



**HAL**  
open science

## A Bayesian framework to objectively combine metrics when developing stressor specific multimetric indicator

Hilaire Drouineau, Jérémy Lobry, C. Delpech, M. Bouchoucha, S. Mahévas, A. Courrat, Stéphanie Pasquaud, Mario Lepage

### ► To cite this version:

Hilaire Drouineau, Jérémy Lobry, C. Delpech, M. Bouchoucha, S. Mahévas, et al.. A Bayesian framework to objectively combine metrics when developing stressor specific multimetric indicator. *Ecological Indicators*, 2012, 13 (1), p. 314 - p. 321. 10.1016/j.ecolind.2011.06.029 . hal-00654282

**HAL Id: hal-00654282**

**<https://hal.science/hal-00654282>**

Submitted on 21 Dec 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **A Bayesian framework to objectively combine metrics when developing stressor specific**  
2 **multimetric indicator**

3

4 Drouineau, H<sup>\*1</sup>; Lobry J<sup>1</sup>, Delpech, C<sup>1</sup>; Bouchoucha, M<sup>2</sup>; Mahévas, S<sup>3</sup>; Courrat, A<sup>1</sup>; Pasquaud,  
5 S<sup>1</sup>; Lepage, M<sup>1</sup>

6

7 Addresses:

8 1 Cemagref - UR EPBX - EPBX

9 50 av. de Verdun, Gazinet

10 33612 CESTAS Cedex – FRANCE

11

12 2 Ifremer - Laboratoire Environnement-Ressources

13 ZP de Bregailon

14 BP n° 330 - 83507 La Seyne sur Mer - FRANCE

15

16 3 Ifremer - Département EMH

17 rue de l'Ile d'Yeu

18 44311 Nantes Cedex 03 – FRANCE

19

20

- 
- corresponding author : [hilaire.drouineau@cemagref.fr](mailto:hilaire.drouineau@cemagref.fr)

tel: +33 (0)5 57 89 27 08

fax: +33 (05)5 57 89 08 01

## 21 **Abstract**

22 In the context of the European Water Framework Directive (WFD), monitoring programs and  
23 related indicators have been developed to assess anthropogenic impacts on various components  
24 of aquatic ecosystems. While great precautions are usually taken when selecting and  
25 calculating relevant core metrics, little attention is generally paid to the generation of the  
26 multimetric indicator, i.e. the combination of the different core metrics. Indeed, most  
27 multimetric indicators are generated by simply averaging or summing metrics, without taking  
28 into account their sensitivity and their variability. Moreover, few indicators provide a rigorous  
29 estimate of the uncertainty of the assessments, while this estimation is essential for managers.  
30 In this context, we developed a Bayesian framework to build multimetric indicators aiming at  
31 improving those two weaknesses. This framework is based on two phases. First, pressure-  
32 impact statistical models are developed to quantify the impact of pressure on various fish  
33 metrics. Then the Bayesian theorem is applied to estimate probabilities of being at a certain  
34 anthropogenic pressure level from fish observation and pressure-impact models outputs. The  
35 Bayesian theorem allows to combine objectively the different core metrics, taking into account  
36 their sensitivity and their variability, and to provide rigorous uncertainty quantification, which  
37 is especially valuable in the WFD context.

38 The method is applied as illustrative example on transitional French water bodies to  
39 demonstrate its relevance, especially in the Water Framework Directive context though the  
40 method is generic enough to be applied in various contexts.

41

42 keywords: multimetric fish-based indicator, Bayesian method, pressure-impact models, Water  
43 Framework Directive, anthropogenic pressure, monitoring program, transitional waters.

44

45

## 46 **Introduction**

47 Multimetric fish-based indicators are often used to assess the ecological quality of aquatic  
48 ecosystems (Hughes and Oberdorff, 1999). Ideally, they are based on a set of non-redundant  
49 core metrics measured on fish assemblages. The main interest of such a tool is to provide an  
50 assessment that integrates several aspects of the fish assemblages through the different core  
51 metrics (Karr and Chu, 1999). Moreover, multimetric indicators are often considered as more  
52 sensitive and robust indicators of ecological quality than any of the individual metrics selected  
53 for their construction (Deegan *et al.*, 1997; Hughes *et al.*, 1998; Karr and Chu, 1999). A  
54 specific class of multimetrics indicator, called stressor-specific multimetrics indicator by  
55 Hering *et al.* (2006), is designed to detect the impact of a specific stressor on the ecosystem.  
56 This approach has been widely used in the context of the Water Framework Directive  
57 (European Water Framework Directive 2000/60/EC; WFD) to detect the impact of  
58 anthropogenic pressures on riverine and estuarine fish assemblages (Borja *et al.*, 2004; Breine  
59 *et al.*, in Press; Breine *et al.*, 2007; Coates *et al.*, 2007; Delpéch *et al.*, 2010; Franco *et al.*,  
60 2009; Martinho *et al.*, 2008; Pont *et al.*, 2009; Pont *et al.*, 2006; Uriarte and Borja, 2009).  
61 Hering *et al.* (2006) presented a “cook-book” to develop multimetric indicators that  
62 distinguishes 6 main steps: (i) selection of the form of the multimetric indicator (general or  
63 stressor specific), (ii) metric selection, (iii) metric calculation, (iv) multimetric indicator  
64 generation by combining the different metrics, (v) setting class boundaries and (vi) results  
65 interpretation. One of the main difficulties rests in step (iv): the indicator should objectively  
66 combine the different metrics and, in the case of a stressor specific indicator, it should take into  
67 account the sensitivity of the metric to the stressor. Additionally, a measure of uncertainty  
68 should be calculated, since bioassessments have little value without any uncertainty measures

69 (Kurtz *et al.*, 2001; Clarke and Hering, 2006; Clarke *et al.*, 2006; Carstensen, 2007; Beliaeff  
70 and Pelletier, 2011). Uncertainty measure is explicitly required by the WFD (European  
71 Communities, 2004). More specifically, it is of particular importance to be able to quantify the  
72 probability to don't be at least in "good" status class or not since the directive aims at  
73 achieving "good" water status for all water bodies.

74 Several approaches have been proposed to combine the core metrics. Delpech *et al.* (2010), for  
75 example, scored each metric independently; the final score being the average of the individual  
76 scores. Breine *et al.* (2007) proposed a statistical method to minimise type I and type II  
77 classification error. Pont *et al.* (2006) used reference sites to build a theoretical distribution of  
78 metrics in a reference state and developed a probabilistic framework to combine the metrics  
79 measured on other sites by calculating probabilities to belong to the reference distribution.

