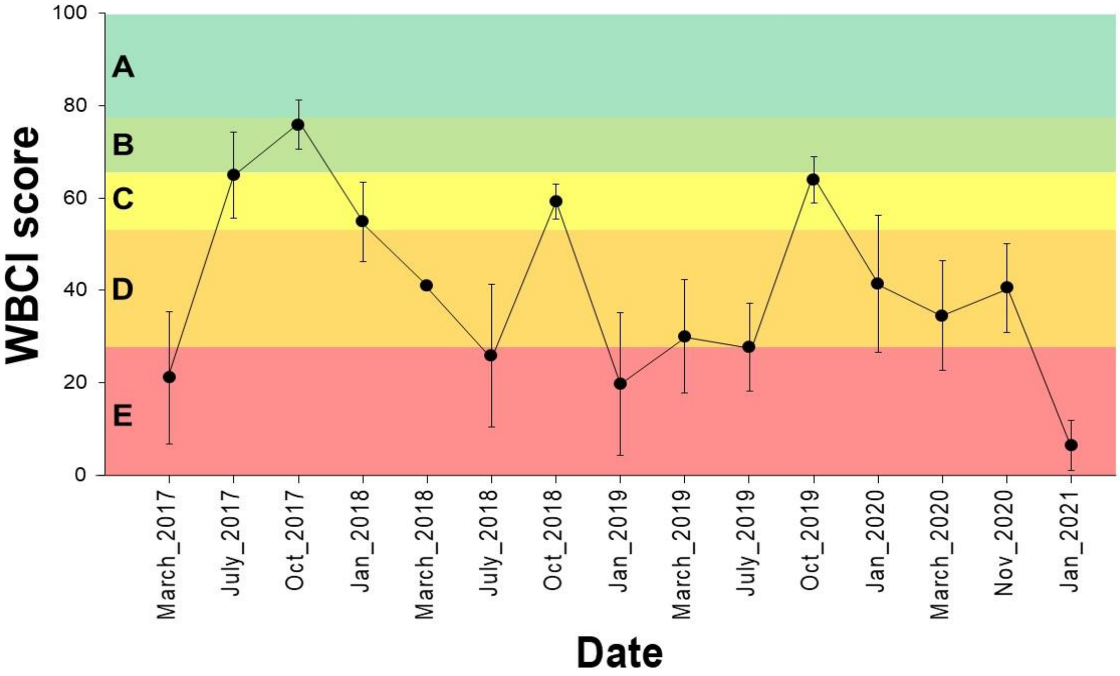


(d) Lower Vasse



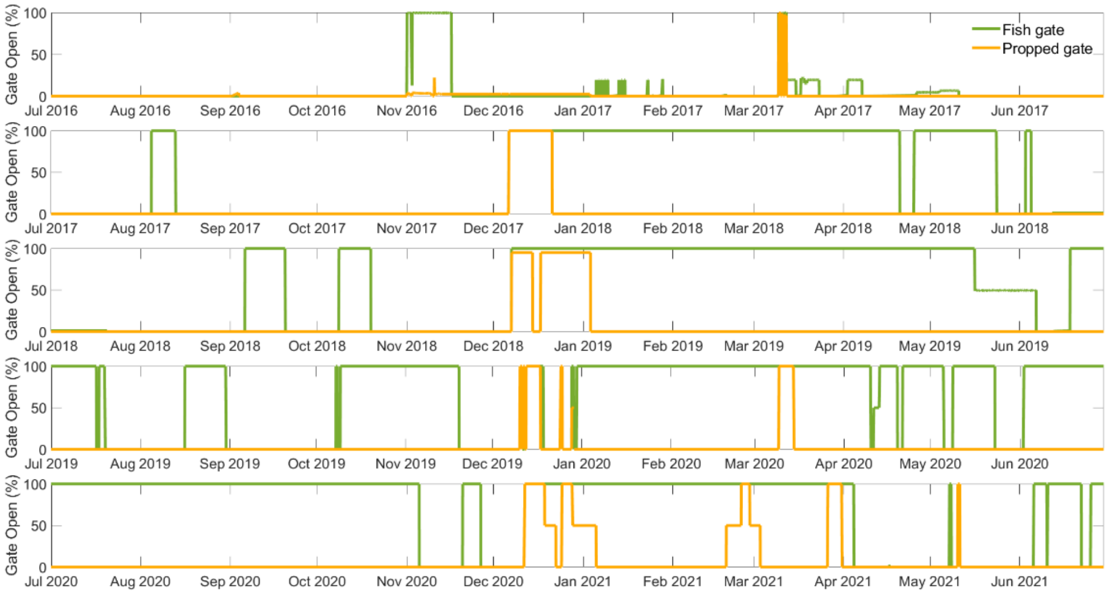
Appendix 5. cont.

