

Bayesian analysis of three indices for lagoons ecological status evaluation

Alessio Pollice · Serena Arima · Giovanna Jona Lasinio · Alberto Basset · Ilaria Rosati

Received: date / Accepted: date

Abstract The Water Framework Directive (WFD) recognizes benthic macroinvertebrates as a good biological quality element for transitional waters as they are the most exposed to natural variability patterns characteristic of these ecosystems, due to their life cycles and space-use behavior. In this paper we consider the performance of three multimetric indices (namely M-AMBI, BITS and ISS) based on benthic macroinvertebrates abundances, aiming at assessing the ecological status of lagoons and likely to respond differently to different sources of stress and natural variability. In order to investigate the possible contrasting behavior of the three multimetric indices, we propose a Bayesian hierarchical model in which they are jointly modeled as functions of abiotic covariates, external anthropogenic pressure indicators and lagoon effects. The proposed model is applied to data from three lagoons in Apulia and assessed using multiple diagnostic tools. The joint sensitivity of lagoon quality evaluations to available covariates is thus investigated.

A. Pollice
Dipartimento di Scienze Economiche e Metodi Matematici, Università degli Studi di Bari Aldo Moro, Largo Abbazia Santa scolastica 53, 70124 Bari
Tel.: +39-080-5049243
E-mail: alessio.pollice@uniba.it

S. Arima
Dipartimento di Metodi e Modelli per l'Economia, il Territorio e la Finanza, Sapienza Università di Roma, via del Castro Laurenziano 9, 00161 Roma

G. Jona Lasinio
Dipartimento di Scienze Statistiche, Sapienza Università di Roma, via del Castro Laurenziano 9, 00161 Roma,

A. Basset, I. Rosati
Dipartimento di Scienze e Tecnologie Biologiche ed Ambientali, Università del Salento, Prov.le Lecce-Monteroni, 73100 Lecce

Keywords Multivariate multiple regression, Correlated responses, Bayesian inference, Ecosystem status

1 Introduction

Lagoons represent important and fragile ecotone ecosystems (Basset *et al.*, 2013) in the coastal landscape, however their geographic position along the coast and their close relation with terrestrial ecosystems make these environments especially vulnerable to anthropogenic pressures (Viaroli *et al.*, 2005; Ji and Chang, 2005). The need to act has been acknowledged by politicians and legislation has been adopted to stop further deterioration and restore lagoons healthy state. The Water Framework Directive (WFD, 2000/60/EC) requires EU Member States to assess the ecological status of each water body in Europe and to ensure a sustainable management such that good ecological quality of all water bodies will be obtained by 2015. Lagoons can be clustered into types (Basset *et al.*, 2006; Lucena-Moya *et al.*, 2009), which display several different internal gradients of physical conditions and hence of the biota associated with them. Here the ecological status of aquatic ecosystems is defined in terms of the quality of the biological community, as well as the systems' hydrological and chemical characteristics. Environmental indices are a common tool used to communicate the overall status of environmental systems (Sadiq and Tesfamariam, 2008; Su *et al.*, 2012). Several simple indicators as the Shannon-Wiener index, the Margalef index and the AMBI index account for the composition and abundance of biological communities and are widely used in the ecological literature. Multimetric indices, combining simple indices as multiple sources of information, focus on benthic macroinvertebrates which are known to be sensitive to both natural and anthropogenic pressures (Borja *et al.*, 2011; Marchini and Marchini, 2006). In this paper, we will focus on three multimetric indices: M-AMBI (Muxika *et al.*, 2007; Borja *et al.*, 2004), BITS (Mistri and Munari, 2008) and ISS (Basset *et al.*, 2012). The so called *a priori* approach to lagoon classification by multimetric ecological quality indices was introduced in Borja *et al.* (2000) and used in Borja *et al.* (2004) and Basset *et al.* (2012). According to some reference samples, the authors choose boundary values of the indices to define ecological status classes and classify the lagoons according to these values. However the proposed indices are likely to respond differently to different sources of stress and natural variability components, adding uncertainty to resulting classifications. An alternative more objective *a posteriori* approach is based on classification boundaries set according to linear mixed models taking lagoons variability into account (Barbone *et al.*, 2012). As multi-stressors can commonly act on the same lagoon, both approaches are leaving anyway some degree of uncertainty in the ecological status classification by different multi-metric indices, with

a same lagoon classifiable into different ecological quality classes using different metrics. In order to emphasize clusters of ecosystems resulting from the integration of different multimetric indices, a multivariate Bayesian hierarchical model was recently proposed (Arima *et al.*, 2013) in which four multimetric indices are jointly modelled through a multivariate normal mixture accounting for the dependence on relevant abiotic covariates. Here we use a statistical model linking values of three multimetric ecological quality indices to indicators of water parameters and anthropogenic pressures, with the purpose of properly understanding what drives indices to disagree in classifying the ecological status of lagoons. A Bayesian hierarchical model, allowing for the inclusion of multiple sources of information and external prior knowledge, is adopted to produce a sensitivity analysis, i.e. to quantify the effects of input variables as abiotic covariates and indicators of anthropogenic pressures on the different performance of multimetric ecological quality indices (Piffady *et al.*, 2013).