80 The aim of this paper is to propose a new approach that provides an original Bayesian  
81 framework to combine objectively the different metrics. The approach is based on two phases,  
82 which corresponds respectively to steps (ii) to (iii) and to step (iv) in Hering *et al.* (2006). In  
83 the first phase, pressure-impact models similar to those developed by Courrat *et al.* (2009) and  
84 Delpech *et al.* (2010) are developed. They enable objective selection and calculation of core  
85 metrics (steps (ii) and (iii) in Hering *et al.* (2006)). Then in the second phase, outputs of the  
86 models are used to compute probabilities of the ecosystem experiencing particular  
87 anthropogenic pressures levels by applying the Bayesian framework (step (iv) in Hering *et al.*  
88 (2006)). In this phase, the method used to combine metrics takes into account their sensitivity  
89 to the stressor and their variability. By providing a measure of uncertainty as required in the  
90 WFD context (European Communities, 2004), this method represents a great improvement  
91 with respect to existing methods, and is a great advance regarding the fourth step described by  
92 Hering *et al.* (2006) for the construction of multimetric indicators. As an illustrative example,  
93 the method is applied to French Mediterranean lagoons, classified as transitional waters for the

94 WFD.

95

96

## 97 **1 Material and methods**

### 98 ***1.1 Methods: description of the original approach***

99 Hering et al. (2006) cook book describes six main steps when developing a multimetric  
100 indicator. The method presented in this paper is partially based on Courrat *et al.* (2009) and  
101 Delpech *et al.* (2009), consequently, only the original part of the method, aiming at improving  
102 steps (ii) to (iv) from Hering *et al.* (2006), will be presented here.

103

#### 104 ***1.1.1 First phase: selection of candidate metrics using pressure-impact*** 105 ***models (step (ii) to (iii) in Hering et al. (2006))***

106 The idea when building a multimetric stressor-specific indicator is to select metrics among a  
107 list of candidate metrics and to combine them in an indicator to detect a gradient of stressor  
108 (Hering *et al.*, 2006). To build the indicator, a dataset with various candidate metrics (for  
109 example fish densities, number of species...) in columns measured on several fishing  
110 operations (in rows) is generally available. A measure or a proxy of the stressor is also required  
111 to check the correlations between the indicator and the stressor (Hering *et al.*, 2006).  
112 Pressure-impact approach (Courrat *et al.*, 2009; Delpech *et al.*, 2010) consists in developing  
113 pressure-impact statistical models by fitting generalized linear models (GLMs) that describe  
114 the impact of the stressor on the different candidate metrics, taking into account other  
115 covariates such as the variability due to the sampling procedure or to environmental  
116 characteristics. This approach is relevant to select candidate metrics since only metrics that are  
117 significantly impacted by the stressor are selected. The GLMs can be written on the matrix

118 form:

119 (1)  $g^{(j)}(E(M^{(j)})) = \alpha^{(j)} \cdot X + \beta^{(j)} \cdot Pr$

120 with  $g^{(j)}(E(M^{(j)}))$  the link transformed (function  $g^{(j)}$ ) expected value of the  $j$ -th metric,  $\alpha^{(j)}$  the  
121 regression parameters for covariables,  $X$  the model matrix for the covariables,  $\beta^{(j)}$  the  
122 regression parameter for the stressor and  $Pr$  the vector of stressor values.

123 Candidate metrics for which  $\beta^{(j)}$  are significantly different from zero, *i.e.* significantly impacted  
124 by the stressor, are potentially relevant metrics to include in the indicator.

125

126 *1.1.2 Second phase: computing probabilities of pressure levels given*  
127 *observed metrics (step (iv) in Herring et al. (2006))*

128 The aim of a stressor specific multimetric fish-based indicator is to evaluate the level of a  
129 specific stressor by observing metrics describing the fish assemblage given particular stressor-  
130 metric relationships (Hering *et al.*, 2006). To fulfil this objective, the present method uses the  
131 pressure-impact statistical models described above incorporating their results in a Bayesian  
132 framework. Applying the Bayes theorem makes possible to calculate the probability that a  
133 waterbody is in an ecological quality class given the fish data. GLMs likelihood functions are  
134 used to convert fish observations into probability densities; so that metrics can be combined on  
135 a common scale. Probability densities account for both the uncertainty of the model (through  
136 variance functions of the GLMs) and the sensitivity of the metric to the stressor (through the  
137 values of regression parameters).

138

139 The probability that the pressure level  $Pr$  is comprised within a given range  $[p_1, p_2]$  given  $I$  new  
140 fishing operations (denoted  $I, \dots, i, \dots, I$ ), which corresponds to an a posteriori probability in a

141 Bayesian framework, is:

142 (2)  $p(p_1 \leq Pr < p_2 | M^{(1)} = \{m_1^{(1)}, \dots, m_n^{(1)}\}, \dots, M^{(J)} = \{m_1^{(J)}, \dots, m_n^{(J)}\})$

143 with  $m_i^{(j)}$  observed values of the metric  $j$  (from 1 to  $J$ ) during fishing operation  $i$  (from 1 to  $n$ ).

144

145 The probability is equal to (cf. Annex):

146 (3)

147 
$$p(p_1 \leq Pr < p_2 | M = m) = \frac{\int_{\alpha} \int_{\beta} \int_{p_1}^{p_2} \left[ \prod_i \prod_j f_{M^{(j)}|Pr} (p, m_i^{(j)}) f_{Pr} (p) \cdot f_{A,B} (\alpha, \beta) \right] dp \cdot d\alpha \cdot d\beta}{\int_{\alpha} \int_{\beta} \int_{p_{min}}^{p_{max}} \left[ \prod_i \prod_j f_{M^{(j)}|Pr} (p, m_i^{(j)}) f_{Pr} (p) \cdot f_{A,B} (\alpha, \beta) \right] dp \cdot d\alpha \cdot d\beta}$$

148 with  $[p_{min}, p_{max}]$  the domain of definition of the stressor.  $f_{M^{(j)}|Pr, A, B} (p, m_i^{(j)}, \alpha, \beta)$  is the density of

149 probability of an observation given the matrix model and the regression parameters. It

150 corresponds to the likelihood of an observation in the GLMs, so that it can be directly

151 calculated as an output of the pressure-impact models (McCullagh and Nelder, 1989).

152  $f_{A,B} (\alpha, \beta)$  corresponds to a prior on the distribution of the regression parameters. Regression

153 parameters' estimates follow a multinomial distribution when fitting a model by maximum

154 likelihood (such as GLMs), consequently a multinormal (or multistudent for small datasets)

155 prior using GLMs estimates and corresponding variance-covariance matrix  $\Sigma$  can be used:

156  $\{A, B\} \sim N(\hat{\alpha}, \hat{\beta} | \Sigma)$ .  $f_{Pr} (p)$  is the equivalent of a prior in a Bayesian framework and enables

157 to include expert knowledge.

158

159 In the absence of precise knowledge, uninformative uniform prior may be used and if standard



160 error of regression parameters estimates are small compared to the GLMs dispersion  
161 parameters, equation 3 can be approximated by:

$$162 \quad (4) \quad p(p_1 \leq Pr < p_2 \mid M = m) = \frac{\int_{p_1}^{p_2} \left[ \prod_i \prod_j f_{M^{(j)}|Pr}(p, m_i^{(j)}) \cdot f_{Pr}(p) \right] dp}{\int_{p_{min}}^{p_{max}} \left[ \prod_i \prod_j f_{M^{(j)}|Pr}(p, m_i^{(j)}) \cdot f_{Pr}(p) \right] dp}$$

163

164 However any expert knowledge or meta-analysis results can be used to build more informative  
165 priors on the stressor, and priors on the regression parameters can be easily incorporated if  
166 estimation error are significant compared to observation errors.

167

168

169 By using this formula, the probability of being in any range of stressor can be easily calculated.

170

## 171 **1.2 Illustrative example**

172 In this illustrative example, the method is applied to generate a multimetric fish-based indicator  
173 for detecting anthropogenic pressure in lagoons.

