A general framework for indicator design and use with application to the assessment of coastal water quality and marine protected area management

Benoît Beliaeff[a, *] and Dominique Pelletier[b]

[a] Ifremer, 101 Promenade R. Laroque, Centre IRD, B.P. 2059, 98846 Nouméa Cedex, New Caledonia
[b] Ifremer, Centre de Brest, Technopôle Iroise, B.P. 70, F-29280 Plouzané, France

[*] Corresponding author : Benoît Beliaeff, Tel.: +687 352581; fax: +687 351177, email address :
benoit.beliaeff@ifremer.fr

Abstract:

Environmental management decisions based upon indicators are the end point of a process involving stakeholders and scientists. These steps should be explicit and follow a chronology. This paper presents a general framework for the design and use of management-oriented indicators, integrating management questions and performance criteria. We first examined the desirable characteristics of indicators aimed at providing decision-support for marine environmental management. Ideally, one should select the indicator that guarantees a safe and unambiguous decision leading to the appropriate measures in terms of regulation, remediation or control. In the present study, indicators are assessed according to two criteria: relevance and effectiveness. Relevance encompasses sensitivity and the existence of quantitative reference values, thereby allowing the selection of potential indicators. Effectiveness is the ability of the indicator to reach its predefined targets based on optimal (or at least improved) data collection protocols. The framework is illustrated by applying it to the European Water Framework Directive and to the Marine Protected Area management contexts.

Research highlights

► We propose a general framework for the design and use of management-oriented indicators. ► Relevance and efficiency are the two indicator performance criteria to be assessed successively. ► A clear linkage is established between management objectives and indicator performance criteria. ► Stakeholders-scientists dialogue is crucial for optimal indicator development and use.

1.   **Introduction**

In the context of marine environmental management, a growing demand exists to develop indicators for increasing threats and actual or potential conflicts of uses. Decisions that have to be taken by managers at various scales should rely on clear information conveyed by appropriate indicators. The implementation of the European Water Framework Directive (WFD) [1] offers an example in which a set of indicators needs to be defined to classify water bodies with the general aim of reaching a good status by 2015 for all continental and coastal bodies of water in Europe. These requirements pointed out to managers and scientists that very few indicators were available to assess the status of the coastal environment and led to the rapid development of *ad hoc* indicators addressing a range of environmental and biological features. For coastal waters, phytoplankton, benthic flora - macroalgae and angiosperms - and invertebrate fauna should be monitored using indicators. A number of related scientific articles exist, in particular on benthic invertebrates [2,3,4,5,6,7], plankton [8,9,10] or seagrass (e.g. [11]), that support the need for this type of monitoring. The purpose for the use of indicators is quite clear within the WFD framework; it should allow one to certify that good status has been reached or to conversely decide upon adequate management measures to reduce the pressure leading to the observed impact. However, in practice, an indicator performs well if it guides one toward the appropriate decision while minimising the risk of error. Although Carstensen [12] showed the importance of considering statistical inference in such a framework, this is rarely taken into account while implementing the WFD. More generally, no investigation exists for indicator performance. Specifically, no investigation is available for indicators of Marine Protected Area (MPA) effectiveness, another current issue for marine environment management [13]. MPAs are increasingly used for managing coastal ecosystems. On the scientific side, a wealth of papers have been published dealing with the assessment of MPA performance, in particular for biodiversity conservation, but none of the studies aimed at defining operational indicators for decision-support systems [14]. From a more management-oriented standpoint, international initiatives aimed at proposing indicators of MPA management effectiveness for MPA managers do not provide guidance about indicator performance, particularly regarding their statistical properties [15].

The Pressure-State-Response (PSR) or DPSIR (Driving forces-Pressure-State-Impact-Response) frameworks [16] are conceptual frameworks for indicator definition and have been widely described and used in environmental assessments, including MPA management [e.g., 17,18] For Mangi et al. [17], they are useful tools in promoting an interaction between social

systems and environmental variability. Most management issues can be framed within these general models, including WFD and MPA management. However, although the Organisation for Economic Co-operation and Development (OECD) [16] proposed criteria for indicator selection, it does not intend to provide insight into the intrinsic properties of indicators. Several papers proposed a list of performance criteria for environmental or ecological indicators [19,20,21,22,23] (see also a review in [24]). Several of these papers were in the field of fisheries management [25,26,27,28,29]. However, these papers did not provide a hierarchy among this list of performance criteria. We feel that indicators would be more operational and, thus, useful if their selection occurred after a logical suite of steps. Ideally, these steps would include explicit prioritisation of performance criteria. Fontalvo-Herazo et al. [30] described and applied a method for the design of a participative indicator system for local coastal management, borrowing from a methodology widely used in forestry management. Through the MESMIS framework, López-Ridaura et al. [31] proposed a method to derive a set of strategic indicators from a thorough diagnosis. However, these authors did not go into detail on the performance of indicators. During our state-of-the-art study, we did not find a paper integrating both aspects: proposing a general framework for indicator use and discussing their properties in detail within this framework.

Aimed at filling this gap, we first examined what should be the desirable characteristics for indicators aimed at decision-support for marine environmental management. We then proposed a general framework for the definition and use of management-oriented indicators. This framework included the identification of potential indicators and the assessment of their performance with respect to managers' needs. Finally, the approach was illustrated using applications to (i) coastal water quality within the WFD context and (ii) MPA management.

2. Indicators for environmental management: definitions and performance criteria

2.1. Definitions

Before presenting the methods for designing indicators, several terms have to be defined. By indicator, we mean a function of observations or of the outputs of a model that indicates the present state and/or dynamics of the system of interest [32] in relation to scientific questions or management objectives. The performance of such an indicator mainly lies in its sensitivity to the question/objective addressed and in its statistical properties (linked to the observation protocol). To stress the importance of validating indicators through performance criteria, we distinguish metrics (variables observed or calculated at a given scale) from indicators

(metrics that have been validated for management purpose). Also we use "metric" as a catch-all term for metric, score or index, among others. For example, in the case of marine communities, indices concerning species richness, diversity or equitability indices may be computed from the observed table of species abundances (see e.g. [21]) for indices developed to assess the ecological quality of marine waters within the WFD). A metric becomes an indicator if it effectively indicates that for which it was designed. There are two conditions for this: (i) to define the appropriate estimation method to compute the metric from quantitative or qualitative data and (ii) to devise unambiguous rules to interpret the values taken by the metric with respect to thresholds and/or reference values in the light of the question at stake. In the following, the term indicator will be used when the corresponding metric has satisfied these two conditions.

2.2. Indicator performance criteria

In this paper, indicators are assessed according to two criteria: relevance and effectiveness, following Nicholson and Fryer [31].

2.2.1. Relevance

Evaluating the relevance of an indicator is the first essential step in the process of indicator selection. As written earlier, there should be a clear and unequivocal link between the indicator and the objective to which it is supposed to point. Relevance ensures that the direction of changes in indicator values under a given pressure is predictable. For example, species richness, the number of species by taxonomic group, has often been used as an indicator of the diversity of fish assemblages in the MPA context [14].