The paper is organized as follows: in Section 2 we describe the design of the experiment leading to data on benthic macroinvertebrates, chemical and physical water parameters and expert's opinion evaluations of external anthropogenic pressures. In the same section we introduce the notation and define a multivariate hierarchical modelling strategy, including some details on the actual implementation of the estimation algorithms, model choice and assessment issues. Section 3 is devoted to describe some general results and discuss the sensitivity of lagoon ecosystems evaluations to available abiotic information. In Section 4 we conclude with a brief discussion and directions for future work.

2 Materials and methods

In order to reveal environmental conditions determining discrepant evaluations of the ecological quality of lagoon ecosystems we considered a database gathering biological data from 8 sampling campaigns at 15 monitoring stations with abiotic information characterizing times and sites. Such data require the formulation of a multivariate model to estimate the effects of water parameters and pressure gradients on multimetric indices, considering correlated responses and lagoon effects. Model parameter estimates summarize the sensitivity of the biotic evaluations of lagoons' quality to abiotic features of the ecosystems.

2.1 Experimental design and data structure

Data on benthic macroinvertebrates colonizing various habitat types were collected in 3 transitional water ecosystems in Apulia, Italy: namely Alimini, Lesina and Varano. Four seasonal field sampling campaigns were performed for each

of the two years 2008 and 2009 in the three ecosystems. On the whole, 15 monitoring sites were considered in the three lagoons. For each sampling site three replicates of benthic samples were collected by manual Reineck box-corer (0.03 m²) in Alimini and an Ekman-Birge grab in Lesina and Varano. In the laboratory, samples were sorted under a stereomicroscope, identified to the lowest possible taxonomic level, counted, measured individually (total length for most taxa) and weighted. For each lagoon, the macrobenthic community was examined in order to build two of the most common indices proposed in the Water Framework Directive and a new index for benthic assessment. The multimetric indices M-AMBI (Muxika *et al.*, 2007), BITS (Mistri and Munari, 2008), and ISS (Basset *et al.*, 2012) were calculated at the replicate level. Chemical and physical water parameters (water temperature, dissolved oxygen, salinity, pH and chlorophyll) were monitored at each station. Experts' opinion evaluations of four external anthropogenic pressures (agricultural diffuse inputs, domestic discharges, industrial discharges and fin fisheries) were also made available at each monitoring station; quantifications were given according to a score scale ranging from 0 (no pressure) to 3 (intense pressure).

2.2 Some data summaries

The three biotic indices have different theoretical range and are usually considered in association with boundaries defined in the literature. Within the mentioned *a priori* approach boundaries allow to classify lagoon ecosystems in five quality categories (bad, poor, moderate, good, high), though they do not necessarily correspond to discontinuities in the data, as Fig. 1 shows for the benthic samples described in section 2.1. In some circumstances M-AMBI, BITS and ISS can be in disagreement and lead to contrasting assessments of the ecological status of the same ecosystem. The relation among the three indices for the data at hand is expressed in Fig. 2: M-AMBI and ISS have stronger linear correlation ($R \simeq 0.84$), M-AMBI has higher variability for increasing values of BITS ($R \simeq 0.25$), while BITS and ISS show a weaker dependence ($R \simeq 0.25$) and a cluster of records with large values of BITS and small values of ISS (area B in Fig. 2). In general terms we can confirm that while M-AMBI and ISS convey similar information, BITS does not. In Fig. 2 points/records within areas A and B correspond to discrepant values for M-AMBI and BITS and for BITS and ISS respectively. Due to the stochastic relation among the three multimetric indices, their use in association with standard boundaries can lead to contrasting ecosystem classifications. In Fig. 2 areas in red correspond to both indices assigning the same ecologic class to the sites, but the majority of observed records fall outside these areas. It is very evident that the "high" BITS category

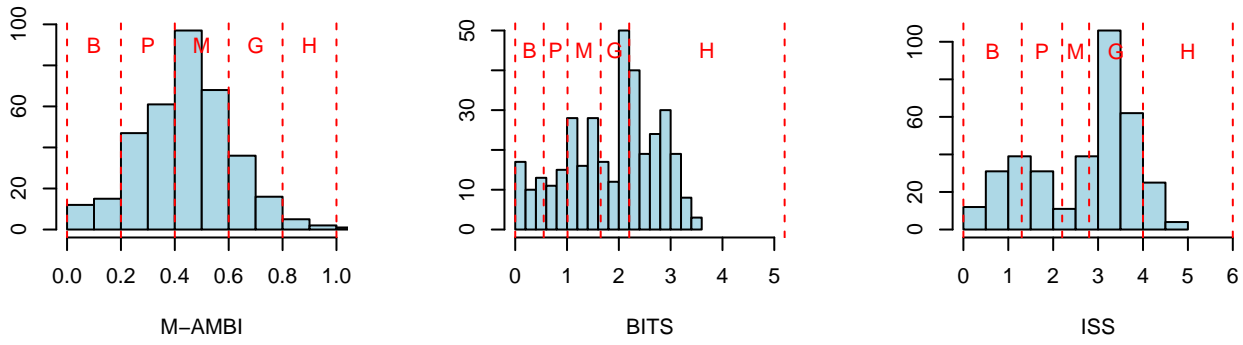


Fig. 1 Distributions of the biotic multimetric indices with standard classification boundaries.