174

### 175 **1.2.1 Fish data and anthropogenic pressure index**

176 Fourteen lagoons along the French Mediterranean coast are considered in this study (Table 1,  
177 Fig. 1). Lagoons were described by two physical factors (Table 1) that have proved to have a  
178 decisive effect on fish assemblages in lagoons (Perez-Ruzafa *et al.* 2007):

179 *Surf*: total surface area (km<sup>2</sup>)

180 *Sect*: channels cross-sectional area 1, 2, 3... (m<sup>2</sup>)

181 A dataset with 348 fishing operations (a fishing operation corresponds to a fyke-net fishing  
182 during 24 hours) carried out in the context of the Water Framework Directive was available.  
183 Fishing operations were conducted in spring and autumn using fyke nets in shallow water  
184 (between 0.5m and 2.5m).  
185 Each captured fish was identified at the species level, and then assigned to different functional  
186 guilds according to Delpech *et al.* (2010).  
187  
188 An anthropogenic pressure index was estimated using an approach similar to Courrat *et al.*  
189 (2009) and Delpech *et al.* (2010). A principal component analysis (PCA) was carried out on  
190 mean concentrations of four heavy metals (Cd, Zn, Cu, Pb) and one organic contaminant  
191 (S16HAPs) provided by the RINBIO biointegrator network set up by IFREMER (Andral *et al.*,  
192 2004) in twelve of the lagoons (data were not available for Biguglia and Grand Bagnas - Fig.  
193 1). All contaminants showed a strong correlation with the first axis of the PCA which  
194 represents 55% of the total variance. We thus used lagoons coordinates on this axis as a  
195 measure of contamination (Fig. 2). As suggested in Courrat *et al.* (2009) and Delpech *et al.*  
196 (2010) for French estuarine areas, we considered contamination as a proxy of global  
197 anthropogenic pressure in each lagoon.  
198 The RINBIO network also provides quality thresholds for each contaminant. Therefore it was  
199 possible to determine the thresholds of five quality classes on our anthropogenic pressure index  
200 by projecting the contaminant thresholds on the PCA factorial map (Fig. 2, Table 2).  
201 Considering this method, the twelve lagoons were in the two best quality classes.

202

### 203 1.2.2 Candidate metrics and pressure-impact models

204 Several metrics were calculated for each fishing operation (Table 3): total abundance, total

205 species richness, total number of fishes per guild, number of distinct species per guild. Total  
206 number of fishes were log-transformed ( $\log(x+1)$ ) to limit the influence of outliers (fishing  
207 events with high catches). Metrics were computed at the fishing operation level rather than at a  
208 larger scale in order to take into account the metric variability due to sampling protocol  
209 (Courrat *et al.*, 2009; Delpech *et al.*, 2010).

210

211 GLMs were built for each metric (Table 3). A stepwise backward procedure was used to select  
212 the most relevant and parsimonious models based on the corrected Akaike Information  
213 Criterion (AICc), which is for small datasets (Burnham and Anderson, 2002). Such pressure-  
214 impact models were fitted on data from the twelve lagoons monitored by the RINBIO network.

215

### 216 1.2.3 *Multimetric indicator generation and probabilities computation*

217 Graphical analysis of the residuals was carried out to check that GLMs assumptions were  
218 respected. In that case, standardised residuals of deviance were normally distributed  
219 (McCullagh and Nelder, 1989) so Pearson's correlation coefficients are appropriate to analyse  
220 residuals independence between core metrics.

221 After checking metric correlations, all 348 fishing operations from all the 14 lagoons  
222 (including lagoons in which pressure index was not available) were used to compute the  
223 probability that each lagoon belong to each stressor quality class using equation 4.

224

225 Results obtained with the above described dataset are presented here only as an illustrative  
226 example of possible outputs of the method given a plausible use in a WFD context.

227

228

229

## 230 **2 Results**

### 231 ***2.1 Selection of core metrics***

232 In the illustrative example presented here, four metrics responded as expected to an increase of  
233 the anthropogenic pressure index (Table 4). Values of estimated regression parameters for the  
234 pressure index were low (Table 4) but significantly different from 0. Both abundance metrics  
235 (*TD* and *DM*) and species richness metrics (*SR* and *NM*) were significantly "impacted" by the  
236 pressure index. Moreover, both pressure-impact models at the functional group level (*DM* and  
237 *NM*) and at the fish assemblage levels (*TD* and *SR*) were significant.

238 Graphical analysis of the residuals confirmed that GLMs assumptions were respected  
239 (consistence with the assumed distribution, with the assumed mean-variance relationship and  
240 independence). Correlation analysis tests theoretically rejected the independence assumptions.  
241 However, Pearson's correlation coefficients for the three metrics *TD*, *DM* and *SR* were very  
242 low (inferior to 0.24) and independence can be reasonably assumed.

243 Since standard-errors were small compared to observation errors and no expert knowledge was  
244 available, equation 4 rather than equation 3 was used to compute probabilities of being in a  
245 given pressure range.

246

### 247 ***2.2 Ecological quality assessment***

248 For each of the 14 surveyed lagoons, the posterior probability to be in each of the five quality  
249 classes given the observed fish metrics were computed using equation 4 (Fig. 3). Moreover, the  
250 fish indicator was very discriminant: for all lagoons, the most probable ecological quality class  
251 was superior >60%, except for Berre where we could not distinguish between good and very  
252 good classes.

253 Results were also analysed metric by metric (Fig. 4). For example, observed *TD* was relatively  
254 low in Or but high in Vaccarès compared to what was expected given the anthropogenic  
255 pressure index. Results at the core metrics level also illustrate that diagnostics based on a  
256 single metric are much more variable than results of the multimetric indicator; some individual  
257 metrics indicating medium or poor quality in some lagoons (for example, *DM* and *RT* in  
258 Biguglia).

259

## 260 **3 Discussion**

### 261 **3.1 Comments related specifically to the illustrative example**

262 The ecological quality of 14 French Mediterranean lagoons was assessed as an illustrative  
263 example, and further analysis should be carried out to improve the current application before  
264 using this multimetric fish indicator for management. More specifically, the present pressure  
265 index should be improved. Indeed, there was no available measure for oligohaline lagoons, and  
266 values for other lagoons were not really discriminant: they were all in the two best quality  
267 classes according to the pressure index leading to a dataset without much contrast. Moreover,  
268 available data were still limited (though they will increase with the implementation of WFD  
269 routine surveys) so it was necessary to use the same fishing observations for both GLMs  
270 construction and probabilities calculations (except for Biguglia and Grand Bagnas) whereas  
271 theoretically, the dataset used to build pressure-impact models should be distinct from the  
272 dataset used in the Bayesian calculation. Consequently, this paper will focus on the method  
273 used to build the indicator rather than on the illustrative example. So, specific results on  
274 lagoons or classification will not be discussed. However, with few modifications, the method  
275 will be relevant to assess the ecological status of those water bodies.