Relevance encompasses two notions: sensitivity and the existence of quantitative reference values. Both of these notions refer to the knowledge that is required before any monitoring and assessment takes place. Sensitivity reflects the indicator's ability to respond to variations in pressure. For example, the abundance of a fish species is sensitive to a major increase in fishing efforts that target that particular species.

Reference values should be provided so that they may be used to gauge the indicator. In the case of a pressure or impact that is limited over time or is recent, it may be possible to ensure that the former situation can be handled properly in the sampling design over time, i.e. reference values can be established before the impact as in a Before-After-Control-

Impact (BACI) design [34]. However, when the pressure has lasted for years or decades, this comparison is impossible, as pristine conditions are lacking. The case for fishing or for eutrophication is made based on this idea. In addition, basic knowledge may not be available to assess the reference conditions. Regarding indicators at the community level, to our knowledge, no available reference points exist. For example, for a number of indicators issued from fisheries modelling, the underlying theory allows the definition of reference points or of the range of acceptable indicator values [28,35]. The WFD requires that reference conditions correspond to pristine conditions for a certain number of biotic (e.g., benthic communities) and abiotic (e.g. nitrate) parameters and for each type of coastal water body (defined from hydrodynamic characteristics). Unfortunately, it is almost impossible to approximate the pristine conditions that existed in the pre-industrial era (mid-19th century). Thus, it is currently recommended to find a similar type of body of water in an area that is not impacted or in a less impacted area. Modelling can also be used in this context (see e.g. [36]).

The problem of finding reference values extends to the issue of categorising metric values into several classes that lead to distinct diagnostics. The simplest classification scheme is binary discrimination; for example, as fishing impacts are assessed within a background of multiple environmental fluctuations and changes, the identification of the main driving factor may be difficult or impossible. In this case, the hypotheses to be tested are restricted to the following two hypotheses: (i) the null hypothesis, no fishing or the sustainable impact of fishing of the resource and (ii) the alternative hypothesis, overexploited resource. Again, it is difficult to have prior knowledge of the reasonable limit beyond which the impact on the resource is considered significant from a fisheries management or ecological point of view.

2.2.2. Effectiveness

Whereas relevance allows for the selection of potential indicators, effectiveness is the condition that allows the indicator to reach its predefined targets. Achieving indicator effectiveness as defined by reaching a quantified objective must rely on statistical inference.

Precision, accuracy and statistical power. Precision refers to the dispersion of the observations. This definition holds whether or not the mean value around which the dispersion is measured approximates the "true" value [37]. Precision is to be differentiated from with accuracy, which is the property of being close (on average) to some target or true value. The precision of a metric is affected by various sources of variability: measurement errors, temporal variability (e.g. between seasons and across years) and spatial variability at

several scales. All of these sources should be sufficiently understood and documented [20]. Otherwise, pilot studies should be conducted to estimate the variance components. Power is the probability of detecting a change of a given magnitude when it exists [1 – (Type II error)]. The power can be increased by increasing the metric precision using an appropriate sample size, an optimal sampling design and a reduced measurement error. Power increases when the likelihood of a Type I error (the probability of detecting a change when it does not exist) increases. A power analysis is a quantitative and objective way to evaluate indicator performance, along with the previously mentioned relevance to management objectives. Few papers go into these methodological aspects in terms of indicator selection; in particular, and somehow surprisingly, optimising sampling strategy is rarely applied as a way of increasing the probability of effectively detecting progress toward management objectives. Peterman [38] and Peterman and M'Gonigle [39] addressed the issue of power analysis for fisheries management, an issue that was subsequently discussed in relation to indicators by Trenkel and Rochet [28]. Nicholson and Fryer [33] incorporated a power analysis into the design of monitoring programs for chemical contamination. Kurtz et al. [20] illustrated power's sensitivity to changes in the environmental monitoring design. Peterman [38] discussed the use of power in depth and also discussed its implications for the design in the context of fisheries research and management for both scientists and decision makers. He showed how failing to account for statistical power can lead to useless monitoring programs and, consequently, to a waste of money. In practice, we should define the magnitude of the changes we want to be able to detect (either a temporal trend or a comparison to previously defined thresholds). Then, a cost-based trade-off should be used to determine the appropriate sampling strategy (sampling strata, sample sizes and pooled samples).

Auxiliary information. A sub-criterion related to effectiveness is used to analyse and/or interpret the metrics' variations using covariables. The metric precision can be increased by accounting for auxiliary information to control the variability linked to covariables, e.g., control of environmental descriptors, either by modelling the link between the metric and the covariable or by adapting the sampling strategy to the distribution of the covariable. In the case of indicators that are applied at a large spatial scale, a similar issue may be raised. For instance, an importance target species whose abundance is monitored to assess fishing impacts might be absent from certain sites in the area under investigation, e.g., due to habitat considerations. One might then choose to consider a species with comparable life history traits and that is subject to a similar fishing pressure. Alternatively, one could account for habitat information when analysing the spatial variations of the metric [40].

Feasibility. Trade-offs between desirable features, costs and feasibility often determine the choice of indicators [41]. Indicators that are constructed using only conceptual models may be useless for operational purposes if the data needed to properly estimate the indicator are too difficult or expensive to obtain. For example, this is for example, the case for indicators that are based on the abundance and/or the diversity of phytoplankton in coastal waters, as recommended by the WFD [1], that are subject to extreme sampling variabilities in time and space (both horizontally and along the water column). Therefore, these indicators require a very large number of samples to achieve a sufficient precision.

3. A general framework for indicator design and use

As stated in the introduction, we focus on management-oriented indicators. From a decision-maker standpoint, the indicator that is selected should guarantee a safe and unambiguous decision leading to the appropriate regulation/mitigation/control measures. A management decision that is based upon indicators is the end-point of a process for which we propose a frame with explicit steps. Although this process is presented in a chronological way, in practice, iterations are needed to define indicators that achieve desirable performance criteria.

With respect to this framework, the first step is to assist decision-makers/managers/stakeholders in formulating objectives and questions in a way that can facilitate indicator selection. This framework has been successfully tested within the Liteau-MPA Project on Marine Protected Areas funded by the French Ministry for Ecology and Sustainable Development [42]. During this project, seminars that brought scientists and MPA managers together resulted in tables crossing objectives, management actions and indicators (see the MPA illustration in Section 4).

The second step of this framework is to find a relevant metric or a set of metrics for a given objective or action. At this stage, the way the metric selected is going to be analysed should also be precisely defined. For instance, let us consider the risk of eutrophication in near-shore reefs with respect to nutrient enrichment from urban wastes or aquaculture. Chlorophyll a concentration in the body of water was the monitored parameter. Although the link with the objective of eutrophication status assessment here is not questionable, one concentration value observed at a given station does not allow an accurate evaluation of the risk, nor does a mean or median concentration that is estimated from data that are regularly sampled. Large individual values are representative of acute situations; therefore they better

reflect the status of the body of water; thus,, risks can be better assessed from high percentiles, as used by Devlin et al. [10], within the WFD context.