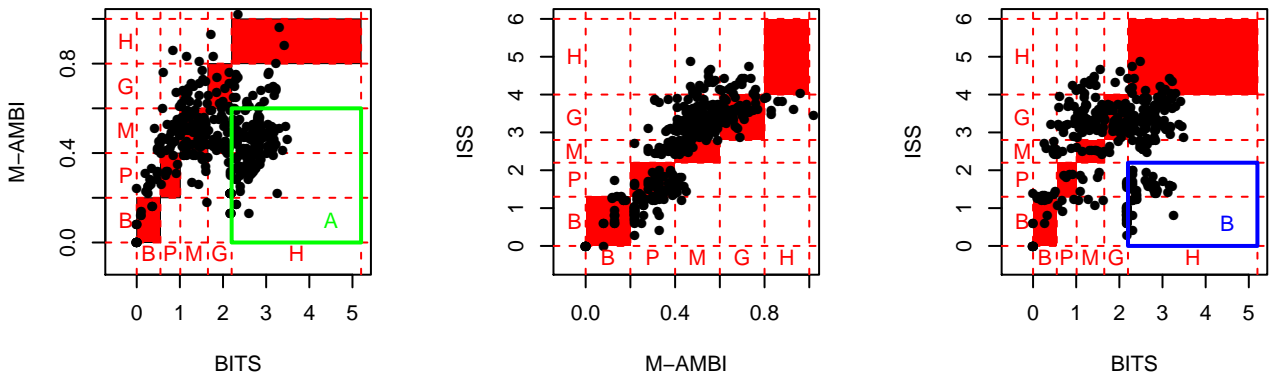


Fig. 2 Joint distributions of pairs of multimetric indices with standard classification boundaries. Areas A and B contain clusters of observation with contrasting behavior of two indices.

corresponds to a very variable behavior of both M-AMBI and ISS. Exploratory attempts to uncover abiotic features of lagoons producing ambiguous ecosystem quality assessments with the three indices only provided a weak evidence for a direct relation of M-AMBI and ISS with both oxygen content and water pH that is not observed for BITS (full results of the exploratory data analysis are available from the authors upon request).

2.3 Joint model for lagoon ecological quality indices

The possible contrasting behaviour of the three multimetric indices is analyzed in the context of Bayesian hierarchical models. As the indices have different theoretical ranges (see section 2.2), in order to compare the influence of each explanatory variable the three response variables were standardized. Explanatory variables were also standardized in order to avoid dimensional issues and focus on the gradients of the three responses along their joint variation. Standardization of both response and explanatory variables leads to a model that lends itself to analyze the sensitivity of the multimetric quality indices to water parameters and anthropogenic pressures. Standardized values of the indices are jointly modeled as functions of water parameters, expert's

opinion on external anthropogenic pressures and lagoon effects. Here the following notation is adopted: $r = 1, \dots, R$ is a replicate of a biotic record taken at time $t = 1, \dots, T$ and location $s = 1, \dots, n_l$ within lagoon $l = 1, \dots, L$. Specifically we have $R = 3$ replicates, $T = 8$ time points, $L = 3$ lagoons with $n_l = 6, 6, 3$ monitoring stations respectively for Lesina, Varano and Alimini. The response vector \mathbf{y}_{rtsl} contains standardized values of the three multimetric indices M-AMBI, BITS and ISS for each replicate, while the P -dimensional vector \mathbf{x}_{rtsl} contains standardized values of P explanatory variables for each monitoring station and every time point. Vectors \mathbf{y}_{rtsl} and \mathbf{x}_{rtsl} are related by the following expression:

$$\mathbf{y}_{rtsl} = \mathbf{1}_3 \alpha + \mathbf{B} \mathbf{x}_{rtsl} + \mathbf{1}_3 w_l + \boldsymbol{\varepsilon}_{rtsl} \quad (1)$$

where α is the intercept term, \mathbf{B} is a $3 \times P$ matrix of index-specific regression coefficients that measure the effect of each abiotic covariate on every multimetric index, $\mathbf{1}_3$ is a unit vector, w_l is a random lagoon effect ($l = 1, \dots, L$) and $\boldsymbol{\varepsilon}_{rtsl}$ is a 3-dimensional correlated random error vector. At the second hierarchical level independent fixed and random

effects are specified as follows:

$$\begin{aligned}\alpha &\sim N(0, 0.01) \\ b_{jp} &\sim N(0, 0.01) \\ w_l &\sim N(\mu_w, \tau_w) \\ \boldsymbol{\epsilon}_{rtsl} &\sim N_3(\mathbf{0}, T_{\boldsymbol{\epsilon}})\end{aligned}\quad (2)$$

where $j = 1, 2, 3$ for M-AMBI, BITS and ISS and the correlation between the three indices is accounted for by the unstructured precision matrix $T_{\boldsymbol{\epsilon}}$. Here we adopt a Bayesian approach and estimate model parameters by a Markov chain Monte Carlo (MCMC) posterior simulation algorithm. Then a third hierarchical level is specified for random effects hyperparameters:

$$\begin{aligned}\mu_w &\sim N(0, 0.01) \\ \tau_w &\sim \text{Gamma}(0.01, 0.01) \\ T_{\boldsymbol{\epsilon}} &\sim \text{Wishart}(\mathbf{I}, 5)\end{aligned}\quad (3)$$

where \mathbf{I} is the identity matrix.