276

### 277 **3.2 Methodological aspects**

278 The approach proposed by Delpech *et al.* (2010) based on pressure-impact statistical models is  
279 especially relevant to describe the impact of a specific stressor such as anthropogenic pressure,  
280 on fish assemblages. However, the method to generate the multimetric indicator by combining  
281 the core metrics was rather subjective and did not provide any uncertainty quantification  
282 (Brind'Amour and Lobry, 2009). The aim of this paper was to propose a method to construct  
283 multimetric indicators that provide an objective scheme for combining metrics based on a  
284 Bayesian framework. The first phase is to develop pressure-impact GLMs (Courrat *et al.*, 2009;  
285 Delpech *et al.*, 2010) or any other models (general additive models, mixed models...) that  
286 provides a likelihood measure. GLMs outputs are then used to calculate probabilities that a  
287 water body belongs to an ecological quality class from fish observations. The present method  
288 presents some similarities with Bayesian discriminant analysis (Geisser, 1964; Keehn, 1965)  
289 since both methods compute probabilities of being in one pre-defined. However, our approach  
290 relies on the construction of pressure-impact models in a first phase, that allows a great  
291 variability on the form of the distribution of the metric given the pressure and the regression  
292 model, and to easily account for environmental covariates. Moreover, this first phase provides  
293 valuable information at the metric level that can be analysed in the light of ecological concepts  
294 (e.g. (Nicolas *et al.*, 2010a; Nicolas *et al.*, 2010b) for examples with fish data on estuaries),  
295 rather than building discriminant functions that are only based on a statistical criterion and are  
296 often less flexible.

297

298 One of the main advantages of the method is that it combines core metrics taking into account  
299 both (i) the sensitivity of the metric to the stressor through the regression parameter in the  
300 pressure-impact statistical model and (ii) the uncertainty of the statistical model through the  
301 variance function of the GLMs included in the likelihood computation (equation 4).

302 Furthermore, while Delpech *et al.* (2010) couldn't include any metric based on species richness  
303 in their final multimetric indicator because they were not powerful discriminant metrics, the  
304 present method can include them and incorporate their weak, though perhaps important,  
305 discriminant ability in the collective. The method also provides a measure of uncertainty by  
306 computing posterior probabilities that the water body is in each pre-defined quality class (Fig.  
307 3). This uncertainty estimation is especially important for managers and is required by the  
308 WFD (European Communities, 2004). More specifically, the fact that it is possible to estimate  
309 the for a water body to be "good or better" ecological status or not is especially important  
310 given the importance of this threshold in the directive. This particular aspect is probably the  
311 greatest advantage of this method.

312 Moreover, in this approach, the problem is considered in a different way to what is usually  
313 done. Generally, pressure indices are used to predict expected fish assemblages metrics values  
314 at various pressure levels in order to estimate metric thresholds (Courrat *et al.*, 2009; Delpech  
315 *et al.*, 2010). In the present approach, the method is similar in a first phase (development of  
316 pressure-impact statistical models), but then the problem is reversed by trying to predict a level  
317 of pressure from fish metrics so that it is possible to define a priori pressure thresholds. This  
318 approach seems much more relevant in an operational monitoring context when stakeholders  
319 have to take appropriate management decisions based upon the latest observations of  
320 ecological state.

321 As mentioned by Hughes *et al.* (1998) and Karr & Chu (1999), multimetric indices are more  
322 robust than single metrics because the metrics combination generally makes the indicators less  
323 variable than each core metric considered individually. This is illustrated by the higher  
324 variability of the results when providing a diagnostic for each metric (Fig. 4) compared to the  
325 results when applying the global multimetric indicator (Fig. 3).

326 An other important aspect to consider when developing an indicator is the consequences of

327 misclassification. For example, in the illustrative example, three lagoons were misclassified by  
328 the fish-based indicator with respect to the pressure index. Three main reasons may explain this  
329 result: (i) the natural variability of biological measures (however the sampling procedure is  
330 designed to minimize this aspect), (ii) pressure index was not perfectly relevant since it is only  
331 a proxy of global anthropogenic pressures or (iii) lagoons are more or less resilient with respect  
332 to anthropogenic pressure. The two last points are especially interesting. The fish indicator for  
333 Or Lagoon is pessimistic compared to the pressure index because  $TD$  is low compared to the  
334 pressure-impact model prediction, *i.e.* it seems that Or is especially sensitive to an increase of  
335 pressure or that pressure index was underestimated. On the other hand, the fish indicator  
336 provides a more pessimistic diagnostic for Vaccarès than the pressure index because  $TD$  is high  
337 compared to what is predicted by the pressure-impact model, possibly because Vaccarès is less  
338 sensitive or because the pressure index was overestimated for this lagoon. Consequently, the  
339 indicator tends to penalize lagoons which are more sensitive to pressure and/or to consistently  
340 deal with pressure index misspecification, which is consistent with the precautionary approach.  
341 Another advantage of the Bayesian approach relies on the possibility to include prior  
342 knowledge in the indicator. Expert knowledge or results of independent analysis can be used to  
343 build informative prior. That knowledge is difficult to incorporate in traditional methods. Yet, it  
344 is interesting to include it by building *a priori* pressure density distribution which would  
345 influence the results (equation 3).

346 Moreover, in the absence of previous expert knowledge and when standard-errors of the GLMs  
347 are small compared to dispersion as it is in the illustrative example, equation 4 can be used  
348 rather than equation 3. Since equation 4 consists in a single integral, it is possible to compute  
349 the probabilities using usual numerical integration algorithm. However, in most situations, this  
350 solution will not be possible, and Bayesian inference algorithm, such as Gibbs Samplings will  
351 be appropriate. In that case, the use of traditional Bayesian diagnostic tools will of course be



352 necessary to check relevance of the results, and more specifically the convergence of the  
353 algorithm.  
354 Finally, the present method is completely generic and can be implemented in many situations,  
355 as long as it is possible to build pressure-impact statistical models (enough data and, sufficient  
356 knowledge on determinant abiotic factors). More specifically, any kind of metrics can be  
357 included in the indicator: density, numbers, proportion (GLM with binomial family) or other  
358 metrics as soon as likelihood of new observations given a pressure can be calculated (the  
359 principle limitation will generally concern the number of available data to fit pressure impact-  
360 models). Consequently, it may be applied for many other stressor-specific indicators, especially  
361 indicators developed in the WFD context. However, a difference with traditional indicators is  
362 that generally, a set of thresholds are defined for each metric in order to get a score by metric  
363 which are then aggregated to get the assessment, whereas in our approach, only one set of  
364 threshold is required and is directly defined on the pressure index. Consequently, ecological  
365 quality ratio (EQR, the WFD defined EQR as a ratio between the actual level of an indicator  
366 and the reference level of the indicator) does not have exactly the same meaning as in other  
367 indicators. However, this approach is more consistent with the initial objective of a multimetric  
368 indicator, *i.e.* assessing levels of pressure by analysing metrics, and more consistent with the  
369 approach recommended in the annexes III and IV of the WFD guidance document (European  
370 Communities, 2009).

371

## 372 **Acknowledgements**

373 We want to acknowledge Dr Patrick Lambert and Dr Daniel Duplisea for fruitful discussions  
374 on the method. We also would like to thank the Rhone-Mediterranean and Corsica Water  
375 Agency for their collaboration and their support in the data collection. This work is part of the

376 WISER European project (Water bodies in Europe: Integrative Systems to assess Ecological  
377 status and Recovery) funded by the European Union under the 7th Framework Programme,  
378 Theme 6 (Environment including Climate Change) (contract No. 226273),  
379 <http://www.wiser.eu>.

