Evaluating guidelines as key components of conservation policies: a conceptual framework and a pilot application

Angela Osorio[1,2] ⓘ, Laurent Schmitt[1] ⓘ, Dominique Badariotti[1] and Yves Meinard[3] ⓘ

[1]Laboratoire Image Ville Environnement (LIVE UMR 7362), University of Strasbourg, Strasbourg, France; [2]Social and Human Sciences Research Institute (ICSH), University of Caldas, Manizales, Columbia and [3]Centre Gilles Gaston Granger (CGGG, UMR 7304), Aix-Marseille Université, Aix-en-Provence, France

Summary

Guidelines for managers of protected areas are an important component of conservation policies, on a par with large-scale frameworks and vehicles for conservation funding. In line with the recent literature proposing evaluations of conservation actions or political strategies to improve them, here we use an innovative, hybrid methodology, based both on an interpretative approach anchored in social sciences and a quantitative literature review, to identify available frameworks for evaluating conservation guidelines. The main result of this analysis is that the relevant literature in conservation is sparse and heterogeneous, but a relevant encompassing framework is provided by the literature in decision sciences based on the policy analytics framework. This evaluation framework consists of three criteria: scientific credibility, operationality and legitimacy. We then implement a pilot application by evaluating guidelines currently used in France to support all of the actors involved in protected areas management. The study concludes that these guidelines are plagued by significant weaknesses that could be overcome by implementing relevant participatory processes.

CAMBRIDGE UNIVERSITY PRESS

Introduction

The elaboration and spreading of guidelines for managers of protected areas (PAs) play key roles in conservation policies, on a par with large-scale frameworks, such as the European Directives (Evans 2012), or vehicles for conservation funding, such as dedicated parts of the Common Agricultural Policy in Europe (Linares et al. 2022). These various components of conservation policies are often interrelated, since institutions and non-governmental organisations (NGOs) financing conservation actions frequently condition funding upon the application of guidelines produced by the same or other institutions, whose legal status and role are entrenched in large-scale conservation frameworks.

The importance of evaluating conservation initiatives, which is increasingly acknowledged in the scientific literature (Álvarez-Fernández et al. 2020a, 2020b, Pearson et al. 2022), is relevant to these various components of conservation policies. Indeed, evaluations can help improve conservation by helping us to learn from past errors and by streamlining funding towards effective actions (Grantham et al. 2009, Bottrill et al. 2011). This logic holds true both for concrete conservation actions in the field on a local scale and for conservation policies on national, regional or even global scales (Baylis et al. 2015). The European Natura 2000 policy is exemplary in this respect, as its implementation involves iterative evaluations (Jeanmougin et al. 2017). More specifically, in the case of conservation guidelines, evaluation is needed because guidelines for managers can play major roles in the success or failure of conservation initiatives at two levels. First, ill-conceived guidelines can mislead managers into setting up and implementing wrongheaded conservation actions. Second, because institutions and NGOs financing conservation actions can condition funding upon their application, ill-conceived guidelines can channel funding towards defective conservation projects.

Evaluating guidelines, however, involves specific methodological challenges. Indeed, although evaluating conservation actions involves numerous difficult technical challenges, it can be done rather unequivocally, in a typical evidence-based approach (De Marchi et al. 2016), by quantifying whether the evaluated actions have had positive impacts on conservation targets, such as the populations of targeted threatened species (e.g., Sanderson et al. 2015). Many such quantitative assessments of conservation effectiveness have been conducted in both the grey and the academic literature in recent years (Courrau et al. 2006, Stolton et al. 2019). They show important promise to improve conservation actions (Courrau et al. 2006, Bottrill & Pressey 2012, Geldmann et al. 2013, Watson et al. 2014, Stolton 2019).

In the case of management guidelines, however, many confounding factors can make it technically difficult – and conceptually questionable – to assess the quality of guidelines based on the success or failure of the actions they guide. Indeed, faultless guidelines can be ill-applied, ill-intentioned actors can undermine their application, unforeseeable political or socio-economic dynamics can render them inapplicable and there may be insufficient relevant data to implement effectiveness assessments. Therefore, evaluating guidelines requires a broader framework, overcoming the limitations of effectiveness assessment, by supplementing this criterion with other criteria.

Against this background, our focal question in the present article is: what framework can be used for evaluating conservation guidelines? To answer this question, we test two hypotheses.

The first hypothesis is that, although the academic conservation literature contains numerous, piecemeal attempts at evaluating various aspects of management guidelines, a general framework, whose relevance is proven in the academic scientific literature, is lacking. Notice that, as it is formulated here, this hypothesis focuses only on evaluation frameworks with proven academic scientific credentials and thereby excludes numerous frameworks produced and used by field experts or expert institutions, whose relevance is entrenched in practice rather than in academia. This reflects a basic assumption of ours in this article, according to which the abovementioned promises of evaluations are predicated on proven academic scientific robustness. This assumption should not be misunderstood as disparaging frameworks produced by field experts and institutions but as a reminder that academic science has an important role to play in validating the credentials of such frameworks.

To test our first hypothesis, we use an innovative, hybrid methodology based both on an interpretative approach anchored in social sciences and a quantitative review of the academic conservation literature.