Notice that the three lagoons are geographically well separated and most likely independent and the small number of monitoring stations within each lagoon (3 to 6) does not allow the estimation of lagoon-specific spatial covariances. Simple forms of temporal autocorrelation or seasonality are even more hardly detectable with only 8 time points and exploratory data analysis didn't show any kind of longitudinal trend. Then only random effects of the lagoons are considered and no spatial and time effects are taken into account in the model specification.

Posterior simulations were implemented using the JAGS software (Plummer, 2003) with 2 chains with different starting points (the code can be obtained from the authors upon request). For each chain, we allowed 250000 iterations with 100000 samples of burn-in and used the remaining 150000 iterations thinned by 20 for posterior estimation. Chain convergence was ascertained by visual inspection of standard convergence diagnostic tools, such as trace plots and autocorrelation plots. Notice that the conjugate nature of this multivariate response model makes Gibbs sampling fully efficient and simulations running in real-time on a standard personal computer.

2.4 Model assessment and sensitivity analysis

Once the model was estimated, we quantified the influence of each explanatory variable on the three indices jointly. Most traditional methods produce influence coefficients for each predictor-output combination, usually not allowing joint evaluation. Typically when response and explanatory variables are standardized we compare the absolute values of model coefficient estimates i.e. $|\hat{b}_{jp}| \geq |\hat{b}_{jp'}|$ where $j = 1, 2, 3$

for MAMBI, BITS and ISS and $p, p' = 1, \dots, P$ refer to explanatory variables. Here to measure the joint sensitivity of response to explanatory variables we compare the norm of estimated coefficient vectors $\hat{\mathbf{b}}_p$ accounting for the residual correlation between response variables (see Brynjarsdóttir and Gelfand, 2013, for details):

$$I_p = \hat{\mathbf{b}}_p^T \hat{T}_{\boldsymbol{\epsilon}}^{-1} \hat{\mathbf{b}}_p \quad (4)$$

where \mathbf{b}_p is the p -th column of \mathbf{B} with $p = 1, \dots, P$. If $I_p > I_{p'}$ then the three indices are jointly more sensitive to (or influenced by) the p -th explanatory variable than to the p' -th. Notice that in the Bayesian MCMC approach samples from the posterior distribution of (4) are available for inferential purposes. Furthermore in Brynjarsdóttir and Gelfand (2013) it is shown that if a prior on the model inputs \mathbf{X} is given, comparison of $P > 2$ inputs using the I_p criterion gives the same results as comparison of marginal conditional variances of the inverse predictive distributions of the inputs, in the spirit of Clark *et al.* (2013).

In order to check the model distributional assumptions and evaluate its goodness-of-fit we consider some posterior predictive approaches based on summaries of the density $f(y^{rep}|y)$. Predictions y^{rep} reflect the expected observations after replicating our experiment in the future, having already observed y and assuming that the model is true. If the model is appropriate for describing the observed data, then vectors y and y^{rep} will be close. Hence a comparison of these two vectors provides information concerning the model fit. We consider summary statistics $D(y, \theta)$ which play the role of tests for checking model assumptions and measure discrepancies between the data and the model (Gelman *et al.*, 1996). Assessment of the posterior distributions of $D(y, \theta)$ and $D(y^{rep}, \theta)$ provides individual and overall goodness-of-fit measures that can be summarized graphically or using tail area probabilities called *posterior predictive p-values* (Meng, 1994) defined as

$$PPPV = P(D(y^{rep}, \theta) > D(y, \theta)|y) \quad (5)$$

Since *PPPV*'s are nothing but probabilities, values around 0.5 indicate that the distribution of the replicated and observed data are close, while values around 0 or 1 indicate a severe difference between them. When the summary statistic $D(y, \theta)$ is ancillary, the distribution of the *PPPV*'s is almost uniform (Gelman *et al.*, 1996). In the following we use tail area probabilities in order to assess the model distributional assumption of normality and to evaluate the model ability to predict some aspects of interest for our application.

Residual analysis is used to check the model structural assumptions of normality and homoscedasticity, evaluating the deviations of the predicted residuals with respect to their observed counterparts. Residuals are defined as

$$r_i = y_i - E(Y_i|\theta) \quad (6)$$

while r_i^{rep} are residuals based on predictive values y_i^{rep} . In order to evaluate the discrepancies between observed and predicted values under the fitted model, we will use the following posterior predictive p -values

$$PPPV_i = P(r_i^{rep} > r_i|y) = P(y_i^{rep} > y_i|y) \quad (7)$$

Since the residuals are ancillary statistics, we expect that the $PPPV$'s follow a uniform distribution.