## References

- Andral, B., Stanisiere, J.Y., Sauzade, D., Damier, E., Thebault, H., Galgani, F., Boissery, P., 2004. Monitoring chemical contamination levels in the Mediterranean based on the use of mussel caging. *Mar. Pollut. Bull.* 49: 704-712.
- 381 Beliaeff, B., Pelletier, D., 2011. A general framework for indicator design and use with  
382 application to the assessment of coastal water quality and marine protected area management.  
383 *Ocean & Coastal Management* 54: 84-92.
- Borja, Á., Franco, J., Valencia, V., Bald, J., Muxika, I., Belzunce, M.J., Solaun, O., 2004. Implementation of the European water framework directive from the Basque country (northern Spain): a methodological approach. *Mar. Pollut. Bull.* 48: 209-218.
- Breine, J., Quataert, P., Stevens, M., Ollevier, F., Volckaert, F., Van den Bergh, E., Maes, J., in Press. A zone-specific fish-based biotic index as a management tool for the Zeeschelde estuary (Belgium). *Mar. Pollut. Bull.*
- Breine, J.J., Maes, J., Quataert, P., Van den Bergh, E., Simoens, I., Van Thuyne, G., Belpaire, C., 2007. A fish-based assessment tool for the ecological quality of the brackish Schelde estuary in Flanders (Belgium). *Hydrobiologia.* 575: 141.
- Brind'Amour, A., Lobry, J., 2009. Assessment of the ecological status of coastal areas and estuaries in France, using multiple fish-based indicators: a comparative analysis on the Vilaine estuary. *Aquat. Living Resour.* 22: 559-572.
- Burnham, K.P., Anderson, D.R., 2002. *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach.* Springer, New York.
- Carstensen, J., 2007. Statistical principles for ecological status classification of Water Framework Directive monitoring data. *Mar. Pollut. Bull.* 55: 3-15.

Clarke, R., Davy-Bowker, J., Sandin, L., Friberg, N., Johnson, R., Bis, B. 2006. Estimates and comparisons of the effects of sampling variation using 'national' macroinvertebrate sampling protocols on the precision of metrics used to assess ecological status. *Hydrobiologia* 566: 477-503.

Clarke, R., Hering, D. 2006. Errors and uncertainty in bioassessment methods - Major results and conclusions from the STAR project and their application using STARBUGS.

*Hydrobiologia* 566: 433-439.

Coates, S., Waugh, A., Anwar, A., Robson, M., 2007. Efficacy of a multi-metric fish index as an analysis tool for the transitional fish component of the Water Framework Directive. *Mar. Pollut. Bull.* 55: 225-240.

Courrat, A., Lobry, J., Nicolas, D., Laffargue, P., Amara, R., Lepage, M., Girardin, M., Le Pape, O., 2009. Anthropogenic disturbance on nursery function of estuarine areas for marine species. *Estuarine, Coastal and Shelf Science.* 81: 179-190.

Deegan, L.A., Finn, J.T., Ayvazian, S.G., Ryder-Kieffer, C.A., Buonaccorsi, J., 1997.

Development and validation of an estuarine biotic integrity index. *Estuaries.* 20: 601-617.

Delpech, C., Courrat, A., Pasquaud, S., Lobry, J., Le Pape, O., Nicolas, D., Boët, P., Girardin, M., Lepage, M., 2010. Development of a fish-based index to assess the ecological quality of transitional waters: The case of French estuaries. *Mar. Pollut. Bull.* 60: 908-918.

European Communities, 2004. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). WFD CIS Guidance Document No7, European Communities, EC, Luxembourg, 153p.

European Communities, 2009. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). WFD CIS Guidance Document No14, European Communities, EC, Luxembourg, 55p.

Franco, A., Torricelli, P., Franzoi, P., 2009. A habitat-specific fish-based approach to assess the

ecological status of Mediterranean coastal lagoons. *Mar. Pollut. Bull.* 58: 1704-1717.

Geisser, S., 1964. Posterior odds for multivariate normal distributions. *Journal of the Royal Society Series B: Methodological.* 26: 69-76.

Hering, D., Feld, C., 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. *Hydrobiologia.* 566: 311-324.

Hughes, R.M., Kaufmann, P.R., Herlihy, A.T., Kincaid, T.M., Reynolds, L., Larsen, D.P., 1998. A process for developing and evaluating indices of fish assemblage integrity. *Canadian Journal of Fisheries and Aquatic Sciences.* 55: 1618-1631.

Hughes, R.M., Oberdorff, T., 1999. Applications of IBI concepts and metrics to waters outside the United States, in: Simon, T.P. (Ed.), *Assessing the Sustainability and Biological Integrity of Water Resource Quality Using Fish Communities*, CRC Press, Boca Raton, Florida, pp. 79–96.

Karr, J.R., Chu, E.W., 1999. *Restoring life in running waters: better biological monitoring.* Island Press, Washington DC.

Keehn, D.G., 1965. A note on learning for Gaussian properties. *IEEE Trans. on Information Theory.* 11: 126-132.

Kurtz, J.C., Jackson, L.E., Fisher, W.S., 2001. “Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency’s Office of Research and Development”. *Ecol. Ind.* 1: 49–60.

Martinho, F., Viegas, I., Dolbeth, M., Leitão, R., Cabral, H., Pardal, M., 2008. Assessing estuarine environmental quality using fish-based indices: Performance evaluation under climatic instability. *Mar. Pollut. Bull.* 56: 1834-1843.

McCullagh, P., Nelder, J.A., 1989. *Generalized Linear Models.* Chapman & Hall, London.

Nicolas, D., Lobry, J., Le Pape, O., Boët, P., 2010a. Functional diversity in European estuaries:

Relating the composition of fish assemblages to the abiotic environment. *Estuarine, Coastal and Shelf Science*. 88: 329-338.

Nicolas, D., Lobry, J., Lepage, M., Sautour, B., Le Pape, O., Cabral, H., Uriarte, A., Boët, P., 2010b. Fish under influence: A macroecological analysis of relations between fish species richness and environmental gradients among European tidal estuaries. *Estuarine, Coastal and Shelf Science*. 86: 137-147.

Perez-Ruzafa A., Mompean M.C., Marcos C., 2007. Hydrographic, geomorphologic and fish assemblage relationships in coastal lagoons. *Hydrobiologia* 577:107-125

Pont, D., Hughes, R.M., Whittier, T.R., Schmutz, S., 2009. A Predictive Index of Biotic Integrity Model for Aquatic-Vertebrate Assemblages of Western US Streams. *Transactions of the American Fisheries Society*. 138: 292-305.

Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., Schmutz, S., 2006. Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages. *Journal of Applied Ecology*. 43: 70-80.

Uriarte, A., Borja, A., 2009. Assessing fish quality status in transitional waters, within the European Water Framework Directive: Setting boundary classes and responding to anthropogenic pressures. *Estuarine, Coastal and Shelf Science*. 82: 214-224.

## 384 **Annex: demonstration of the probability computation**

385 The objective of a stressor specific multimetric index is to detect the effect of the stressor by  
 386 observing some metrics on the fish assemblage. Consequently, we are interesting in computing  
 387 the probability that the pressure is included in a certain range given  $I$  new fishing operations  
 388 (denoted  $I, \dots, i, \dots, I$ ):

$$389 \quad (A1) \quad p(p_1 \leq Pr < p_2 \mid M^{(1)} = \{m_1^{(1)}, \dots, m_n^{(1)}\}, \dots, M^{(J)} = \{m_1^{(J)}, \dots, m_n^{(J)}\})$$

390 with  $m_i^{(j)}$  the value of metric  $j$  measured in the  $i$ -th fishing operation,  $n$  the number of fishing  
 391 operation,  $J$  the number of selected metrics.