The second hypothesis, which is inspired by Jeanmougin et al. (2017) and Choulak et al. (2019), amongst others, is that 'policy analytics' (Meinard et al. 2021), a framework introduced in, and currently mainly confined to, the literature in decision sciences, provides the general framework needed by encompassing all of the relevant piecemeal contributions found in the academic conservation literature. Policy analytics is a multicriteria framework that champions the use of three criteria, in addition to effectiveness, in policy evaluation:

- Scientific credibility – this criterion refers to the need for the evaluated objects (in our case, guidelines) to be based on scientific findings. This echoes the numerous arguments advocating for the need for conservation policies to be anchored in conservation science (Dubois et al. 2020) in order to overcome knowledge or implementation gaps (Knight et al. 2008, Arlettaz et al. 2010).
- Operationality – this criterion refers to the idea that it should be possible to use the knowledge and approaches proposed in the assessed guidelines for day-to-day management (Jeanmougin et al. 2017, Choulak et al. 2019).
- Legitimacy – this criterion refers to the fact that, because managers of PAs devise and implement actions that can, in some cases, conflict with other public policies, such as urban planning or economic development (and, in some cases, even use public financial and human resources for that purpose), conservation policies and actions should be acceptable to stakeholders (Meinard 2017, Arpin & Cosson 2021).

Following our demonstration that this framework is relevant to evaluating conservation guidelines, we then illustrate this relevance by developing a pilot application to a particular guide for PA managers: the French 'Guide for the elaboration of management plans for natural areas' (http://ct88.espaces-naturels.fr). These guidelines were introduced by the French Biodiversity Agency, an institution entrusted by the central government with orchestrating all of the conservation policies devised and implemented at the national scale. Its purpose is to provide a single reference supporting in a coherent manner the work of all of the actors involved in elaborating management plans for PAs in France.

Materials and methods

Literature review

To identify a framework for evaluating conservation guidelines, we performed a literature review of studies devoted to evaluating conservation documents – not only guidelines but also management plans and programmes. This literature review uses a quantitative approach to capture relevant contributions and then analyses them in an interpretative approach based on a thorough reading of the selected papers and an interpretation of their content. It thereby combines the strengths of both quantitative methods and social-science interpretative reasoning.

As explained above, our hypotheses, to be tested thanks to this literature review, were (1) that there is no commonly accepted framework available in the academic literature to perform the kind of evaluation needed, and (2) that the three 'policy analytics' criteria encompass all of the relevant approaches available in the literature despite their diversity, and therefore constitute a general and robust evaluation framework.

The literature review was conducted using a standard four-stage process (Barreto et al. 2020): (1) definition of the objectives guiding the review; (2) definition of the search protocol (database and search terms); (3) selection of articles based on predetermined criteria; and (4) analysis of the selected literature.

For this bibliographic research, we used the Web of Science (WoS) core collection database, one of the two main publication databases currently used by academic researchers (Pranckutė 2021). Several other databases could have been chosen, the most prominent being Google Scholar and Scopus. The former was excluded because, although it appears to have a wider coverage, it includes both academic publications and numerous other resources, such as unpublished reports or manuscripts. If the point had been to identify a diversity of contributions, including the grey literature, this would have been an asset. However, as explained in the 'Introduction' section, the hypotheses that we were concerned to test are only focused on the academic scientific literature, which made Google Scholar inappropriate. Scopus's coverage is also considered broader than WoS's according to recent analyses; however, for searches based on keywords, such as those we wanted to perform (see below), WoS is considered more efficient (Pranckutė 2021). We therefore worked with WoS; a comparative analysis using several databases might have yielded pertinent results but fell beyond our scope (see 'Discussion' section).

An initial search was done on the Web of Science core collection database on 6 February 2023, for the period before 2021, using the following request: [protected areas* AND (management OR ecological restoration*) AND (assessment* OR evaluation* OR analysis) AND (guide* OR manual OR tool OR plan)] on abstracts,

titles and keywords. The set of articles obtained was then manually screened in an interpretative, social-science approach to select all of those articles that contained evaluation criteria liable to constitute a usable framework for evaluating conservation guidelines. This second step in the search was designed to eliminate articles (1) devoted to evaluating stakeholders' perceptions of specific documents rather that evaluating the documents themselves, (2) focusing on specific, limited aspects of the document at issue without proposing criteria for evaluating management documents as a whole and (3) presenting an analysis of the management document on the basis of the topics covered without proposing transposable evaluation criteria.

For all of the articles selected using this procedure, we then identified the criteria on which their evaluations were based. In many cases, these criteria were not explicitly stated as such, and the identification was therefore to some extent interpretative. We then reformulated these criteria as synthetic questions. Lastly, the formulation of these synthetic questions was screened to identify keywords associated with the various criteria constituting the policy analytics framework. Lists of keywords were not defined ex ante but rather elaborated as we went along with the interpretation of criteria. Some keywords could refer to several of the policy analytics criteria; in these cases, the larger context provided by whole sentences articulating the criteria was used to identify interpretatively the policy analytics criterion or criteria (if any) to which the different occurrences of these keywords refer in each case. Similarly, an interpretative reading of the whole sentence was used when several keywords referring to different policy analytics criteria were present in the formulation of a single criterion from the literature at issue. This interpretative process eventually allowed us to determine whether the various criteria are encompassed or not in one policy analytics criterion or criteria.