The third step of the framework is to define an appropriate sampling design to ensure the effectiveness of the indicator selected (see subsection 2.2.2). Here it should be looked for a compromise between data collection costs and the need to account for the sources of variation that affect the data that are used for indicator estimation. Data collection costs can increase, either due to the number of samples required to achieve the desired effectiveness and/or because sampling/analytical equipment is expensive. Alternatively, costs can increase to provide more human resources because the current resources are beyond capacity. In some instances, collection costs can be accommodated, e.g., reducing the number of samples and increasing the estimate precision using another observation device. An important step of the indicator selection process is to reject those indicators that are ineffective given the resource constraints. An example will be discussed in the WFD case study (Section 4).

At each step of the framework's process, communication between scientists and managers is essential. First, scientists should help decision-makers formulate their questions and identify management actions that can be triggered by indicator values. A continuous dialogue throughout the process should avoid i) questions that cannot be answered because they are not properly formulated and ii) indicators that are useless for decision-making because they are ineffective, too costly or infeasible. At the end of the process, scientists should communicate their results in a simple way. The underlying ecological or biochemical rationale, the indicators and the corresponding monitoring strategies may be quite complex, as long as the coding process leading from this underlying complexity to the final presentation is made explicit.

Figure 1 summarises the general framework from the definition of objectives to the indicator selection based on their relevance and effectiveness.

4. Applications

4.1. Assessing the status of a coastal body of water under the European Water Framework Directive (WFD) regulatory context

The general objective of the WFD [1] is to reach a good ecological and chemical status for all European bodies of water, including fresh-, ground-, coastal and transitional waters.

Phytoplankton is one parameter used to assess the ecological status and is categorised in five classes ("High", "Good", "Moderate", "Poor" or "Bad"). In order for a body of water to be given a good status, the indicator value must fall into at least the "high" or "good" class. Bodies of water that do not have a good status require management action that is enforced using regulation, which consists of setting up a program of measures that will guarantee a good status by 2015. Therefore, it is extremely important to select relevant and effective indicators. The WFD makes the parameters explicit for deriving indicators for "Phytoplankton" in the list of biological quality elements [1]: (i) the composition and abundance of phytoplanktonic taxa; (ii) the phytoplanktonic biomass and (iii) the frequency and duration of phytoplankton blooms. The real challenge for the scientist is to discriminate a "good" from a "medium" status based on these parameters, e.g., for "Phytoplankton", the challenge is to differentiate "slight" from "moderate" changes in composition and abundance.

Selecting the right indicator will depend on the related management. The issue raised here is that of eutrophication and of the management actions that need to be undertaken to reduce it. It has not been demonstrated that phytoplanktonic composition and abundance or bloom frequency and duration are relevant. Classically, chlorophyll a concentrations, phytoplanktonic biomass and dissolved oxygen concentrations are the most commonly considered parameters for eutrophication (e.g. [38]) (Table 1), and are found in the list of physico-chemical elements [1].

During meetings that were aimed at intercalibrating metrics and scales, in particular for chlorophyll a, state representatives agreed on using chlorophyll a 90th percentiles (see [10]). Taking into account the sources of variation of chlorophyll a measurement, Carstensen [12] showed that to be able to detect a 10% deviation in the limit between "good" (green) and "moderate" (yellow) with an 80% power level, 300 samples per year would be needed. Collecting this many samples would be extremely difficult, if not impossible, using laboratory analyses of field samples. Using satellite images, Gohin et al. [44] computed percentiles at the body of water level that compared very well with statistics derived from *in situ* measurements. Therefore, this remote sensing information was promoted as an adequate tool for the large-scale assessment of bodies of water within the WFD and the European Marine Strategy Framework Directive (MSFD) [45]. Not only is the spatio-temporal resolution much higher using this tool, which allows the collection of a comprehensive data set, but it can cover the spatial extent of the area under scrutiny. This leads to a more precise metric and a subsequent increase in statistical power that also depends on the satellite images that are available during the phytoplanktonic production period.

As for dissolved oxygen, a clear link exists between eutrophication and oxygen availability for living pelagic or benthic organisms. These organisms are found within some physiological *preferenda*, and beyond a value of 5 mg/L, the dissolved oxygen concentration that decreases with respiration, the nutrient enrichment is considered harmful to organisms [43]. We will consider this value a threshold beyond which the oxygen concentration value should not surpass too often, e.g., for less than 10% of the data set. Under this assumption, the 10th percentile is a convenient metric evidencing undesirable events, and the 5 mg.L^{-1} threshold is chosen as a reference point (Table 2). If the metric value falls beyond this threshold, the body of water is not given a good status. With a sensible link with the pressure and having defined this threshold, we can consider this metric relevant. The metric's effectiveness will partly rely on the number of samples collected to assess the eutrophication level in the body of water. The cost to send staff into the field to sample water quickly increase with the number of samples. Alternatively, one could use instrumented buoys and follow their locations using auxiliary information such as hydrodynamics characteristics like currents and/or residence time. High-frequency sampling would provide comprehensive information on the dissolved oxygen dynamics in the targeted part of the body of water and yield the number of hypoxic events (N) occurring in the area (Table 2). Further, managers may decide that N_c is the good-to-moderate threshold that serves to qualify the ecological status. As an example, we performed 300 simulations of monthly sampling from a Summer bi-weekly dissolved oxygen time series that were observed in the Thau lagoon (Hérault, France) in the summer of 2007 (data provided by Th. Laugier), with a major hypoxic event in July. Results showed that monthly sampling fails to detect this event two-thirds of the time. As a result, this sampling frequency fails to evaluate the detrimental effects to the body of water's ecological status and to the lagoon oyster farming industry. By selecting the right metrics and adequate measurement methods, we greatly improved power and effectiveness, and therefore, minimised bias in body of water qualification.

In this case study, discussions with institutional counterparts were crucial to effectively meet WFD recommendations by selecting metrics and observation/analytical means that fulfil indicator performance criteria, such as the use of instrumented buoys or satellite images. This way, we identified potential cost-effective alternatives to field water samples that provide the required statistical power at the scale of the spatial entity to be managed (Table 2).

4.2. Marine Protected Areas

The following example illustrates a context that is somewhat different from § 4.1 because there is no institutional prescription for indicators arising from national or international commitments. Indicators of MPA management effectiveness (see Pelletier 2010 for a description of the context) are being constructed as a result of a collaboration between MPA managers and scientists.

Marine Protected Areas (MPAs) have become widely advocated tools for biodiversity conservation, with quantitative targets regarding the coverage of a global network of MPAs to be reached in the next years (Convention for Biological Diversity (CBD), http://www.biodiv.org/defaults.html). These objectives are reflected in other international texts such as the European "Habitat Directive" [46] and the European MSFD. The Programme of Work for Protected Areas of the CBD specifically called on parties to develop and adopt appropriate methods and standards, criteria and indicators for evaluating management effectiveness and governance by 2008. They also called on parties to assess at least 30% of their protected areas by 2010. Similar objectives are found in the MSFD with different timetables.

