Development of macroinvertebrate multimetric indices for a heavily modified water body using quantile regression

Canobbio S.[1], Cabrini R.[1], Fornaroli R.[1], Sartori L.[1], Piana A.[2], Mezzanotte V.[1], Azzellino A.[3]

[1]DISAT, Università degli Studi di Milano-Bicocca, Piazza della Scienza 1, 20126 Milano (IT)
[2]ARPA Lombardia, Via I. Rosellini 17, 20124 Milano (IT)
[3]DIIAR, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milano (IT)

Introduction

A critical factor for the better comprehension of Heavily Modified Water Bodies (HMWBs) and for decision-making in restoration programs is the individuation of which factors **set limits** to biological community development. HMWB ecosystems suffer from the severe effects of multiple stressors. In these conditions it is hard to assess **causal relationships** among specific stressors and responses of biological communities using the most commonly used statistical tools. Usually, hypotheses about the central response of organism abundance and/or richness to environmental gradients are tested, although the effects of other stressors, and even data stochasticity, may also influence such response and decrease the fit of the model, which can become uninformative. In this perspective, **quantile regression** (Cade & Noon, 2003) enables the various stressors to be considered as "constraints" to the distribution of biological communities (fig. 1), specifically dealing with **high data variability** (Downes, 2010).

In our study we analysed over 220 samples of macroinvertebrate assemblages and environmental variables coming from a ten-year long survey in the HMWBs of the Lambro-Seveso-Olona basin, one of the most densely populated watersheds of Europe, in the conurbation area of Milan, Italy (fig. 2).

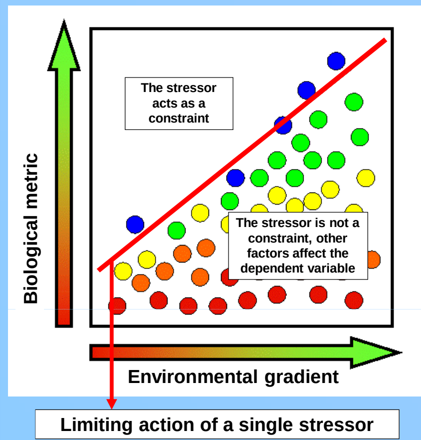
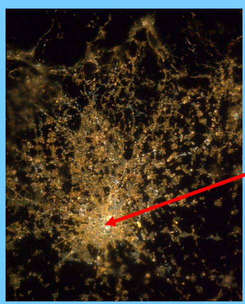


Figure 1 - Estimates of extreme quantiles allow to detect constraints along gradients.



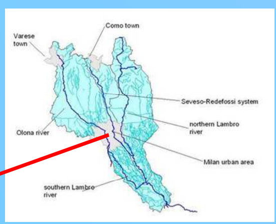


Figure 2 – The Lambro-Seveso-Olona basin and the sprawling conurbation of Milan (courtesy Paolo Nespoli – ESA/NASA).

Results

Relationships and patterns among the collected environmental variables were analyzed using a Principal Component Analysis (PCA). The water quality parameters and the morphological indicators basically cluster in two different groups, identified by the first two components explaining about 55% of the total variance. The factor scores of the first two principal components were subsequently used as new variables. The first one represents the gradient of hydromorphological integrity (*habitat gradient*). The second one represents the overall water quality gradient (*pollution gradient*). Due to the mathematical properties of principal components, these are gradients that maximize variation and are independent from each other.

Four screening criteria, plus a fifth qualitative criterion, were used to screen 53 biological metrics and their response to stressor gradients. These criteria were adapted from Purcell *et al.* (2009) and was used for the elimination of both non-informative and redundant metrics (table 1).

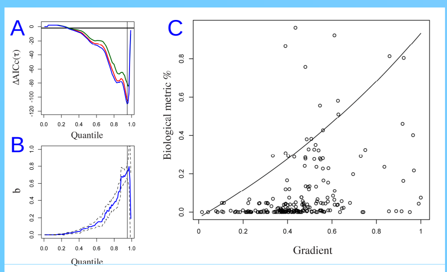


Figure 3 - In panel A $\Delta AIC_c(\tau)$ for each model for all τ are shown. In this case, the exponential model presents the minimum value of $\Delta AIC_c(\tau)$ at $\tau = 0.93$ (vertical black line). Simultaneously, at the same τ, the confidence interval of b parameter for exponential model gets its minimum (panel B); for larger quantiles confidence intervals are wider and can included zero. In the panel C, the selected 0.93 exponential quantile regression (black line curve) was superimposed to a scatterplot between a biological metric and a gradient..