Application

Based on the results from the bibliographic analysis, we then applied the identified relevant evaluation framework to the latest version of the French guidelines to develop management plans for PAs ('Guide for the elaboration of management plans for natural areas'; http://ct88.espaces-naturels.fr). These guidelines were chosen because they represent an attempt at orchestrating all of the conservation policies devised and implemented at a relatively large scale (that of the whole of France). This analysis illustrates the applicability of our framework and points to the strengths and weaknesses of this particular document. This analysis also enables us to suggest means to improve this document.

Results

Article selection

The initial search yielded 3593 articles (Table S1); however, the ensuing interpretative selection procedure filtered out 3204 of them that in fact do not tackle the evaluation of conservation documents. Although they do perform evaluations of conservation initiatives or documents, 367 of them failed to propose transferable criteria and 60 focus only on effectiveness. In the end, only 22 articles provide a possibly transferable framework based on clearly articulated evaluation criteria other than effectiveness (Table S1). This first result echoes the intrinsic difficulty of evaluating guidelines.

The 22 articles finally selected are mainly relatively recent contributions to the literature (63% (n = 14) published after 2013).

These articles are published in 10 journals, the most frequent being *Environmental Management* (n = 5) and *Ocean and Coastal Management* (n = 4). Articles with case studies are the most frequent, and they concern 15 countries, the most represented being France (n = 6), Spain (n = 4), Portugal (n = 4) and England (n = 4). They concern 36 types of PAs, the most frequently covered being national parks (n = 8) and marine nature parks (n = 7). The conservation documents evaluated are mainly management plans (n = 13); others are work programmes, guides for the elaboration of management plans and monitoring programmes.

Identification and analysis of evaluation criteria

The criteria used in the various selected articles, rearticulated in synthetic questions, are presented in Table 1. Column C2 lists these synthetic questions, article by article (column C1). When similar criteria are shared by different articles, these articles are grouped together in column C1 (e.g., line L6, which groups three papers using the same criteria). If a commonly accepted set of evaluation criteria had been available, it would have appeared as a single cell or group of cells in column C2, attached to a single cell in column C1 grouping an important part of the population of papers (Table 1). This is not the case, as the most populous cell in column C1 contains only three papers, four cells contain two papers and 14 cells out of 18 contain only one paper. This first analysis allows validation of our first hypothesis, as it shows that no commonly accepted set of evaluation criteria currently exists.

The subsequent analysis, striving to identify whether the various proposed criteria can be encompassed in one or several of the policy analytics criteria, shows that the criteria used in the 22 articles can all be interpreted as special cases of the policy analytics criteria (Table 1, column C3). Among the 22 articles, 18 champion criteria that can be interpreted as variants of a general criterion of operationality. These criteria refer to requirements to take administrative, legal or financial constraints into account, to cogently organize human and material resources or to use relevant organizational tools. Nineteen of the 22 articles put forward criteria capturing aspects of legitimacy. These criteria mention the need to include various types of stakeholders, the importance of discussing and/or assessing values and the ways by which the public or different relevant communities were involved. Lastly, 12 articles promote criteria reflecting scientific credibility requirements. Such criteria mention the need to anchor management in updated knowledge and data, to implement relevant monitoring schemes or to use concepts and framework accepted in the scientific community.

Application of the evaluation framework

The application of the evaluation framework constituted by the scientific credibility, operationality and legitimacy criteria to the French 'Guide for the elaboration of management plans for natural areas' highlights considerable weaknesses with respect to the three criteria (see also Osorio et al. 2023, which expands on some of these issues).

In terms of scientific credibility, this analysis shows that:

- The French 'Guide for the elaboration of management plans for natural areas' promotes the use of Red Lists and similar species lists (p. 29 of the downloadable pdf file) without mentioning the uncertainties and biases affecting them (Yang et al. 2013, Beck et al. 2014, Meyer et al. 2015, McRae et al. 2017, Jarić et al. 2019; problem S1).

Table 1. Evaluation criteria used in the selected articles (excluding criteria that are irrelevant for our purpose; e.g., effectiveness). Italicized are keywords associated with the 'policy analytics' criteria.