Methodologies for assessing the effectiveness of MPA management are still being developed, and several initiatives have been undertaken to evaluate them [47]. Pelletier et al. [35] reviewed assessment methods to track the progress toward the ecosystem conservation objectives in MPAs and discussed the ability of these methods to produce indicators for this purpose. The authors pointed out the lack of articles dealing with indicators of MPA effects, particularly management-oriented indicators. The MPA Management Effectiveness Initiative (MEI) created by the IUCN World Commission on Protected Areas and World Wildlife Fund was tasked with developing a set of marine-specific natural and social indicators to evaluate MPA management effectiveness, based on both scientific and practitioner expertise. The resulting guidebook [15] provides clear insight into indicator construction, but it is not intended to propose assessment methods, guidelines for indicator interpretation or means to evaluate indicator performance. In contrast to Pomeroy et al. [15], the present application does not focus on the definition of metrics but focuses on their use within the framework described in Section 3. We do not intend to be exhaustive and will select two examples of indicators for illustrative purposes.

The two major management goals for MPAs are the conservation of biodiversity and the sustainable exploitation of fish resources [48]. The first example deals with the objective of

restoring and maintaining the abundance of target species under the goal of sustainable exploitation. One of the proposed metrics is the mean size of the target species, which is expected to increase in a no-take area or in an area under reduced fishing pressure (Table 3). This metric might be estimated using Underwater Visual Censuses (UVC) or the number of fish caught; however, both of these values may be experimental if the MPA is a no-take zone. If fishing is allowed in the MPA, the number of fish caught can be used. This metric is very sensitive to the protection of target species [14], but no mean size value exists that can be linked to the sustainable exploitation of the species. In the absence of reference values, diagnostics usually rely and/on spatial and temporal trends. Using a design that involves data collection before and after MPA establishment and within and outside the MPA, it is possible to calculate the following: (i) the evolution over time of the within-outside difference in the mean size, which is denoted by D and (ii) the temporal trend in the mean size outside of the MPA, which is denoted as Tr. We listed plausible cases of trends for D and Tr and linked their variations to management actions (Table 4). If D increases and Tr increases or is stable, then no regulation is needed, and the species is deemed to be in good status. If D increases or is stable and Tr decreases, an excessive fishing pressure outside of the MPA can be stressed. If fisheries management is in the MPA manager's range of action, he or she might use regulatory action; otherwise, the negative diagnostic about the species should be reported by the MPA manager to the competent authorities. In contrast, if D decreases and Tr increases, a problem obviously exists within the MPA. If spillover occurs from the MPA to the surrounding areas, this may be a signal of a potential future problem outside the MPA. In the worst case, if D decreases and Tr decreases, an additional problem inside the MPA is likely, e.g., a problem that is linked to the degradation of essential habitats for the target species. In this case, additional action is needed within the MPA, and with the help of auxiliary information, such as habitat monitoring (see next example), actions to mitigate the cause of the degradation should be implemented by the MPA manager.

The second example deals with the objective of maintaining representative habitats under the goal of biodiversity restoration and conservation. Two complementary metrics are proposed, the seagrass cover (SC) and the number of boats moored in the MPA (NM). Seagrass beds are an essential habitat of many juvenile fish, and its cover may evolve over time as a result of changes in water quality or detrimental effects of moorings. SC is usually estimated by UVC, although other observation techniques such as video or photo may be considered. Seagrass beds worldwide are known to suffering from mechanical damage caused by boat anchoring, particularly in sites highly frequented by boaters [50].Yet, there is no quantified link between the number of boats moored in a given location and seagrass

cover; therefore, no reference point is available. NM may be estimated through a frequentation survey (Table 3).

A decrease in seagrass cover indicates a problem, but it is not necessarily linked to a cause that is under control by the MPA manager, e.g., it may be due to water quality. By considering the number of boats moored in the MPA, especially those nearby or in the seagrass area, we get more information about the cause of the evolution of SC. In the absence of a reference value, comparing seagrass areas with different frequentation levels by boats and observing a correlation that is independent of other sources of variation for SC is necessary to attribute a decrease in SC to an excessive NM. We listed plausible combinations of trends for SC and NM (Table 4). When SC increases and NM does not increase, then the status of the seagrass is satisfactory and no further management action is required. If SC is stable and NM increases, then it might be relevant to anticipate a potential degradation of seagrass cover by installing permanent moorings. In a more serious situation, if SC decreases and NM decreases, a temporary prohibition of boat access to the MPA together with a tighter monitoring of SC might be an option.

The SC and NM metrics and the corresponding interpretation rules approach were tested in a protected site of the New Caledonian South lagoon (South Pacific). The SC was calculated in two contiguous areas near the MPA of Laregnere Islet: one close to the islet where boats moor or anchor (on average 4.9±4.8 boats present per day), and the other further from the shore, where boats never anchor. The average SC measured by underwater video was 12.7% in the first location and 27.6% in the second location. The SC was significantly greater in the area where boats do not anchor (Student test, t = -3.2, df = 53, p-value<0.002). In this case, no data are available to date to test temporal trends, but the spatial difference between these two proximate locations clearly indicates an impact of anchoring. As the SC was 46% less, on average, in the impacted area, the result of the diagnostic is serious and should lead to the implementation of measures to control either anchoring or frequentation in the area.

In both examples, the different diagnostics are not guided by metrics thresholds but by the statistical significance, magnitude and direction of the metrics considered. Determining the appropriate significance level, magnitudes and direction stem from the scientific knowledge about ecological dynamics and also from the time frame required for restoration, which in turn depends on the initial ecological status. Regarding SC, a significantly positive short-term trend in the zone where moorings are restricted (and where NM significantly decreases as a result) indicates that the restriction has detectable consequences on SC. This trend may be considered as an acceptable outcome for MPA managers. In other situations, for instance

when forceful restoration strategies are implemented, more demanding thresholds may be sought by managers, e.g., a 10% increase in SC in five years. The ability to interpret the variations of SC and NM confirms the relevance of these two metrics for the considered management objective. Establishing this kind of diagnostic requires several years of data because natural fluctuations of exploited populations and other anthropogenic causes of variation of seagrass might make changes due to mooring management difficult to observe due to MPA protection. Therefore, the interpretation of indicator values may become more sophisticated by accounting for values taken over successive years. Once many years of data have been collected that correspond to a larger number of protection years, there may be enough information in the monitoring data to determine reference values. For instance, a formal relationship may be identified between seagrass cover and frequentation by boats. Similarly, for target species, larger sets of data might be used to determine the range of plausible values for the mean size that corresponds to an *a priori* safe status of the species.