3 Results

While the posterior estimate of the common intercept parameter α obtained by the MCMC simulations from the multivariate continuous response model has value -0.499 with strictly negative 95% credibility interval $(-0.80, -0.198)$, table 1 shows the posterior estimates of index-specific regression coefficients and the corresponding 95% credibility intervals. Abiotic indices and expert's opinion evaluations of four external anthropogenic pressures (due to the presence of agricultural inputs, domestic discharges, industrial discharges and fisheries) have been considered as explanatory variables. Estimates corresponding to strictly positive or negative 95% credibility intervals are highlighted in red. First we report about overall effects shared by the three multimetric quality indices. pH and the severity of domestic discharges positively influence the average value of the three indices. On the other hand the average value of the indices decreases when the pressure of industrial discharges and the amount of chlorophyll increase. The expected effects of agricultural inputs and of salinity are probably masked by the presence of other informative predictors (as the pressure of domestic discharges and the amount of dissolved oxygen). Domestic and diffuse agricultural discharges, as far as mainly characterized by organic and inorganic inputs, are likely to have idiosyncratic impacts at the community and ecosystem levels in lagoon ecosystems, being strongly affected by the overall abiotic context (e.g., oxygen concentration and hydrodynamics) and differing among biological quality elements.

Turning to effects that differentiate the behavior of the BITS index with respect to M-AMBI and ISS, we have that the pressure of fisheries has a negative influence on the average value of M-AMBI and ISS but does not affect BITS which is inversely proportional to the amount of dissolved oxygen. On the other hand, the amount of dissolved oxygen positively affects M-AMBI and ISS. This is probably due to an increasing oxygen content causing an increase in the number of species and in the size structure, but reducing the number of families. Water temperature negatively influences the BITS index but it is not relevant for the other two indices. The sign and significance of the effects of pressures and abiotic variables are all supported by sound ecological knowledge and have plausible interpretation.

Figure 3 shows the posterior distribution of the correlation coefficients between the three indices: the model confirms a known higher correlation between M-AMBI and ISS and a lower correlation of these two with BITS. The model provides estimates of lagoon effects corresponding to higher values of the three indices at Alimini and Lesina, possibly due to Varano being deeper and more eutrophic. Considering the lagoon effects within the model improves the overall fit in terms of DIC (the chosen model DIC is 488.699 while the DIC of an alternative model without lagoon-specific random effects is 498.300).

In order to quantify the influence of explanatory variables on the three indices jointly, we computed the influence measures in Equation (4). Figure 4 shows the 95% credibility intervals of these influence measures. Agriculture is the least influential variable while the pressure due to the fishery activity is the most relevant. Oxygen and Chlorophyll have very similar effects on the joint sensitivity of the three indices: $I_{oxy} = 0.135$ (0.050, 0.267), $I_{clo} = 0.124$ (0.058, 0.216). The influence measures are in substantial agreement with the marginal estimates in Table 1 in terms of both the sign and the absolute value of the estimates.

Model distributional assumptions and goodness-of-fit have been checked using posterior predictive p -values. First, in order to evaluate the discrepancies between observed and predicted values under the fitted model, we used the posterior predictive p -values defined in Equation (7). Since the residuals are ancillary statistics, we expect that the $PPPV$'s follow a uniform distribution. Figure 5 shows the posterior predictive p -values computed for each multimetric index with the proposed model: the three distributions are acceptably uniform and their averages are all around the desired value of 0.5, suggesting a good fit of the proposed model to the available data.

In order to evaluate the normality assumption, we check that the posterior distribution of the residuals is approximately normal with zero mean. Figure 6 shows the distributions of the posterior MC means of the residuals for each index and the 95% credibility intervals for their mean: all intervals contain the value zero. The Shapiro-Wilk test does not lead to reject the assumption of normality of the residuals for the three indices, with p -values respectively equal to 0.724, 0.191, 0.751 for M-AMBI, BITS and ISS.

The assumption that the error precision matrix $T_{\boldsymbol{\epsilon}}$ is unstructured and equal between lagoons was assessed comparing the proposed model with an alternative one with lagoon-specific precision matrices in (2) and (3), i.e. modeling the error term in Equation (1) as $\boldsymbol{\epsilon}_{rstl} \sim N_3(\mathbf{0}, T_{l,\boldsymbol{\epsilon}})$, with $T_{l,\boldsymbol{\epsilon}} = \tau_{l,\boldsymbol{\epsilon}} \mathbf{I}$ where \mathbf{I} is a 3×3 unit diagonal matrix. The chosen model DIC is 488.699 while the DIC of the alternative model is 771.125.

Once model assumptions are validated, we use posterior predictive p -values to evaluate the model ability to predict

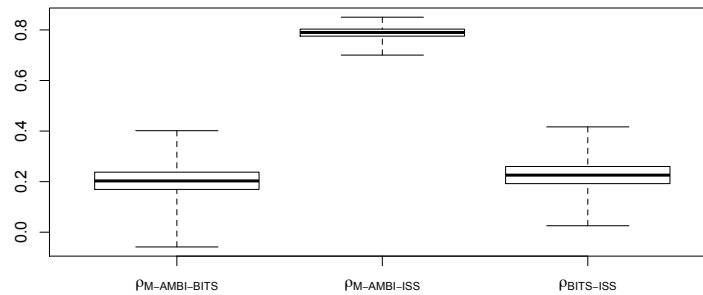


Fig. 3 Boxplots of the estimated posterior distribution of the correlation between indices.