C1. Selected article	C2. Questions synthetizing the evaluation criteria promoted by the selected article	C3. 'Policy analytics' criteria
L1. Alder (1996), p. 101	Is the management strategy perceived as *effective and flexible*, and does the management plan participate in educational *programmes*?	Op.
	Have *stakeholders* been *involved* in planning?	Le.
L2. Allen et al. (2019), pp. 12–14	Does the plan identify and prioritize *needs*? Does it *organize* communication amongst staff members? Does it explain the *organisational structure*? Does it identify *specific targets* to achieve objectives? Does it explain the long-term *financial outlook*? Does it explain the *allocation of expenditures and financial management practices*? Does it *ensure funding* for management activities? Does the plan foresee adequate *facilities* for the level of visitors? Does it state the specific biodiversity-related *objectives*? Is it *consistent* with the objectives? Does it *enshrine* long-term protection *in law*? Does it specify the *staff and financial resources* required and *skills needed* and provide *training and development opportunities* for staff? Does it include regular reviews of *staff performance and progress* in achieving objectives? Does it define a zoning system to *achieve objectives* and explain how different zones relate to other protected areas? Does it *specify means* of communication between field and office staff? Does it identify the ecological and socio-economic *data needed and means of collecting new data and systems to process and analyse them*? Does it ensure *effective communication* with local communities? Does it specify the *transport infrastructure, field equipment and personal facilities* needed? Does it ensure *maintenance and care of equipment*?	Op.
	Does the plan enable staff to *engage with local communities and other organizations*? Does it allow local communities to *participate* in decisions? Does it explain how *conflicts* are addressed?	Le.
	Does the plan contain a *comprehensive inventory* of natural and cultural resources, an *analysis* of and strategy for dealing with threats and pressures, actions *monitoring* the impact of uses, investigation of key ecological and social issues and an explanation of how the results of *research and monitoring* are incorporated into planning? Does it allow officials to access *scientific research and advice*?	Sc.
L3. Álvarez-Fernández et al. (2017), annex A	Does it include a threat analysis, a zoning *plan* and an *atlas* of the protected area? Does it use a *geographic information system*? Does it *report past results*? Does it define quantitative and qualitative *objectives* and *expected results*? If yes, are the objectives detailed? If yes, is there an associated *programme*? Is there a *budget* for each action? Does it describe the *management tools* to be used for biodiversity conservation and restoration, cultural heritage, recreation and economic activities, environmental education and awareness-raising? Does it include an action *plan* for cooperation with other protected areas? Does it provide *contact details* for the protected area manager? Does it include an overall *budget with details* for human resources, operational costs and equipment and different activities? Does it foresee *agreements* with other institutions *for control* missions? Is there a *specific process* for elaborating the plan? Does it describe regulation of activities, a monitoring action plan and control *tools* (warnings, fines, etc.)? Does it include current *staffing*, future *staffing needs*, future *staff training needs* and the *cost* of these needs? Does it define the duration and cost of *implementation* and *renewal*?	Op.
	Does it describe the conservation *value* of the protected area, the legal framework, the *governance organization* and inter-administrative arrangements for the management of the site? Does it contain a communication plan for *stakeholders*, a *conflict* analysis and a specific *validation* process? Does in include *stakeholder participation*? Does it specify who is *responsible* for the elaboration of the plan and who *approves* it?	Le.
	Does the plan contain an analysis of current *knowledge gaps*? Does the management plan describe the *study action plan*?	Sc.
L4. Álvarez-Fernández et al. (2020a), p. 10	Was there a *time lag* between designation of the protected area and implementation of the management plan? Is the plan based on specific *objectives*? Were management plans duly *renewed*?	Op.
	Were *stakeholders involved* in all critical phases of plan design and implementation?	Le.

(Continued)

Table 1. (Continued)

C1. Selected article	C2. Questions synthetizing the evaluation criteria promoted by the selected article	C3. 'Policy analytics' criteria
L5. Álvarez-Fernández et al. (2020b; based on Hocking et al. 2000)	Is the plan based on specific *objectives*? Does it provide *resources* for the management? Does it explain how *resources* will be used? Are *context* and *constraints* considered? Are relevant *policies* and *procedures* accounted for? Does the plan define which *partners* the project will work with? Does it identify which areas need specific attention to improve the capacity of managers to *carry out their work* (more resources, staff training, etc.)?	Op.
	Is the management plan based on an *assessment* of significant hazards, vulnerabilities and context?	Sc.
L6. Anthony and Shestackova (2015), PAME framework: Table 3; Ayivor et al. (2020), Lu et al. (2012), RAPPAM methodology: Ervin (2003)	Is the plan based on *effective administration*, *governance* and *leadership*? Does it include *adequate infrastructure*, *funding* and *human resources*? Does it contain a communication *programme*?	Op.
	Does the tool *involve* the *community* and *stakeholders*? Does the tool contain an adequate *community benefits/assistance* programme?	Le.
	Is the plan based on an *assessment* of significant threats and context? Does the management plan include *research* and *monitoring*?	Sc.
L7. Arpin (2019)	Does the management plan include a strategy for managing conflict?	Leg.
L8. Barker and Stockdale (2008), Table 1	Does the plan encourage consideration of different *sources* in addition to central government, and does it propose to explore 'visitor reimbursement' approaches to *raising funds* from tourism?	Op.
	Does the plan integrate *conflict* as inevitable? Does it promote the concentration of management efforts at the point of *interaction between people and place* to reduce the degree of conflict? Does it champion communication to encourage *accountability and participation*? Does it emphasize *co-management and the role of local people as stewards* of the landscape? Does it affirm *individual and community freedoms*? Does it encourage the creation of advisory groups to foster *political integration* and include a *partnership approach* (if relevant)?	Le.
L9. Claudet and Pelletier (2004), pp. 130–132	Is the management plan based on specific *objectives*? Does it include management *operations planning*?	Op.
	Is the management plan based on a *participatory* approach?	Le.
	Have the expected results been formalized and prioritized, and is the management plan *based on scientific knowledge*?	Sc.
L10. Ernoul et al. (2015), Table 1	Is the plan based on the evaluation of *previous plans*? Does it present the *legal context, clear* management *objectives, factors* influencing the ecological evolution of the site, the *identified activities*, a *results chain analysis*, the existing *infrastructure*, the socio-economic *activities* within and near the site, *interaction* between stakeholders and the natural heritage, an *organizational chart* and direct and indirect threats? Does it include a *hierarchy of threats and problems* (in terms of heritage, socio-economic issues and potential, environmental education, ecosystem services, existing planning scenarios and management objectives)? Does it analyse *interaction* between stakeholders and the natural heritage? Does it account for existing *human resources*, human resource *needs*, existing *funding* and *funding needs*? Does it include a *funding plan*? Does it account for *training*? Does it present *partners*? Does it include *SMART indicators*? Does it specify *data collection methods*, a data collection *plan* and means for *data storage* and *processing*? Does it foresee *knowledge and information needed* for communication? Does it ensure *funding* for communication?	Op.
	Does the tool present the *stakeholders*, management *authorities* and the *groups* likely to affect or influence the site? Does it include *stakeholders'* objectives, roles and responsibilities?	Le.
	Does the tool consider existing *knowledge* on natural heritage? Does it identify additional *knowledge needs*? Does it present the historical evolution of the site (ownership, land use, etc.)? Does it contain an overall *analysis* of the site and a description of the heritage? Does it explain *adaptive management*?	Sc.
L11. Hockings (1998), p. 340	Does the plan detail the *actions/policies to be implemented*? Does it include a management *information system*? Does it specify data requirements for indicators? Are the data needs of the whole strategy reviewed? Does the plan include *monitoring projects* based on actions/polices to be implemented? Are priority monitoring *programmes* selected and implemented?	Op.
	Does the plan explain the identification and use of information on reserve *values*?	Le.