The effectiveness of these indicators is first linked to the possibility of detecting MPA-related effects using the analysis of the ecological metrics considered, such as the abundance of target species and seagrass cover. In this respect, accounting for auxiliary information such as habitat is very important. In this paper, habitat is meant to represent the biotic and abiotic environment of target species and seagrass. It is well-known that it is meaningless to assess MPA effects while ignoring the influence of habitat on the spatial distribution of the ecological metric of interest [35,51]. The feasibility of the considered metrics is determined by monitoring means and staff capacity. Abundance of target species might be obtained through UVC, remote underwater video or the number of fish caught (in this case leading to a relative abundance index). UVC requires expert divers that are able to identify fish species, and each diver may achieve 2 or 3 observations per day (in the case of transects). Additional observations may be needed to describe the habitat. Remote underwater video does not require divers. More observations may be realised per day and per video system used, and habitat information can be collected together with fish information; however, more work is required at the office for image analysis. In addition, the values representing the number of fish that are caught are generally obtained by logbook systems or interviews, whether in landing sites or at sea. Unlike UVC and underwater video, the catch location is often not precise. In most cases, the choice of the observation system is dictated by the staff capacity. Achieving a given effectiveness is then a matter of devising the appropriate observation protocol while accounting for monitoring the budget and the allocated staff time. Systems that enable more stations to be carried out per day are more likely to satisfy managers' needs provided that they can be programmed and accomplished using the human and financial resources at hand.

5. Discussion - Conclusions

Our proposed framework is designed for a management context in which it is expected that monitoring efforts will guide toward adequate management actions. We consider reporting to be the least that can be done using data generated by such monitoring programs. A wide range exists in the kind of information that can be asked from managers. First, there may be an obligation to report the monitoring effort; managers may have to show that monitoring is implemented and that data have been collected by displaying the results, without further requirement. Second, monitoring data are collected and exploited, and the manager reports the results, possibly at a larger scale; for example, the Global Coral Reef Monitoring Network proposes a simple monitoring tool, the Reef Check, which aims to obtain a representative image of reef status at global level [52]. The user of this tool does not intend to trigger local management actions; however, in an opportunistic way, it might point out small-scale undesirable events. Using this tool, invasions of the sea urchin *Acanthaster* have been recently observed in the New Caledonian lagoon through the Reef Check data (Wantiez, pers. comm.). Third, monitoring data are collected and exploited in the form of indicator values that help the manager to evaluate the effectiveness of the actions that have been undertaken and to possibly guide toward additional actions. The assessment of the MPA's performance mostly falls into this third category because a positive MPA effect clearly shows that the MPA management is achieving its objectives, which in turns reasserts the soundness of the MPA's existence and justifies its budget. In addition, the WFD and its objective to achieve a final good status based on European body of water quality assessment is justified, which should lead to measures of pollution reduction.

The main feature of our framework proposal is to provide an integrated chronological (although iterative) procedure at two levels: (i) using indicators in the management context and (ii) considering performance criteria within the indicator selection process. The application of the framework to these two distinct contexts shows that it is general and simple to implement. In a very important paper [24], more complexity was brought to formalising interactions among indicators by means of causal chain networks, which should lead to an optimal selection of indicators within an enhanced DPSIR framework. In our work, we consider this aspect as part of the first step because the formulation of objectives requires such an analysis. As shown in the MPA case study, this level of complexity can be achieved by deriving optimal interpretation rules from the combination of individual indicator results.

In practice, indicator development is seldom the result of a sustained collaboration between managers and scientists; however, most of the studies on ecological indicator characteristics agree that indicators should address management questions [53]. In our opinion, such collaboration is essential to a successful match between indicators and management actions. First, this collaboration helps managers to formulate objectives and actions in a way that enables scientists to select the most relevant metrics. Second, it ensures that the categorisation of indicator values guides managers toward adequate actions. Third, it should lead to effective indicators because they are derived from a sampling strategy that is optimised with respect to capacity and cost constraints. The metrics proposed in Section 4 stem from workshops organised with the concerned management bodies, water agencies, the Ministry of Environment [54] and MPA managers [42].

The proposed framework should lead to concrete outputs, from the explicit (re-) formulation of objectives and subsequent management actions to the specification of the monitoring protocol, including the sampling design and observation systems. Ultimately, the process is successful when the derived indicators are explicitly included into a management plan. This approach is currently being implemented in the PAMPA LITEAU III project (http://www.ifremer.fr/pampa) in which a set of indicators for MPA performance are being developed.

Generally, the proposed criteria are similar to relevance and effectiveness (precision, measurement, costs), but there is no recommendation as to the order in which this list of criteria should be considered in the selection process, and no priorities are set for these criteria. First, the relevance of a given metric should be established from prior knowledge using research works either from literature or from pilot studies, and the effort required to achieve a given effectiveness should then be assessed. Pilot studies from previous research provide components of variance, allowing one to minimise errors when categorising indicator values, e.g., by better discriminating natural fluctuations or changes from values observed under an expected impact. Evaluating the overall cost required to make an indicator effective constitutes the second step of the process, and this step is under the constraints of capacity, costs and technical feasibility. This last point is essential because sampling costs often determine the choice of observation systems (see subsection 2.2.2).

A large scope for methodological investigations on indicators still exists. For example, a recurring question is how to synthesise information from several individual indicators [23]. Borja et al. [55] identified this question as one of the challenges in coastal water quality management for the next decade. Dauvin et al. [56] proposed a report card within a WFD

context. Rice and Rochet [29] reviewed different methods for combining indicators, and Trenkel et al. [57] provided an alternative to classical fish stock assessment by defining a suite of indicators and their combination. The WFD [1] recommends that one defines the ecological status on the basis of the worst class that is obtained across all of the indicators. This recommendation is a precautionary approach that bears the advantage of being simple, though perhaps too simplistic. Borja et al. [58] relaxed this rule by allowing at least one element to not meet the corresponding standard in an integrative process. The approach we presented here bears the advantage of involving managers at almost each step, which seems important to synthesise and communicate indicators, as raised by Peterman [59]. approach could be pursued and enriched using exchanges between managers and scientists about this multidimensional aspect of indicator dashboards. The price to pay for aggregation is a necessary loss of relevance because a synthetic score cannot be directly linked to some pressure or impact. Thus, in accordance with our framework, it is preferable to direct efforts toward the collaborative interpretation of a set of indicators. For instance, it would be useful to assess managers' perceptions facing a variety of warning signals to select the appropriate panel of indicators. It would also be useful to take into account the governance issues (socio-economic and political issues and participative management, among others). The fact that indicators are easily understood by managers and by non-scientists is often proposed as a criterion for indicator selection (see e.g. [25]). In our opinion, the ability to communicate about indicators should not be regarded as a criterion for indicator selection, and the proposed framework should enable one to distinguish the issues of indicator development and communication. However, communication should involve the scientists and managers that worked for indicator construction in collaboration with communication experts. Neglecting this part of the process might result in an undesirable situation in which an expensive program that was aimed at collecting data that fulfil manager needs fails to impact manager decisions because the proper way of conveying results to end users has been overlooked. Delivering a simple answer to management questions should be viewed as a subsequent and challenging task to complete this indicator approach.

Acknowledgments

References

[1] European Union. Water Framework Directive. Directive 2000/60/CE of the European Parliament and Council of 23 October 2000. N° L 327, 22/12/2000, 2000.

[2] A. Borja, I. Muxika, J. Franco. The application of a Marine Biotic Index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Mar. Pollut. Bull. 46 (2003) 835-845.