392

393 We use the following notations

394  $M^{(j)}$  : random variable representing metric  $j$  (from 1 to  $J$ )

395  $m_i^{(j)}$ : value metric  $j$  in fishing operation  $i$  (from 1 to  $n$ )

396  $Pr$ : random variable of the pressure index

397  $p_1$  and  $p_2$ : thresholds of an ecological quality class corresponding to thresholds on the pressure  
 398 index

399  $f_X$ : density function of random variable  $X$

400  $A$  and  $B$  random variables corresponding to the GLM regression parameters

401

402 By definition of a density function, we have

$$403 \quad (A2) \quad p(p_1 \leq Pr < p_2 \mid M = m) = \int_{p_1}^{p_2} f_{Pr \mid M^{(1)}, \dots, M^{(J)}}(p, m_1^{(1)}, \dots, m_n^{(J)}) dp$$

404 with  $f_{Pr|M^{(1)}, \dots, M^{(J)}}$  the density function of the variable  $Pr | M^{(1)}, \dots, M^{(J)}$ . Given that a  
 405 conditional density is equal to the ratio of the joint density function over the marginal density.

$$p(p_1 \leq Pr < p_2 | M = m) = \int_{p_1}^{p_2} \frac{f_{Pr, M^{(1)}, \dots, M^{(J)}}(p, m_1^{(1)}, \dots, m_n^{(J)})}{f_{M^{(1)}, \dots, M^{(J)}}(m_1^{(1)}, \dots, m_n^{(J)})} dp$$

406 (A3)

$$= \frac{\int_{p_1}^{p_2} f_{Pr, M^{(1)}, \dots, M^{(J)}}(p, m_1^{(1)}, \dots, m_n^{(J)}) dp}{f_{M^{(1)}, \dots, M^{(J)}}(m_1^{(1)}, \dots, m_n^{(J)})}$$

407

408 Because of the total conditional probability:

$$p(p_1 \leq Pr < p_2 | M = m) = \frac{\int_{p_1}^{p_2} f_{Pr, M^{(1)}, \dots, M^{(J)}}(p, m_1^{(1)}, \dots, m_n^{(J)}) dp}{\int_{p_{min}}^{p_{max}} f_{M^{(1)}, \dots, M^{(J)}|Pr}(p, m_1^{(1)}, \dots, m_n^{(J)}) f_{Pr}(p) dp}$$

409 (A4)

410 with  $p_{min}$  and  $p_{max}$ , the minimum and maximum value of the pressure index.

411 Given that a conditional density is equal to the ratio of the joint density function over the  
 412 marginal density.

$$p(p_1 \leq Pr < p_2 | M = m) = \frac{\int_{p_1}^{p_2} f_{M^{(1)}, \dots, M^{(J)}|Pr}(p, m_1^{(1)}, \dots, m_n^{(J)}) f_{Pr}(p) dp}{\int_{p_{min}}^{p_{max}} f_{M^{(1)}, \dots, M^{(J)}|Pr}(p, m_1^{(1)}, \dots, m_n^{(J)}) f_{Pr}(p) dp}$$

413 (A5)

414 If the observations are independent:



$$415 \quad (A6) \quad p(p_1 \leq Pr < p_2 \mid M = m) = \frac{\int_{p_1}^{p_2} \left[ \prod_i \prod_j f_{M^{(j)}|Pr} (p, m_i^{(j)}) \cdot f_{Pr} (p) \right] dp}{\int_{p_{min}}^{p_{max}} \left[ \prod_i \prod_j f_{M^{(j)}|Pr} (p, m_i^{(j)}) \cdot f_{Pr} (p) \right] dp}$$

416

417 GLMs provides a measure of  $f_{M^{(j)}|Pr,A,B} (p, m_i^{(j)}, \alpha, \beta)$ , therefore we reformulate equation A6 as

418 (A7)

$$419 \quad p(p_1 \leq Pr < p_2 \mid M = m) = \frac{\iint_{\alpha \beta} \int_{p_1}^{p_2} \left[ \prod_i \prod_j f_{M^{(j)}|Pr} (p, m_i^{(j)}) \cdot f_{Pr} (p) \cdot f_{A,B} (\alpha, \beta) \right] dp \cdot d\alpha \cdot d\beta}{\iint_{\alpha \beta} \int_{p_{min}}^{p_{max}} \left[ \prod_i \prod_j f_{M^{(j)}|Pr} (p, m_i^{(j)}) \cdot f_{Pr} (p) \cdot f_{A,B} (\alpha, \beta) \right] dp \cdot d\alpha \cdot d\beta}$$

420

421 **Tables**

**Table 1.** Main characteristics of the different lagoons, which are located along the French Mediterranean coast (Fig. 1) and number of fishing operations per season (a fishing operation corresponds to a fyke-net fishing during 24 hours).

*Surf*: lagoon surface (km<sup>2</sup>). *Sect*: cross-sectional area of the inlets (m<sup>2</sup>).

Grand Bagnas and Méjean does not have inlet directly connected to the sea

Lagoon	<i>Surf</i>	<i>Sect</i>	Number of fishing operations	
			Spring	Summer
Bages-Sigean	38	600	12	15
Berre	133	517	15	15
Biguglia	14	3	8	8
Diana	5	61	8	10
Grand Bagnas	2	0	15	16
La Palme	5	25	16	16
Méjean	7	0	8	8
Or	33	15	7	8
Palo	1	10	8	8
Prévost	3	53	8	8
Salses-Leucate	54	367	23	29
Thau	69	237	30	15
Urbino	8	31	8	12
Vaccarès	102	15	7	7

422

423

**Table 2.** Value of  $p_1$ ,  $p_2$ ,  $p_{min}$  and  $p_{max}$  for the 5 classes.

Pressure class	$p_1$	$p_2$	$p_{min}$	$p_{max}$
Very good	-6.00	0.65	-6.00	25.00
Good	0.65	6.72	-6.00	25.00
Medium	6.72	12.81	-6.00	25.00
Bad	12.81	18.90	-6.00	25.00
Very Bad	18.90	25.00	-6.00	25.00