Table 1 Posterior estimates of index-specific regression coefficients and 95% credibility intervals for model (1). Estimates corresponding to strictly positive or negative credibility intervals are highlighted in red.

	M-AMBI		BITS		ISS	
Water temperature	-0.028	(-0.195, 0.141)	-0.217	(-0.396, -0.041)	-0.038	(-0.203, 0.130)
Dissolved oxygen	0.149	(0.020, 0.280)	-0.220	(-0.356, -0.085)	0.167	(0.039, 0.298)
Salinity	-0.017	(-0.219, 0.196)	0.148	(-0.061, 0.370)	-0.025	(-0.228, 0.190)
pH	0.238	(0.136, 0.337)	0.160	(0.057, 0.264)	0.217	(0.119, 0.318)
Chlorophyll	-0.207	(-0.303, -0.113)	-0.259	(-0.363, -0.157)	-0.211	(-0.305, -0.114)
Agricultural inputs	-0.000	(-0.134, 0.134)	-0.028	(-0.166, 0.109)	0.069	(-0.064, 0.200)
Domestic discharges	0.258	(0.131, 0.384)	0.225	(0.092, 0.357)	0.252	(0.128, 0.375)
Industrial discharges	-0.313	(-0.458, -0.168)	-0.253	(-0.405, -0.102)	-0.215	(-0.356, -0.069)
Fisheries	-0.438	(-0.645, -0.235)	-0.093	(-0.300, 0.115)	-0.433	(-0.635, -0.232)

relevant aspects of the process under investigation. Among the main goals of our model is investigating the variability of the multimetric indices and the correlation between them. Hence we define *PPPV*'s using the variance of each index and the correlation between pairs of indices as the summary function D in (5). The left panel of Figure 6 shows the distributions of the posterior predicted variances of each multimetric index with the observed variances corresponding to the red lines. M-AMBI and ISS have posterior predicted variances almost centered around their observed counterpart and the right tail area probabilities, corresponding to the *PPPV*'s are around 0.5. On the other hand the variance of BITS is slightly overestimated by the model, corresponding to a tail area probability equal to 0.736, though this behavior is not to be considered so extreme as to worry about the capability of the model to capture the variability of the data. The right panel of Figure 6 shows the posterior predicted correlations between pairs of indices: all predictive distributions are centered around the observed correlations and all the tail area probabilities are around 0.5 (0.511, 0.431, 0.534 respectively for M-AMBI, BITS and ISS).

4 Concluding remarks and directions for future work

In this paper we compare the performance of three benthic macroinvertebrates-based multimetric indices (M-AMBI, BITS and ISS) in describing the ecological status of three lagoons in Apulia. We investigate the joint behavior of the indices and evaluate their different responses to sources of stress and natural variability, in order to highlight inconsistencies among multimetric indices in assessing the ecological status. A Bayesian hierarchical model has been used as a natural parametrization of the ecological process. The indices have been jointly modeled as function of abiotic covariates and indicators of external anthropogenic pressures. The proposed machinery is at the same time easily understood by practitioners and very rich in terms of interpretative tools that are obtained as by product for model assessment and sensitivity analysis.

A comparative analysis of the responses of the three indices to different stressors supports the original assumption of the paper that different multi-metric indices have differentiated responses to external stressors. However, the close agreement between M-AMBI and ISS responses to the tested natural and anthropogenic sources of variation strongly sup-

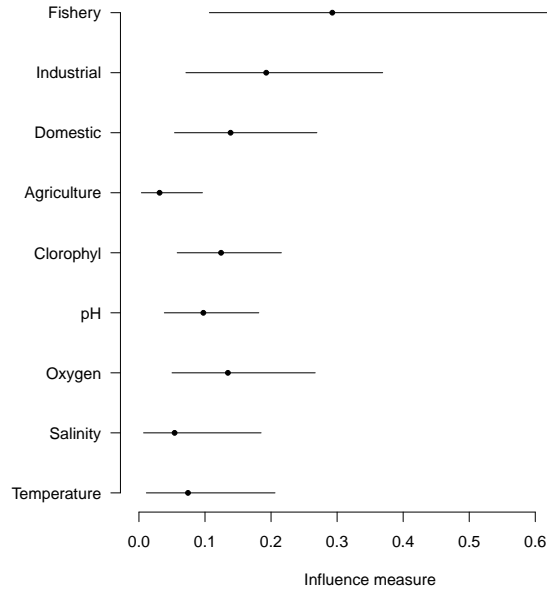


Fig. 4 Posterior credibility intervals of the influence measure for explanatory variables.