[3] P. Muniz, N. Venturini, A.M.S. Pires-Vanin, L.R. Tommasi, A. Borja. Testing the applicability of a Marine Biotic Index (AMBI) to assessing the ecological quality of soft-bottom communities, in the South America Atlantic region. Mar. Pollut. Bull. 50 (2005) 624-637.

[4] H.L. Rees, S.E. Boyd, M. Schratzberger, L.A. Murray. Role of benthic indicators in regulating human activities at sea. Environ. Sci. Policy 9 (2006) 496-508.

[5] J.C. Dauvin, 2007. Paradox of estuarine quality: benthic indicators and indices, consensus or debate for the future. Mar. Pollut. Bull. 50 (2007) 107-110.

[6] H. Blanchet, N. Lavesque, T. Ruellet, J.-C. Dauvin, P.-G. Sauriau, N. Desroy, C. Desclaux, M. Leconte, G. Bachelet, A.-L. Janson, C. Bessineton, S. Duhamel, J. Jourde, S. Mayot, S. Simon, X. de Montaudouin. Use of biotic indices in semi-enclosed coastal ecosystems and transitional waters habitats – Implications for the implementation of the European Water Framework Directive. Ecol. Indic. 8 (2008) 360-372.

[7] A. Borja, D.M. Dauer, R.J. Díaz, R.J. Llansó, I. Muxika, J.G. Rodríguez, L.C. Schaffner. Assessing estuarine benthic quality conditions in Chesapeake Bay: a comparison of three indices. Ecol. Indic. 8 (2008) 395-403.

[8] F. Bianchi, F. Acri, F. Bernardi-Aubry, A. Berton, A. Boldrin, E. Camatti, D. Cassin, A. Comaschi. Can plankton communities be considered as bio-indicators of water quality in the Lagoon of Venice? Mar. Pollut. Bull. 46 (2003) 964-971.

[9] S. Sagert, D.K. Jensen, P. Henriksen, T. Rieling, H. Schubert. Integrated ecological assessment of Danish Baltic Sea coastal areas by means of phytoplankton and macrophytobenthos. Estuarine Coastal Shelf Sci. 63 (2005) 109-118.

[10] M. Devlin, M. Best, D. Coates, E. Bresnan, S. O'Boyle, R. Park, J. Silke, C. Cusack, J., Skeats. Establishing boundary classes for the classification of the UK marine waters using phytoplankton communities. Mar. Pollut. Bull. 55 (2007) 91-103.

[11] C. Pergent-Martini, V. Leoni, V. Pasqualini, G.D. Ardizzone, E. Balestri, R. Bedini, A. Belluscio, T. Belsher, J. Borg, C.F. Boudouresque, S. Boumaza, J.M. Bouquegneau, M.C. Buia, S. Calvo, J. Cebrian, E. Charbonnel, F. Cinelli, A. Cossu, Di G. Maida, B. Dural, P. Francour, S. Gobert, G. Lepoint, A. Meinesz, H. Molenaar, H.M. Mansour, P. Panayotidis, A. Peirano, G. Pergent, L. Piazzi, M. Pirrotta, G. Relini, J. Romero, J.L. Sanchez-Lizaso, R. Semroud, P. Shembri, A. Shili, A. Tomasello, B. Velimirov. Descriptors of *Posidonia oceanica* meadows: use and application. Ecol. Indic. 5 (2005) 213-230.

[12] J. Carstensen. Statistical principles for ecological status classification of Water Framework Directive monitoring data. Mar. Pollut. Bull. 55 (2007) 3-15.

[13] B. Cicin-Sain, S. Belfiore. Linking marine protected areas to integrated coastal and ocean management: a review of theory and practice. Ocean Coast. Manag. 48 (2005) 847-868.

[14] D. Pelletier, J. García-Charton, J. Ferraris, G. David, O. Thébaud, Y. Letourneur, J. Claudet, M. Amand, M. Kulbicki, R. Galzin. Designing indicators for evaluating the effects of Marine Protected Areas on coral reef ecosystems: a multidisciplinary standpoint. Aquat. Living Resour. 18 (2005) 15-33.

[15] R.S. Pomeroy, J.E. Parks, L.M. Watson. How is your MPA doing? A guidebook of natural and social indicators for evaluating marine protected area management effectiveness. IUCN, WWF, Gland and the US National Oceanic and Atmospheric Administration (NOAA): Gland and Cambridge, 2004.

[16] OECD. OECD Environmental Indicators – Development, measurement and use – Reference paper. OECD, Paris, 1993.

[17] S.C. Mangi, C.M. Roberts, L.D. Rodwell. Reef fisheries management in Kenya: preliminary approach using the drievr-pressure-state-impacts-response (DPSIR) scheme of indicators. Ocean Coast. Manag. 50 (2007) 463-480.

[18] P. Vella, R. Bowen, A. Frankic. An evolving protocol to identify key stakeholder-influenced indicators of coastal change: the case of Marine Protected Areas. ICES J. Mar. Sci. 66 (2009) 203-213.

[19] J. Hilty, A. Merenlender. Faunal indicator taxa selection for monitoring ecosystem health. Biol. Conserv. 92 (2000) 185-197.

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

[21] A. Borja, D.M. Dauer. Assessing the environmental quality status in estuarine and coastal systems: Comparing methodologies and indices. Ecol. Indic. 8 (2008) 331-337.

[22] J. Rice. Environmental health indicators. Ocean Coast. Manag. 46 (2003) 235-259.

[23] F. Müller, R. Lenz. Ecological indicators: theoretical fundamentals of consistent applications in environmental management. Ecol. Indic. 6 (2006) 1-5.

[24] D. Niemeijer, R.S. de Groot, 2008. A conceptual framework for selecting environmental indicator sets. Ecol. Indic. 8 (2008) 14-25.

[25] S.M. Garcia, J.M. Staples. Sustainability reference systems and indicators for responsible marine capture fisheries: a review of concepts and elements for a set of guidelines. Mar. Freshw. Res. 51 (2000) 385-426.

[26] S.M. Garcia, D.J. Staples, J. Chesson. The FAO guidelines for the development and use of indicators for sustainable development of marine capture fisheries and an Australian example of their application. Ocean Coast. Manag. 43 (2000) 537-556.

[27] M.J. Rochet, V., Trenkel. Which community indicators can measure the impact of fishing? A review and proposals. Can J. Fish. Aquat. Sci. 60 (2003) 86-99.

[28] V. Trenkel, M.J. Rochet. Performance of indicators derived from abundance estimates for detecting the impact of fishing in fish community. Can J. Fish. Aquat. Sci. 60 (2003) 67-85.

[29] J. Rice, M.J. Rochet. A framework for selecting a suite of indicators for fisheries management. ICES J. Mar. Sci. 62 (2005) 516-527.

[30] M.L. Fontalvo-Herazo, M. Glaser, A. Lobato-Ribeiro. A method for the participatory design of an indicator system as a tool for local coastal management. Ocean Coast. Manag. 50 (2007) 779-775.