**Table 3.** List of selected candidate metrics and corresponding characteristics of the pressure-impact GLMs.

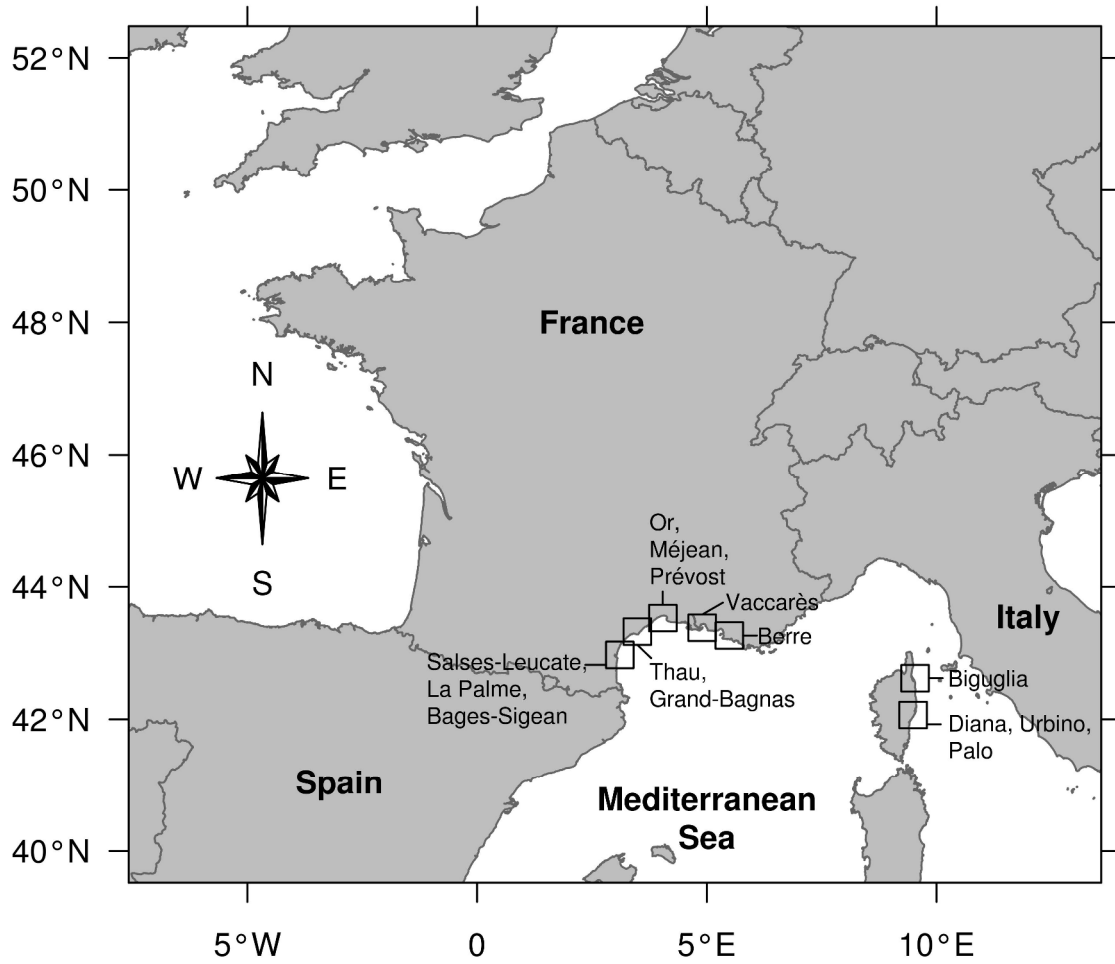
Type of metrics	Metrics	Definition	GLMs: Family, link function
Fish assemblage	SR	Total species richness	Poisson, log
	TD	Log(Total number of captured fishes +1)	Gaussian, identity
Ecological guilds	NMIG	Number of distinct migrant species	Poisson, log
	DMIG	Log(number of migrant fishes+1)	Gaussian, identity
	NFW	Number of distinct freshwater species	Poisson, log
	DFW	Log(number of freshwater fishes+1)	Gaussian, identity
	NMJ	Number of distinct marine juveniles species	Poisson, log
	DMJ	Log(number of marine juveniles fishes+1)	
	NMS	Number of distinct marine seasonal migrants species	Poisson, log
	DMS	Log(number of marine seasonal migrants fishes+1)	Gaussian, identity
	NM	Number of distinct marine species	Poisson, log
DM	Log(number of marine fishes+1)	Gaussian, identity	
Trophic	NIB	Number of distinct benthic invertebrate predators species	Poisson, log
	DIB	Log(number of benthic invertebrate predators fishes+1)	Gaussian, identity
	NF	Number of distinct fish feeders species	Poisson, log
	DF	Log(number of fish feeders fishes+1)	Gaussian, identity
Vertical distribution	NB	Number of distinct benthic species	Poisson, log
	DB	Log(number of benthic fishes+1)	Gaussian, identity

**Table 4.** Kept candidate metrics and corresponding pressure-impact models, pressure index regression parameter and corresponding p-value

Metric	Definition	Model	Regression parameter	P-value
SR	Total species richness Log(Total number of captured fishes +1)	<i>Sal_Class+Season+Pr</i>	-0.04	0.01
TD		<i>Sect+Pr</i>	-0.32	0.00
NM	Number of distinct marine species	<i>Sect+Surf+Pr</i>	-0.16	0
DM	Log(number of marine fishes+1)	<i>Pr</i>	-0.12	0.05