(Continued)

Table 1. (Continued)

C1. Selected article	C2. Questions synthetizing the evaluation criteria promoted by the selected article	C3. 'Policy analytics' criteria
	Are the objectives used as a basis for developing *outcome evaluation*, and are the objectives reviewed to establish appropriate indicators of achievement?	Sc.
L12. Hockings et al. (2009), Table 4, Jones (1994), p. 152	Does the plan include a *work programme*? Does the information used in the plan include *visitors*? Does the plan contain a regular *maintenance programme*, a *monitoring* and *evaluation programme* and objectives that consider the *economic background*? Does the plan account for its adequacy with *management guidelines*? Does the plan explain how information is used to *support management decision-making* and how *law is abided by* in preventing illegal activities? Does it ensure the adequacy of visitor *facilities*? Does it explain how basic information is provided to *visitors* and how *interpretive and educational services* are provided? Does the plan foresee the *requirements necessary to achieve the objectives*?	Op.
	Does the information used in the plan account for the *Aboriginal heritage* (sites and places) *and historical heritage*? Is the management of Aboriginal sites and historical heritage one of the key issues addressed in the plan? Does the plan explain how the *Aboriginal community* is consulted and how the *wider community* is consulted? Does the plan account for the impact of park *values* on natural features, on Aboriginal sites and on historical heritage? Does the management plan contain objectives that consider the *ethical background*?	Le.
	Does the management plan contain objectives that consider the *scientific background*?	Sc.
L13. Kovács et al. (2017), Table 2	Is planning based on a *participatory* process?	Le.
L14. Maestro et al. (2020), Tables 5–7, Morris et al. (2014), p. 44	Is there an *operational* plan, a sustainable *education programme* and a *communication programme*? Was a *diagnosis* of the ecosystem carried out before drawing up the management plan?	Op.
	Was *the public involved*? Was *public participation* representative? Are *stakeholders involved*? Is information available to *stakeholders* and the *public*? Is the plan *perceived* to be effective?	Le.
	Is the management plan based on an *identification of the problems* at the site? Is the *necessary funding* secured?	Op.
	Is the *governance* structured and effective and are conditions in place to ensure *stakeholder participation*?	Le.
L15. Muñoz and Hausner (2013), p. 2380	Does the management plan have clear *objectives*, and is its *implementation* supported by appropriate regulations/legislation?	Op.
L16. Scianna et al. (2018), p. 177	Does the management plan specify the necessary *competences* of the management body? Does it include a *monitoring system*?	Op.
	Is the management plan associated with a specialized *management agency*?	Le.
	Does the management plan propose an *evaluation of results*?	Sc.
L17. Stori et al. (2019), p. 333	Is the management plan *consistent with other frameworks and regulations* in place in the area? Is the governance of the site *integrated with other relevant frameworks and regulations*?	Op.
L18. Wyatt et al. (2011), p. 2256	Were *Aboriginal Nations involved*, and did they have an effective influence on the development of the management plan?	Le.

Le. = legitimacy; Op. = operationality; PAME = Protected Area Management Effectiveness; RAPPAM = Rapid Assessment and Prioritization of Protected Area Management; Sc. = scientific credibility; SMART = Specific, Measurable, Achievable, Relevant and Time-bound.

- It wrongly assumes (p. 29) that experts in the field can perform analyses of ecosystem functioning (Pe'er et al. 2014, Jeanmougin et al. 2017, Troudet et al. 2017, Jarić et al. 2019, Sutherland et al. 2019; problem S2).
- It ignores that, according to the academic literature in conservation, assessing representativeness is a major global challenge (p. 29; Anthamatten & Hazen 2007), mainly because of a lack of data in inventories (Fedorov et al. 2020; problem S3).
- It ignores difficulties in choosing how to aggregate various dimensions or criteria to produce overall assessments of the value of natural sites (p. 30; e.g., Schwartz et al. 2018, Choulak et al. 2019; problem S4).