[31] S. López-Ridaura, O. Masera, M. Astier. Evaluating the sustainability of complex socio-environmental systems. The MESMIS framework. Ecol. Indic. 2 (2002) 165-148.

[32] FAO. Indicators for sustainable development of marine capture fisheries. FAO, Roma, 1999.

[33] M. Nicholson, R. Fryer. Developing effective environmental indicators - does a new dog need old tricks? Mar. Pollut. Bull. 45 (2002) 53-61.

[34] A.J. Underwood. Beyond BACI: the detection of environmental impacts on populations in the real, but variable, world. J. Exp. Mar. Biol. Ecol. 161 (1992) 145–178.

[35] D. Pelletier, J. Claudet, J. Ferraris, L. Benedetti-Cecchi, J.A. García-Charton. Assessing ecological and fisheries-related effects of Marine Protected Areas: Current status and perspectives. Can. J. Fish. Aquat. Sci. 65 (2008) 765-779.

[36] D. Pont, B. Hugueny, U. Beier, D. Goffaux, A. Melcher, C. Rogers, N. Roset, S. Schmutz. Assessing river biotic condition at at continental scale: a European approach using functional metrics and fish assemblages. J. Appl. Ecol. 43 (2006) 70-80.

[37] Y. Dodge. The Oxford Dictionary of Statistical Terms, sixth ed., Oxford University Press, Oxford, 2003.

[38] R.M. Peterman. Statistical power analysis can improve fisheries research and management. Can. J. Fish. Aquat Sci. 47 (1990) 2-15.

[39] R.M. Peterman, M. M'Gonigle. Statistical power analysis and the precautionary principle. Mar. Pollut. Bull. 24 (1992) 231-234.

[40] J. Ferraris, D. Pelletier, M. Kulbicki, C. Chauvet. Assessing the impact of removing reserve status on Abore Reef fish assemblage in New Caledonia. Mar. Ecol. Prog. Ser. 292 (2005) 271-286.

[41] V.H. Dale, S.C. Beleyer. Challenges in the development and use of ecological indicators. Ecol. Indic. 1 (2001) 3-10.

[42] D. Pelletier. Développement d'outils diagnostics et exploratoires d'aide à la décision pour évaluer la performance d'Aires Marines Protégées. Rapport final de contrat Liteau II, Ministère de l'Ecologie et du Développement Durable. Rapport scientifique 49 p., Rapport final 32 p. et Annexes 286 p. (http://www.liteau.ecologie.gouv.fr/IMG/pdf/LITEAU2-Pelletier.pdf), 2007.

[43] S.B. Bricker, C.G. Clement, D.E. Pirhalla, S.P. Orlando, D.R.G. Farrow. National Estuarine Eutrophication Assessment: Effects of nutrient enrichment in the Nation's estuaries. NOAA, National Ocean Service, Special Projects Office and the National Centers for Coastal Ocean Science. Silver Spring, MD: 71 p., 1999.

[44] F. Gohin, B. Saulquin, H. Oger-Jeanneret, L. Lozac'h, L. Lampert, A. Lefebvre, Riou Ph., F. Bruchon. Towards a better assessment of the ecological status of coastal waters using satellite-derived chlorophyll-a concentrations. Remote Sens. Environ. 112 (2008) 3329-3340.

[45] European Union. Marine Strategy Framework Directive. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy. N° L 164, 25/06/2008. 22 p., 2008.

[46] European Union. Habitats Directive. Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. N° L 206, 21/05/1992. (Last amended by Directive 2006/105/EC, n° L 363, 20/12/2006, p. 368), 1992.

[47] S. Wells, N. Dahl-Tacconi. Table: Methodologies for Evaluating MPA Management Effectiveness. MPA News, 2-3, 2006.

[48] J. Claudet, D. Pelletier. Marine protected areas and artificial reefs: A review of the interactions between management and scientific studies. Aquat. Living Resour. 17 (2004) 129-138.

[49] A. Stewart-Oaten., W.W. Murdoch, S.J. Walde. Estimation of temporal variability in populations. Am. Nat. 146 (1995) 519–535.

[50] M. Milazzo, F. Badalamenti, G. Ceccherelli, R. Chemello. Boat anchoring on *Posidonia oceanica* beds in a marine protected area (Italy, western Mediterranean): effect of anchor types in different anchoring stages . J. Exp. Mar. Biol. Ecol. 299 (2004) 51-62.

[51] J.A. García-Charton, A. Pérez-Ruzafa. Ecological heterogeneity and the evaluation of the effects of marine reserves. Fish. Res. 42 (1999) 1-20.

[52] C. Wilkinson. Status of Coral Reefs of the World: 2008. Global Coral Reef Monitoring Network and Reef and Rainforest Research Centre, Townsville, Australia, 296 p, 2008.

[53] E.C. Wicks, Longstaff B.J., B. Fertig, W.C. Dennison. Ecological indicators – Assessing ecosystem health using metrics. In Integrating and Applying Science. In: B.J. Longstaff, T.J.B. Carruthers, W.C. Dennison, T.R. Lookingbill, J.M. Hawkey, J.E. Thomas, E.C. Wicks, J. Woerner, editors. A practical handbook for effective coastal ecosystem assessment. Ian Press and University of Maryland Center for Environmental Science (2010) p. 61-77.

[54] MEEDDM. Guide relatif à la définition du bon état des eaux littorales (eaux côtières et eaux de transition), en application de la directive cadre sur l'eau. Version novembre 2009, 54 p., 2009.

[55] A. Borja, A. Ranasinghe, S.B. Weisberg. Assessing ecological integrity in marine waters, using multiple indices and ecosystem components: Challenges for the future. Mar. Pollut. Bull. 59 (2009) 1-4.

[56] J.C. Dauvin, C. Fisson, J. Garnier, R. Lafite, T. Ruellet, G. Billen, J. Deloffre, R. Verney. A report card and quality indicators for the Seine estuary : From scientific approach to operational tool. Mar. Pollut. Bull. 57 (2008) 187-201.

[57] V.M. Trenkel, M.J. Rochet, B. Mesnil. From model-based prescriptive advice to indicator-based interactive advice. ICES J. Mar. Sci. 64 (2007) 768-774.

[58] A. Borja, J. Bald, J. Franco, J. Larreta, I. Muxika, M. Revilla, J.G. Rodríguez, O. Solaun, A. Uriarte, V. Valencia. Using multiple ecosystem components, in assessing ecological status in Spanish (Basque Country) Atlantic marine waters. Mar. Pollut. Bull. 59 (2009) 54-64.

[59] R.M. Peterman. Possible solutions to some challenges facing fisheries scientists and managers. ICES J. Mar. Sci. 61 (2004) 1331-1343.