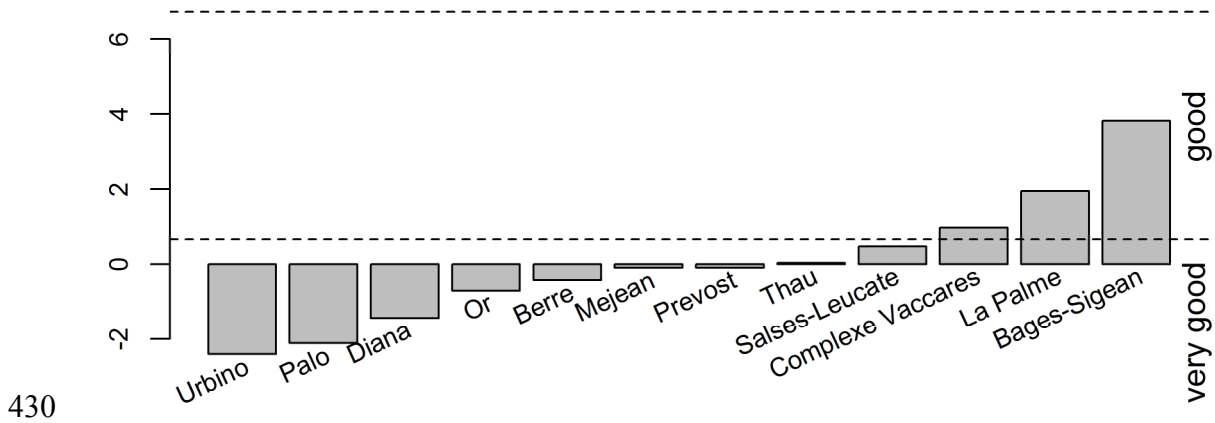
426

427 **Figures**



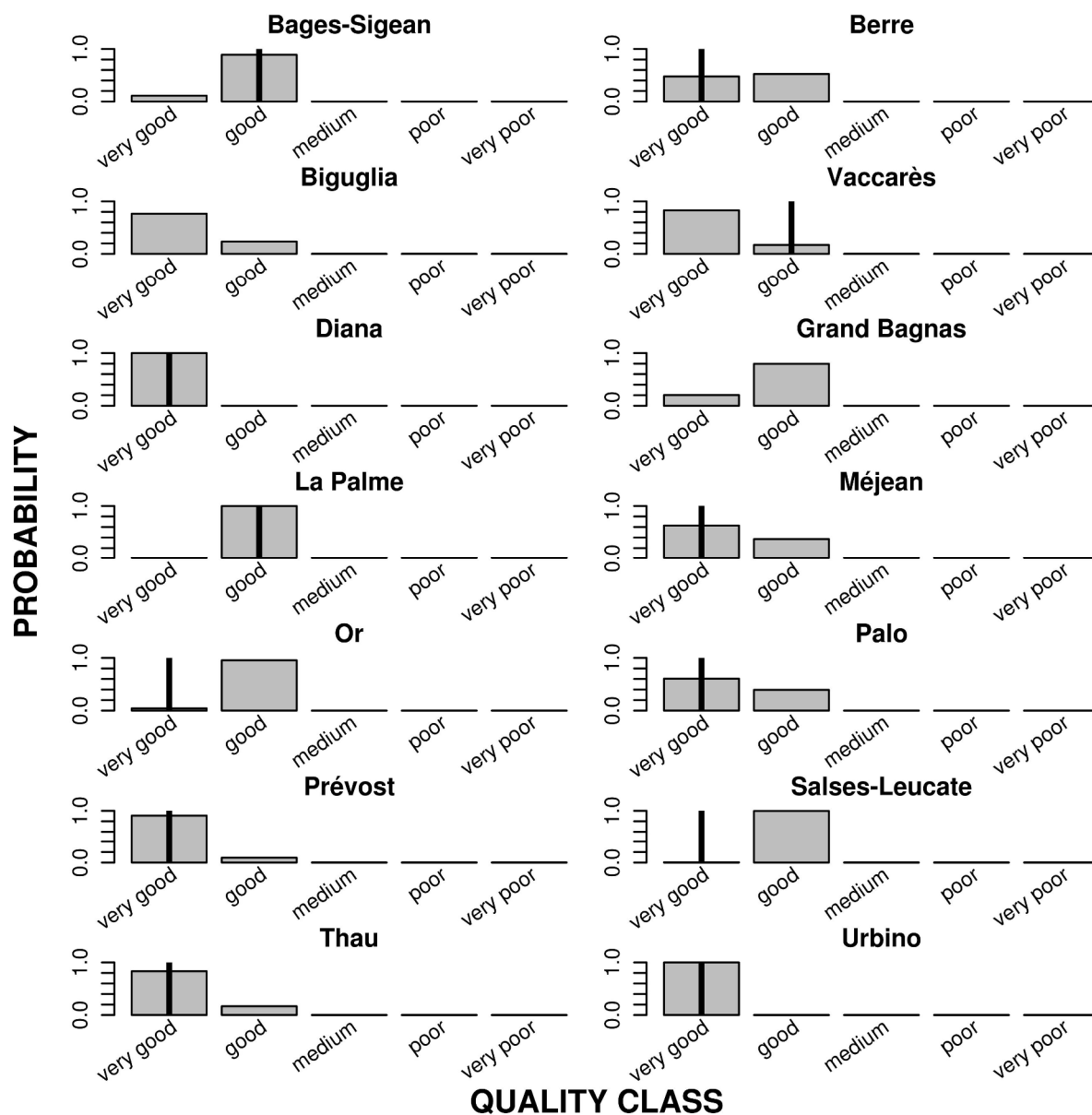
428

429 **Fig. 1.** Maps of the different lagoons considered in this study.



430

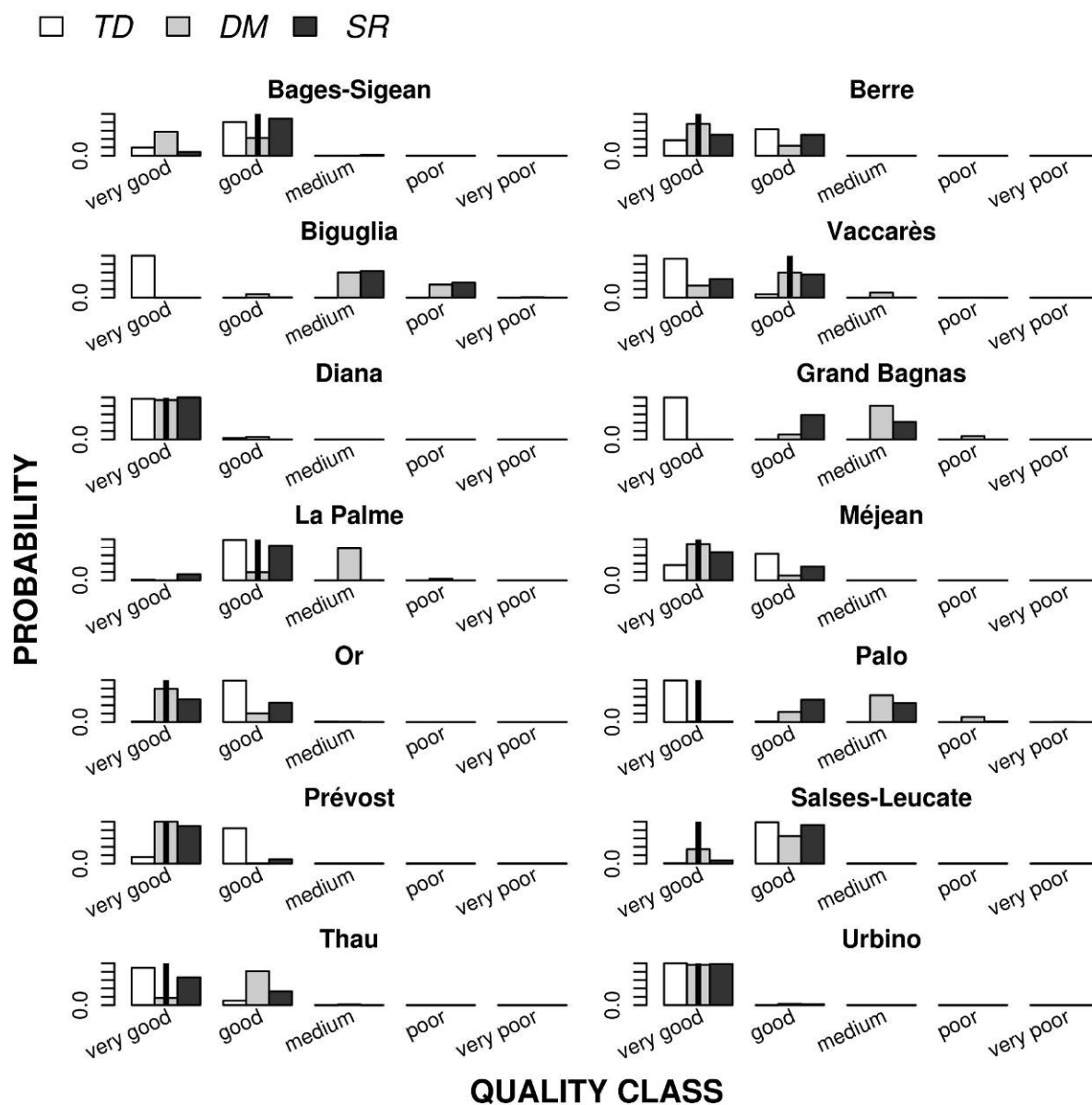
431 **Fig. 2.** Calculated anthropogenic pressure index for each lagoon (a high PCA coordinate  
432 implies a high level of pressure) and corresponding quality classes (“very good”, “good”,  
433 “medium”, “poor”, “very poor”) thresholds (dotted lines).



434

435 **Fig. 3.** Posterior probability to be in a quality class given the observations (barplot) and  
 436 pressure index quality class (from RINBIO, vertical bold line).





437

438 **Fig. 4.** Posterior probability to be in a quality class given *TD* (white bars), *DM* (light grey  
 439 bars), *RT* (dark grey bars) considered individually, and pressure index quality class (from  
 440 RINBIO, vertical bold line).

441