- It articulates recommendations on how to frame objectives that are at odds with the acknowledged importance of assessing the achievement of targets (p. 35; e.g., Ferraro & Pattanayak 2006; problem S5).
- It ignores the literature highlighting the need to assess the influence of external factors (p. 35; Holling 1996, Apitz 2008, Santos & Schiavetti 2014, Bennett et al. 2017, Sendzimir et al. 2018; problem S6).
- It fails to promote the ongoing flexibility and adaptation of practices as well as the cooperation between experts, scientists and managers and mutual learning emphasized in the literature on adaptive management (one single reference to adaptive management, p. 39, without explanation nor

operational details; Folke et al. 2005, Bormann et al. 2007, Ananda & Proctor 2013; problem S7).

- It downplays the difficulties in choosing or constructing indicators (pp. 42–43; Bouyssou et al. 2000, Hallam et al. 2020; problem S9).
- It ignores the literature on the importance and complexity of stakeholder identification and participatory processes (p. 62; Luyet et al. 2012, Paletto et al. 2015, Kovács et al. 2017; problem S10).

In terms of operationality:

- The French 'Guide for the elaboration of management plans for natural areas' fails to discuss operational procedures for assessing representativeness (p. 29; Mingarro & Lobo 2018, Fedorov et al. 2020, Milla-Figueras et al. 2020; problem O1).
- It fails to explain how the analysis of 'influencing' or 'stress' factors should be carried out (p. 36; problem O2).
- It fails to explain how managers should choose indicators to structure monitoring and evaluation (p. 42; problem O3).
- It fails to explain how stakeholders should be identified and recruited (p. 62; Paletto et al. 2015; problem O4).

In terms of legitimacy:

- The French 'Guide for the elaboration of management plans for natural areas' fails to discuss the various actors' responsibilities and strategies as well as actions to strengthen accountability (p. 39; problem L1).
- It fails to justify the key choices underlying the definition it gives to operational objectives (p. 35; problem L2).
- It fails to promote discussions on the values underlying the tools used, such as Red Lists and similar species lists (p. 29; problem L3).
- It promotes the search for consensus (p. 62), thereby ignoring that consensus-seeking can nullify the possibility of debating different positions without having to resort to violence, prevent an in-depth analysis of conflicts and obscure the hegemony of certain actors (Mouffe 2005, Arpin 2019; problem L4).

Recommendations

The literature suggests that the weaknesses identified by our evaluation can all be addressed by implementing relevant participatory processes involving both local communities and a diversity of knowledge-holders, including experts and scientists. Indeed, by involving scientific experts, participation can help strengthen scientific robustness (scientific credibility), and the co-construction with local actors and operational workers can help fix operational problems (operationality). In addition, the inclusion of stakeholders with diverse views and values can strengthen legitimacy by initiating constructive discussions on values (García-Montes & Monreal 2019) and, depending on the specific situation, either by enabling stakeholders to build a shared vision of the future (Santana-Medina et al. 2013) or by enabling the open acknowledgement of irreducible disagreements.

The fact that guidelines such as those analysed here are plagued by problems that participatory processes can fix shows that participation, although routinely and repeatedly referred to in guidelines, is insufficiently dealt with in such documents, which underestimate the difficulty of setting up and implementing participatory processes (Osorio et al. 2023).

Discussion

The main result of this analysis is that the relevant academic literature in conservation is sparse and heterogeneous, but a relevant encompassing framework is provided by the literature in decision sciences on the 'policy analytics' framework. Like most scientific studies based on literature reviews, this analysis admittedly neglects the grey literature, because the latter is excluded from large-scale homogeneous bibliographic databases such as the one used here. However, as explained above, excluding the grey literature is justified when the aim is to identify frameworks for which the robustness is buttressed in the academic scientific literature.

In addition, most of the articles analysed in Table 1 refer to and are based on important contributions to the grey literature, which are duly referred to. This suggests that our analysis indirectly encompasses at least part of the relevant grey literature. That said, the grey literature certainly contains other useful frameworks that are ignored by the academic scientific literature. This conjecture suggests that academic scientific evaluations of such contributions to the grey literature are needed to entrench their scientific credentials and, incidentally, to increase their visibility. A systematic review of evaluation frameworks published in the grey literature and a systematic meta-evaluation of their scientific credentials would accordingly be major contributions. Dedicated methodologies will have to be devised for that purpose, as identifying and screening the grey literature involves numerous major challenges. All of this falls beyond the scope of the present paper.

Comparing our results with those obtained using other large-scale bibliographic databases, such as Scopus, could also bring complementary insights. However, a similar analysis of Scopus could not possibly invalidate our key message, according to which there is no dominant evaluation framework for conservation guidelines in the academic literature. Indeed, although Scopus is known to be more extensive in some domains, even if all of the records included in Scopus but not Web of Science were to share a unique framework, which seems unlikely, such a framework would not dominate the Scopus plus Web of Science corpus.

Another improvement that future studies could take upon themselves is to test the robustness of the interpretative steps of our analyses. We characterize as 'interpretative' the operations that consisted in reformulating criteria in synthetic questions and in identifying keywords referring to the various policy analytics criteria. Empirical robustness tests could be implemented by asking a diverse set of experts to propose their own reformulations and keywords.

Another, possibly more promising refinement of our analysis would be to test whether the 'policy analytics' criteria can be rendered more precise whilst retaining their ability to encompass the criteria we identified in the scientific literature. Indeed, a plausible criticism that could be raised against our approach is that the 'policy analytics' criteria are exceedingly vague, and that this vagueness alone explains why they encompass all of the criteria proposed in the literature. This suspected vagueness of the framework has been discussed in the literature in the decision sciences and management (e.g., Meinard et al. 2021), with proposals given of more precise definitions of especially complex

concepts, such as legitimacy. This literature can be used to identify directions for testing more precise variants of the framework.