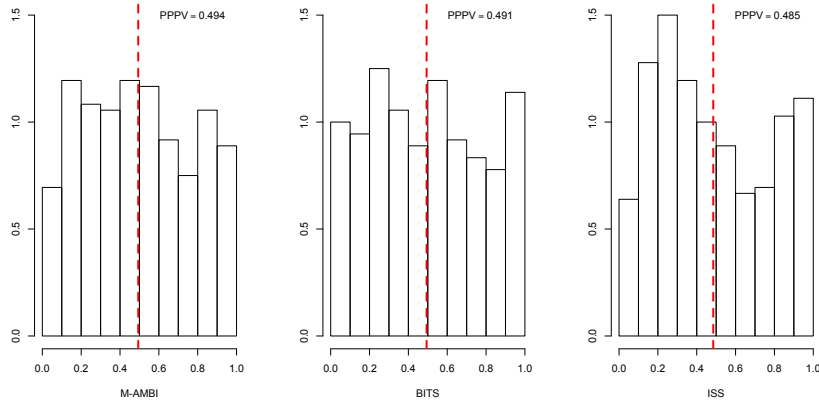


Fig. 5 Posterior predictive p -values of the residuals for the three indices. The red line highlights the mean of the posterior predictive p -values.

ports the accuracy of the two metrics in evaluating lagoon ecological status. In fact, these two multi-metric indices show significant and coherent responses to the same group of tested sources of variation, all of them related to anthropogenic pressures, even though they rely on different types of data, i.e., taxonomic-based data the M-AMBI and body-size based data the ISS. The close agreement, even though not perfect, observed between M-AMBI and ISS and the independence of the data sets on which they are based suggest that the accuracy in the ecological status evaluation might be increased by integrating the indices in a joint Bayesian classification model as proposed in (Arima *et al.*, 2013). On the other hand, the responses of BITS to natural stress agents do not

fully support its integration in the model and suggests possible refinements of the model based classification integrating multi-metric indices for the evaluation of the ecological status of lagoon ecosystems. Future work will be devoted to plan deeper studies, possibly involving non-aggregated abundances, defining a sampling scheme that allows to study the effect of spatial variation inside lagoons and to evaluate their evolution in time. For such a larger database the proposed model could easily be extended including a spatio-temporal term $w_{t,sl}$ with $\mathbf{w}_l \sim N_{Tn_l}(\boldsymbol{\mu}_{\mathbf{w},l}, \tau_{\mathbf{w},l}^2 \mathbf{H}(\phi_l)^{-1})$, where $\mathbf{w}_l = (w_{11l}, \dots, w_{Tn_l l})'$. Here $\tau_{\mathbf{w},l}^{-2}$ and $\mathbf{H}(\phi_l)$ are the process variance component and space-time structured correlation matrix at the l -th lagoon, $l = 1, \dots, L$. This model, allowing

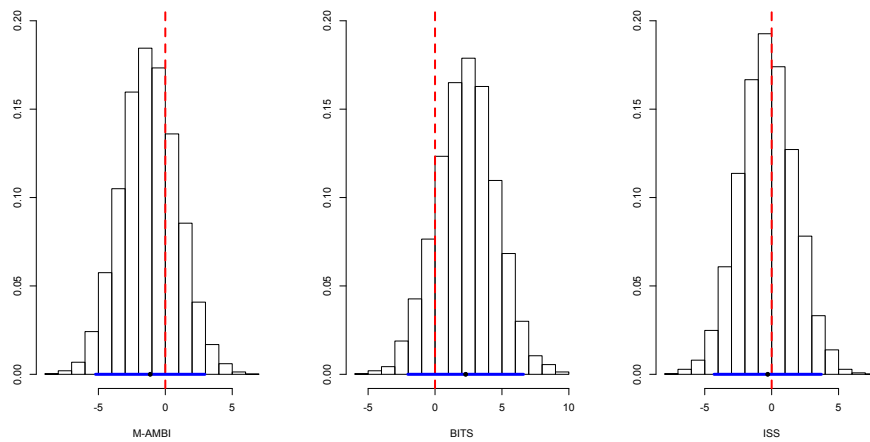


Fig. 6 Posterior means of the residuals for each index: blue segments correspond their 95% credibility intervals and red lines correspond to the zero value. The bullets highlight the overall posterior means of the residuals.

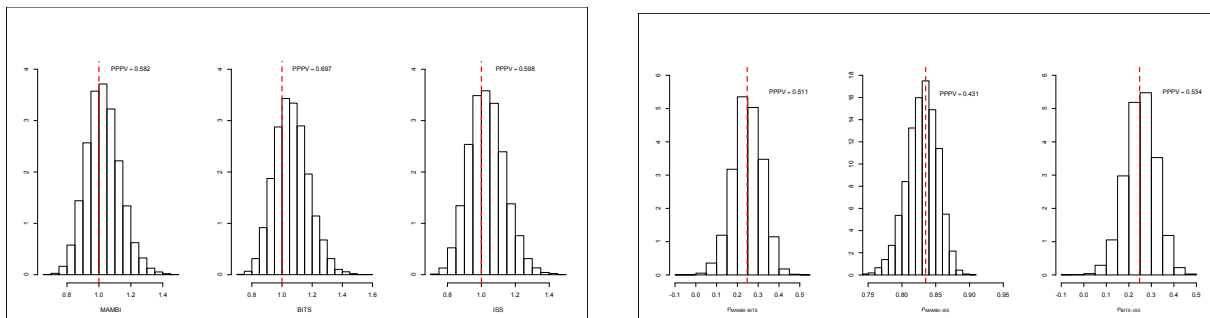


Fig. 7 Left panel: predicted and observed variances of each index. Red lines correspond to the observed variances. Right panel: predicted and observed correlations between indices. Red lines correspond to the observed correlations.

some insight on the spatial dependence within lagoons, may help to clarify the role of local effects, currently completely ignored, in determining multimetric indices and subsequent lagoon classification.