Figure 1: General operational framework linking management questions to indicator performance criteria. Finding a relevant metric to a predefined objective appears as a highly iterative process. Dialogue between decision makers and scientists is essential at each step. As symbolised by the dashed line, this framework is subject to recycling as objectives can be reformulated conditionally to the obtained results.
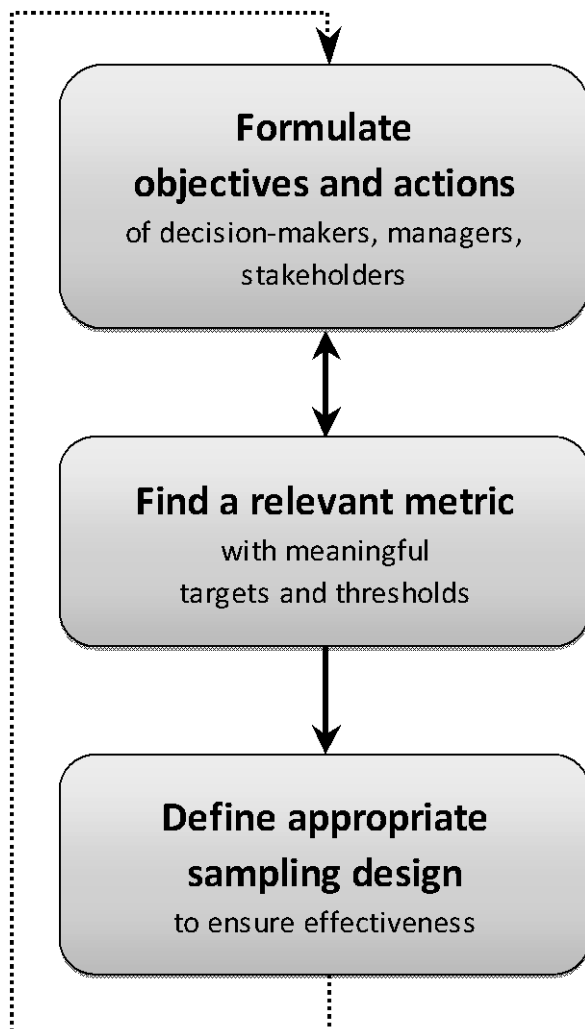
Formulate
objectives and actions
of decision-makers, managers,
stakeholders

Find a relevant metric
with meaningful
targets and thresholds

Define appropriate
sampling design
to ensure effectiveness

Table

Table 1. Three metrics related to coastal water eutrophication: i) 90th percentile chlorophyll a (Chl*a*) concentration ii) 10th percentile of dissolved oxygen concentration measured in estuarine and coastal waters; and, iii) number of hypoxic events, with estimated relevance and effectiveness.

| Metrics | Indicator attributes | | | | | |
| | Relevance | | | Effectiveness | | |
	Rationale	*Sensitivity*	*Ref. value(s)*	*Required precision*	*observation and measurement system*	*Cost*
[Chl*a*] mg.L^{-1} percentile 90%	Indirect link with eutrophication through respiration	high	yes	high	Water samples / Fluorimetry in laboratory	high
[O$_2$] mg.L^{-1} percentile 10%		high	yes	medium	Water samples / *in situ* measurement	medium
Number of hypoxic events		low	no	high	*In situ* measurement	low

Table

Table 2. Conceptual (chronological) framework for indicator use in a management context with an illustration from the European Water Framework Directive (WFD). Definitions of general goal and detailed objectives are followed by management actions linked to the indicator(s) through interpretation rule(s); then the definition of the observation system (sampling and measurement characteristics) leads to its cost estimation, allowing to assess the feasibility of the indicator. Communication between managers and scientists is essential at every step. $\widetilde{X}_{90}$ is the chlorophyll a concentration 90% percentile and $\widetilde{X}_{90c}$ the corresponding threshold value. $\widetilde{X}_{10}$ is the dissolved oxygen concentration 10% percentile and $\widetilde{X}_{10c}$ the corresponding threshold value. N is the number of hypoxic events and N_c the corresponding threshold value.

Goal	Objective	Management actions	Interpretation rule	Metrics	Observation system and sampling strategies	Feasability / Cost
European water in good ecological status in 2015	Assessment of water body ecological status – eutrophication	Reduce water body nutrient enrichment No mitigation measures taken	$\widetilde{X}_{90}>\widetilde{X}_{90c}$ $\widetilde{X}_{90}\leq\widetilde{X}_{90c}$	Chlorophyll a concentration 90th percentile	Satellite images provides high temporal and spatial frequency at the water body scale	Images and data processing algorithms available at low cost
		Reduce water body nutrient enrichment No mitigation measures taken	$\widetilde{X}_{10}<\widetilde{X}_{10c}$ $\widetilde{X}_{10}\geq\widetilde{X}_{10c}$	Dissolved oxygen concentration 10th percentile	Water samples and analysis at the laboratory	High cost related to the number of samples required to reach a given power
		Reduce water body nutrient enrichment No mitigation measures taken	$N>N_c$ $N\leq N_c$	Number of hypoxic events	*In situ* instrumentation allows high temporal frequency	Technically feasible Good ratio number of data/acquisition cost

Table

Table 3. Three metrics related to MPA major management objectives: i) sustainable exploitation of resources (first metric); and ii) conservation of biodiversity (second and third metrics), with estimated relevance, effectiveness and cost. Relevance and effectiveness were evaluated from Pelletier et al. [14]. UVC stands for Underwater Visual Census.

| Metrics | Indicator attributes | | | | | |
| | Relevance | | | | Effectiveness | |
	Rationale	*Sensitivity*	*Ref. value(s)*	*Required precision*	*Observation system*	*Cost*
Mean size of target species	Increases in no-take areas	High	No	High	UVC Catch	High Medium
Seagrass percent cover	Should increase in MPA	Medium	No	Medium	UVC	High
Nb of boats moored in MPA	Moorings destroy seagrass	High	No	Medium	Frequentation study	Low

Table

Table 4. Application of the framework to the MPA management context. Definitions of general goal and detailed objectives are followed by management actions derived in relation with the indicator(s) selection; then the definition of the observation system (sampling and measurement characteristics) leads to its cost estimation, allowing to assess the feasibility of the indicator. Communication between managers and scientists is essential at every step. For the mean size of target species, D is the difference between inside and outside the MPA, while Tr is the interannual difference outside the MPA. BACIPS stands for Before-After-Control-Impact-Paired Series [49].

Goal	Objective	Management actions	Interpretation rule	Metrics	Observation system and sampling strategies	Feasability / Cost
Sustainable exploitation of resources	Restore and maintain target species	No action	$D\uparrow \& (Tr\uparrow \text{ or } \rightarrow)$	Mean size of target species	UVC	Diving feasibility & costs
		Fishing regulation	$D\,(\uparrow\text{or}\rightarrow)\,\& Tr\downarrow$		Catch sampling	Sampling costs
		Fishing regulation + Restrict MPA access	$D\downarrow \& Tr\downarrow$		BACIPS design in each case	
Restore and conserve biodiversity	Maintain representative habitats	No action	$SC\uparrow \& (NM\uparrow \text{ or } \rightarrow)$	Seagrass cover (SC)	UVC	Diving feasibility & costs
		Install permanent moorings in MPA	$SC\rightarrow \& NM\uparrow$	Nb of boats moored in MPA (NM)	Frequentation study stratified per month	Field costs
		Prohibit boat access to MPA	$SC\downarrow \& NM\downarrow$			