After the screening, two biological metrics (relative abundance of predators and Habitat FFG) showed a clear response to habitat gradient as a limiting factor (Fig. 4). Six biological metrics (relative abundance of *Baetis spp.*, Oligochaeta and Predators, Clinger richness, Family richness, Shannon Family level) showed non redundant response to the water quality gradient as a constraint (Fig. 5). Multimetric biological indices for the studied area were developed on the basis of the relationships previously identified. The scaled metrics can be useful to create basin-specific multimetric indices to evaluate complex situations such as those of HMWBs and the role of single stressors.

Table 1 – Criteria for the selection of informative biological metrics

1	Range of relative abundance metrics must be > 10% and range of richness must be > 5 in the whole dataset.
2	Area-based effects examined using linear regression, and significantly related metrics discarded.
3	The relationships of each biological metric to the two environmental gradients have been examined using quantile regression criteria (fig. 3). Models (linear, logarithmic or exponential) were selected for each biological metric using the Akaine Information Criterion corrected for small sample size (AICc) and the evaluation of the model parameters (Appendix C in Cade *et al*., 2005).
4	Redundancy tested (Pearson correlation). From a group of redundant metrics only the metric showing the best relationship with the environmental gradients according to point (3) was considered.
5	Reconsideration of the eliminated metrics (point 4) if showing ecological importance.

$$\text{Index} = \frac{\sum_{i=1}^{n} \text{ScaledMetric}_i (0-1)}{n}$$

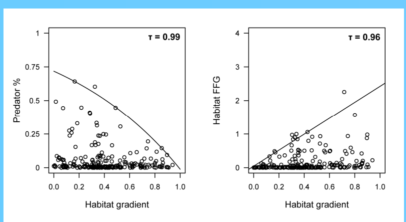


Figure 4 - Scatterplots of invertebrate metrics against habitat gradient.

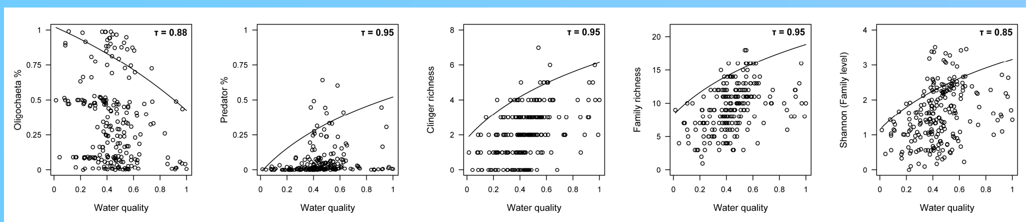


Figure 5 - Scatterplots of invertebrate metrics against water quality gradient.

Conclusions

Basin-specific indices, based on biological metrics that can highlight the role of stressors as constraints, can be helpful to **disentagle the sources of data variability** in HMWBs.

Quantile regression, applied to extreme quantiles of data distribution, allows to analyse how a specific stressor influences the biological communities and, thus, **how a specific restoration effort can potentially increase the ecosystem quality**, reducing or removing the constraint, to a settable **biological potential**. This information can be used for management purposes, and could allow to set pragmatic restoration goals.

References

Cade B.S. & Noon B.R. (2003) A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1, 412-420.

Downes B.J. (2010) Back to the future: little-used tools and principles of scientific inference can help disentangle effects of multiple stressors on freshwater ecosystems. Freshwater Biology 55, 60-79.

Cade B.S., Noon B.R. & Flather C.H. (2005) Quantile regression reveals hidden bias and uncertainty in habitat models. Ecology 86, 786–800.

Purcell A.H., Bressler D.W., Paul M.J., Barbour M.T., Rankin E.T., Carter J.L., et al. (2009) Assessment Tools for Urban Catchments: Developing Biological Indicators Based on Benthic Macroinvertebrates. Journal of the American Water Resources Association 45, 306-319.