References

- Arima, S., Basset, A., Jona Lasinio, G., Pollice, A., and Rosati, I. (2013). A hierarchical bayesian model for the ecological status classification of lagoons. *Ecological Modelling*, **263**, 187–195.
- Barbone, E., Rosati, I., Reizopoulou, S., and Basset, A. (2012). Linking classification boundaries to sources of natural variability in transitional waters: a case study of benthic macroinvertebrates. *Ecological Indicators*, **12**, 105–122.
- Basset, A., Sabetta, L., Fonnesu, A., Mouillot, D., Do Chi, T., Viaroli, P., Reizopoulou, S., and Carrada, G. (2006). Typology in mediterranean transitional waters: new challenges and perspectives. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **16**, 441–445.
- Basset, A., Barbone, E., Borja, A., Brucet, S., Pinna, M., Quintana, X., Reizopoulou, S., Rosati, I., and Simboura, N. (2012). A benthic macroinvertebrate size spectra index for implementing the water framework directive in coastal lagoons in mediterranean and black sea. *Ecological Indicators*, **1240**, 72–83.
- Basset, A., Barbone, E., Elliott, M., Li, B., Jorgensen, S., Lucena-Moya, P., Pardo, I., and Mouillot, D. (2013). A unifying approach to understanding transitional waters: Fundamental properties emerging from ecotone ecosystems. *Estuarine, Coastal and Shelf Science*, **132**, 5–16.
- Borja, A., Franco, J., and Perez, 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**, 110–114.
- Borja, A., Franco, J., Valencia, V., Bald, J., Muxika, I., Belzunce, M., and Solaun, O. (2004). Implementation of the european water framework directive from the basque country (northern spain): a methodological ap-

- proach. *Marine Pollution Bulletin*, **48**, 209–218.
- Borja, A., Barbone, E., Basset, A., Borgersen, G., Brkljacic, M., Garmendia, J., Marques, J., Mazik, K., Muxika, I., Neto, J., Norling, K., Rodriguez, J., Rosati, I., Rygg, B., Teixeira, H., and Trayanova, A. (2011). Response of single benthic metrics and multi-metric methods to anthropogenic pressure gradients, in five distinct european coastal and transitional ecosystems. *Marine Pollution Bulletin*, **62**, 499–513.
- Brynjarsdóttir, A. and Gelfand, A. (2013). Collective sensitivity analysis for ecological regression models with multivariate response. *Submitted*, pages 1–37.
- Clark, J. S., Bell, D. M., Kwit, M., Powell, A., and Zhu, K. (2013). Dynamic inverse prediction and sensitivity analysis with high-dimensional responses: Application to climate-change vulnerability of biodiversity. *Journal of Agricultural, Biological, and Environmental Statistics*, **Accepted**.
- Gelman, A., Meng, X. L., and Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, **6**, 733–807.
- Ji, J.-H. and Chang, N.-B. (2005). Risk assessment for optimal freshwater inflow in response to sustainability indicators in semi-arid coastal bay. *Stochastic Environmental Research and Risk Assessment*, **19**, 111–124.
- Lucena-Moya, P., Pardo, I., and Alvarez, M. (2009). Development of a typology of transitional waters in mediterranean ecoregion: the case of the islands. estuarine, coastal and shelf. *Science*, **82**, 61–72.
- Marchini, A. and Marchini, C. (2006). A fuzzy logic model to recognise ecological sectors in the lagoon of venice based on the benthic community. *Ecological Modelling*, **193**, 105–118.
- Meng, X. L. (1994). Posterior predictive p-values. *The Annals of Statistics*, **22**(3), 1142–1160.
- Mistri, M. and Munari, C. (2008). Bits: a smart indicator for self-bottom, non-tidal lagoon. *Marine Pollution Bulletin*, **56**, 597–599.
- Muxika, I., Borja, A., and 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.
- Piffady, J., Parent, E., and Souchon, Y. (2013). A hierarchical generalized linear model with variable selection: studying the response of a representative fish assemblage for large european rivers in a multi-pressure context. *Stochastic Environmental Research and Risk Assessment*, **27**, 1719–1734.
- Plummer, M. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20–22, Vienna, AUSTRIA.
- Sadiq, R. and Tesfamariam, S. (2008). Developing environmental indices using fuzzy numbers ordered weighted averaging (fn-owa) operators. *Stochastic Environmental Research and Risk Assessment*, **22**, 495–505.
- Su, S., Zhang, Z., Xiao, R., Jiang, Z., Chen, T., Zhang, L., and Wu, J. (2012). Geospatial assessment of agroecosystem health: development of an integrated index based on catastrophe theory. *Stochastic Environmental Research and Risk Assessment*, **26**, 321–334.
- Viaroli, P., Mistri, M., Trousselier, M., Gierzoni, S., and Cardoso, A. (2005). Structure, functions and ecosystem alterations in southern european coastal lagoons. *Hydrobiologia*, **550**, 57–71.