The second task performed in this study consisted in applying the three criteria of legitimacy, operationality and scientific credibility to specific guidelines for managers of PAs. This application illustrates that, although the criteria proposed in our framework are arguably more abstract than those identified in the conservation literature, this abstractness does not come at the expense of applicability. The main conclusion of the application was that the evaluated guidelines are plagued by significant weaknesses that could be overcome by implementing relevant participatory processes. Some initiatives arguably go in the direction of implementing participation that might be able to address the kind of problems that we pinpointed in this analysis. For example, the German procedure to draw up management plans for Natura 2000 sites (e.g., in Baden-Württemberg State, Germany; https://pd.lubw.de/69643) involves the wide diffusion of preliminary layouts of management plans associated with public hearings, on-site debates with stakeholders and websites presenting management actions. However, the associated guidelines do not detail how such mechanisms should be chosen and implemented. This loophole echoes the multiple weaknesses in the application of participation that generally plague current PA management in Europe (Piwowarczyk & Wróbel 2016, Kovács et al. 2017, Álvarez-Fernández et al. 2020a, 2020b). The lesson learnt from our analysis of management guidelines hence appears to hold true more generally for a vast array of conservation policy tools.

However, the very idea that participation should be encouraged in conservation decision-making, which constitutes the backbone of our recommendations, is not without its critics. Indeed, participation does not always strengthen conservation (Young et al. 2013): it increases the time needed to develop management strategies and their costs (Paletto et al. 2015), and it can be used as a manipulative tool to reproduce unequal power relations and reinforce the dominance of certain forms of knowledge (Turnhout et al. 2020). To overcome such problems, Osorio et al. (2022) champion 'counter-argumentative participation', defined as a process by which different stakeholders influence decision-making by expressing criticisms and counter-arguments. How such recommendations can be integrated into conservation guidelines such as those analysed here remains to be formally established; so is the extent to which they can solve the problems facing conservation practitioners in the field.

References

Alder J (1996) Have tropical marine protected areas worked? An initial analysis of their success. *Coastal Management* 24: 97–114.

Allen L, Holland KK, Holland H, Tome' S, Nabaala M, Seno S, et al. (2019) Expanding staff voice in protected area management effectiveness assessments within Kenya's Maasai Mara National Reserve. *Environmental Management* 63: 46–59.

Álvarez-Fernández I, Fernández N, Sánchez-Carnero N, Freire J (2017) The management performance of marine protected areas in the north-east Atlantic Ocean. *Marine Policy* 76: 159–168.

Álvarez-Fernández I, Freire J, Naya I, Fernández N, Sánchez-Carnero N (2020a) Failures in the design and implementation of management plans of marine protected areas: an empirical analysis for the North-east Atlantic Ocean. *Ocean & Coastal Management* 192: 105178.

Álvarez-Fernández I, Freire J, Sánchez-Carnero N (2020b) Low-quality management of marine protected areas in the north-east Atlantic. *Marine Policy* 117: 103922.

Ananda J, Proctor W (2013) Collaborative approaches to water management and planning: an institutional perspective. *Ecological Economics* 86: 97–106.

Anthamatten P., Hazen, H. (2007) Unnatural selection: an analysis of the ecological representativeness of natural world heritage sites. *The Professional Geographer* 59: 256–268. https://doi.org/10.1111/j.1467-9272.2007.00611.x

Anthony BP, Shestackova E (2015) Do global indicators of protected area management effectiveness make sense? A case study from Siberia. *Environmental Management* 56: 176–192.

Apitz SE (2008) Adaptive management principles and sediment management: a call for discussions. *Journal of Soils and Sediments* 8: 359–362.

Arlettaz R, Schaub M, Fournier J, Reichlin TS, Sierro A, Watson JEM, et al. (2010) From publications to public actions: when conservation biologists bridge the gap between research and implementation. *BioScience* 60: 835–842.

Arpin I (2019) The rise of planning in nature conservation and the practitioners' approach to conflicts. The inspiring case of the northern French Alps nature reserves. *Journal for Nature Conservation* 48: 54–60.

Arpin I, Cosson A (2021) Seeking legitimacy in European biodiversity conservation policies: the case of French national parks. *Environmental Science & Policy* 116: 181–187.

Ayivor JS, Gordon C, Tobin GA, Ntiamoa-Baidu Y (2020) Evaluation of management effectiveness of protected areas in the Volta Basin, Ghana: perspectives on the methodology for evaluation, protected area financing and community participation. *Journal of Environmental Policy & Planning* 22: 239–255.

Barker A, Stockdale A (2008) Out of the wilderness? Achieving sustainable development within Scottish national parks. *Journal of Environmental Management* 88: 181–193.

Barreto GC, Di Domenico M, Medeiros RP (2020) Human dimensions of marine protected areas and small-scale fisheries management: a review of the interpretations. *Marine Policy* 119: 104040.

Baylis K, Honey-Rosés J, Börner J, Corbera E, Ezzine-de-Blas D, Ferraro PJ, et al. (2015). Mainstreaming impact evaluation in nature conservation. *Conservation Letters* 9: 58–64.

Beck J, Böller M, Erhardt A, Schwanghart W (2014) Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics* 19: 10–15.