(e) Vasse Exit Channel

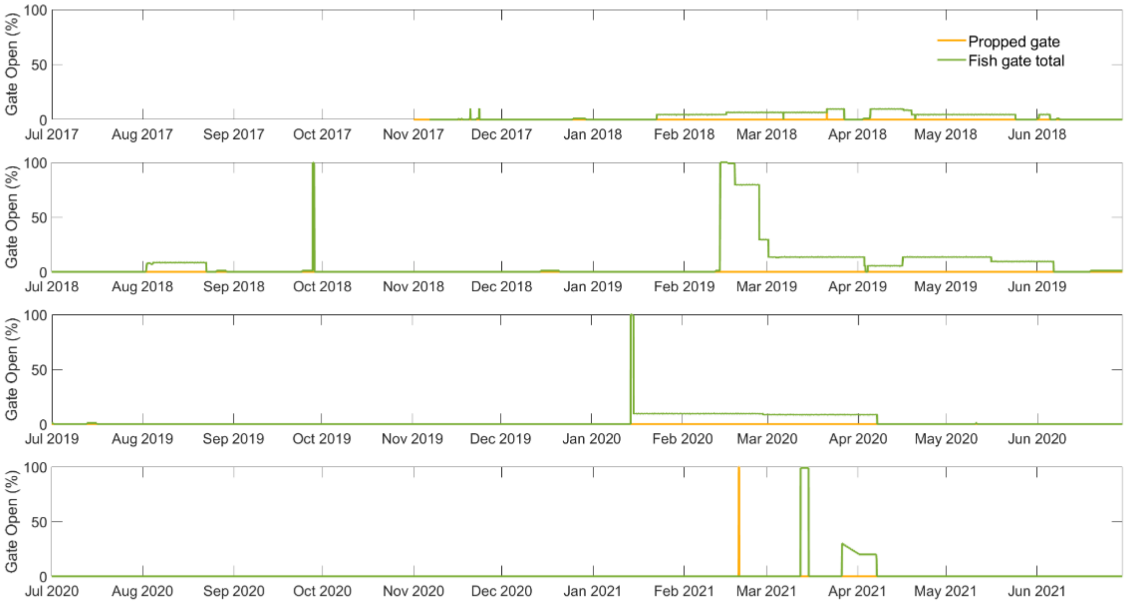


Appendix 6. Opening status of the (a) Vasse and (b) Wonnerup storm surge barriers in the estuary over the course of monitoring period (2017-2021). Source: Department of Water and Environmental Regulation.

(a) Vasse surge barrier



(b) Wonnerup surge barrier



Appendix 7. Average scores of the WBCI, their standard error (SE), and associated health grade (A-E) over the 5-year monitoring period for all regions (*excluding the Vasse Exit Channel) and each region separately. Each season contained four samples per region except for November 2020 and January 2021, where 8 samples were taken per region. Regions: UW, Upper Wonnerup; LW, Lower Wonnerup; LV, Lower Vasse; UV, Upper Vasse; VC, Vasse Exit Channel**.**

	All regions*			UW			LW			UV			LV			VC		
	WBCI	SE	Grade	WBCI	SE	Grade	WBCI	SE	Grade	WBCI	SE	Grade	WBCI	SE	Grade	WBCI	SE	Grade
Mar-17	64.4	4.7	C	66.3	2.6	B	64.0	3.7	C	64.6	7.3	C	62.6	5.7	C	21.1	14.3	E
Jul-17	64.7	5.5	C	64.4	4.7	C	67.0	4.6	B	73.6	4.2	B	53.9	4.7	D	65.0	9.4	C
Oct-17	73.9	4.0	B	77.6	1.5	A	81.4	2.1	A	70.2	2.2	B	66.5	4.7	B	75.9	5.3	B
Jan-18	65.4	5.0	C	72.9	4.2	B	55.0	3.7	C	68.6	3.8	B	65.2	4.6	C	54.9	8.5	C
Mar-18	51.1	11.2	D	60.7	2.0	C	21.9	12.7	E	52.6	3.6	D	69.2	4.6	B	41.0	0.9	D
Jul-18	56.4	8.0	C	55.0	16.7	C	54.4	2.2	C	60.7	4.1	C	55.4	3.6	C	25.8	15.5	E
Oct-18	76.0	4.1	B	76.4	3.8	B	84.0	2.0	A	75.8	2.3	B	67.9	3.9	B	59.3	3.8	C
Jan-19	60.5	8.0	C	75.7	1.0	B	66.5	6.4	B	40.1	2.9	D	59.9	6.3	C	19.7	15.4	E
Mar-19	29.3	10.7	D	26.3	13.7	E	3.2	3.2	E	45.3	1.4	D	42.6	2.0	D	30.0	12.3	D
Jul-19	51.1	8.0	D	58.7	2.7	C	56.8	6.1	C	55.5	3.1	C	33.3	11.1	D	27.6	9.6	E
Oct-19	59.9	10.0	C	73.6	3.6	B	67.3	5.3	B	44.4	16.4	D	54.2	4.0	D	63.9	5.0	C
Jan-20	53.1	8.4	D	45.1	13.7	D	55.5	4.6	C	66.3	4.0	B	45.6	5.4	D	41.4	14.9	D
Mar-20	38.4	12.9	D	18.2	16.2	E	32.7	12.0	D	53.9	12.6	D	48.6	2.0	D	34.5	11.8	D
Nov-20	69.8	5.2	B	73.4	2.5	B	63.4	9.6	C	71.0	2.3	B	71.3	2.7	B	40.5	9.6	D
Jan-21	61.0	5.6	C	68.6	2.8	B	62.5	4.5	C	49.7	7.4	D	63.3	5.5	C	6.3	5.5	E

Appendix 8. Average scores of the EBCI, their standard error (SE), and associated health grade (A-E) over the 5-year monitoring period for the Wonnerup Inlet. Each season contained four samples per region except for November 2020 and January 2021, where 8 samples were taken.

	EBCI	SE	Grade
Mar-17	67.3	5.2	B
Jul-17	73.1	12.4	B
Oct-17	52.6	6.8	C
Jan-18	55.2	7.5	C
Mar-18	35.7	12.2	D
Jul-18	53.0	19.8	C
Oct-18	44.1	11.1	D
Jan-19	57.0	11.9	C
Mar-19	47.3	9.8	D
Jul-19	78.5	8.6	B
Oct-19	40.2	11.3	D
Jan-20	39.8	11.9	D
Mar-20	56.3	9.7	C
Nov-20	70.5	8.8	B
Jan-21	56.4	10.1	C