Bennett NJ, Roth R, Klain SC, Chan K, Christie P, Clark DA, et al. (2017) Conservation social science: understanding and integrating human dimensions to improve conservation. *Biological Conservation* 205: 93–108.

Bormann BT, Haynes RW, Martin JR (2007) Adaptive management of forest ecosystems: did some rubber hit the road? *BioScience* 57: 186–191.

Bottrill MC, Hockings M, Possingham HP (2011) In pursuit of knowledge: addressing barriers to effective conservation evaluation. *Ecology and Society* 16: 14.

Bottrill MC, Pressey RL (2012) The effectiveness and evaluation of conservation planning: evaluation and conservation planning. *Conservation Letters* 5: 407–420.

Bouyssou D, Marchant T, Pirlot M, Perny P, Tsoukiàs A, Vincke P (2000) *Evaluation and Decision Models: A Critical Perspective.* Boston, MA, USA: Springer US.

Choulak M, Marage D, Gisbert M, Paris M, Meinard Y (2019) A meta-decision-analysis approach to structure operational and legitimate environmental

policies – with an application to wetland prioritization. *Science of the Total Environment* 655: 384–394.

Claudet J, Pelletier D (2004) Marine protected areas and artificial reefs: a review of the interactions between management and scientific studies. *Aquatic Living Resources* 17: 129–138.

Courrau J, Dudley N, Hockings M, Leverington F, Stolton S (2006) *Evaluating Effectiveness: A Framework for Assessing Management Effectiveness of Protected Areas*. Gland, Switzerland: IUCN.

De Marchi G, Lucertini G, Tsoukiàs A (2016) From evidence-based policy making to policy analytics. *Annals of Operations Research* 236: 15–38.

Dubois NS, Gomez A, Carlson S, Russell D (2020) Bridging the research–implementation gap requires engagement from practitioners. *Conservation Science and Practice* 2: e134.

Ernoul L, Beck N, Cohez D, Perennou C, Thibault M, Willm L, et al. (2015) Trends in management plans and guides: 25 years of experience from southern France. *Journal of Environmental Planning and Management* 58: 1096–1112.

Ervin J (2003) WWF rapid assessment and prioritization of protected area management (RAPPAM) methodology. Gland, Switzerland: WWF [www document]. URL https://wwfeu.awsassets.panda.org/downloads/rappam.pdf

Evans D (2012) Building the European Union's Natura 2000 network. *Nature Conservation* 1: 11–26.

Fedorov NI, Muldashev AA, Martynenko VB, Baisheva EZ, Shirokikh PS, Elizaryeva OA, et al. (2020) Identifying highly diverse areas of rare plant species as a basis for assessing representativeness and improving the network of protected areas. *Contemporary Problems of Ecology* 13: 418–428.

Ferraro PJ, Pattanayak SK (2006) Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *PLoS Biology* 4: e105.

Folke C, Hahn T, Olsson P, Norberg J (2005) Adaptive governance of social-ecological systems. *Annual Review of Environment and Resources* 30: 441–473.

García-Montes N, Arnanz Monreal L (2019) Metodologías participativas para la planificación de la sostenibilidad ambiental local. El caso de la Agenda 21. *Empiria. Revista de metodología de ciencias sociales* (epub ahead of print) doi: 10.5944/empiria.44.2019.25354.

Geldmann J, Barnes M, Coad L, Craigie ID, Hockings M, Burgess ND (2013) Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. *Biological Conservation* 161: 230–238.

Grantham HS, Bode M, McDonald-Madden E, Game ET, Knight AT (2009) Effective conservation planning requires learning and adaptation. *Frontiers in Ecology and the Environment* 8: 431–437.

Hallam CD, Wintle BA, Kujala H, Whitehead AL, Nicholson E (2020) Measuring impacts on species with models and metrics of varying ecological and computational complexity. *Conservation Biology* 34: 1512–1524.

Hockings M (1998) Evaluating management of protected areas: integrating planning and evaluation. *Environmental Management* 22: 337–345.

Hockings M, Cook CN, Carter RW, James R (2009) Accountability, reporting, or management improvement? Development of a state of the parks assessment system in New South Wales, Australia. *Environmental Management* 43: 1013–1025.

Hockings M, Stolton S, Dudley N (2000) *Evaluating Effectiveness: A Framework for Assessing the Management of Protected Areas. Gland*, Switzerland and Cambridge, UK: IUCN [www document]. URL https://portals.iucn.org/library/efiles/documents/pag-014.pdf

Holling CS (1996) Surprise for science, resilience for ecosystems, and incentives for people. *Ecological Applications* 6: 733–735.

Jarić I, Quétier F, Meinard Y (2019) Procrustean beds and empty boxes: on the magic of creating environmental data. *Biological Conservation* 237: 248–252.

Jeanmougin M, Dehais C, Meinard Y (2017) Mismatch between habitat science and habitat directive: lessons from the French (counter) example: evaluating the European habitat policy. *Conservation Letters* 10: 634–644.

Jones PJS (1994) A review and analysis of the objectives of marine nature reserves. *Ocean & Coastal Management* 24: 149–178.

Knight AT, Cowling RM, Rouget M, Balmford A, Lombard AT, Campbell BM (2008) Knowing but not doing: selecting priority conservation areas and the research–implementation gap. *Conservation Biology* 22: 610–617.

Kovács E, Kelemen E, Kiss G, Kalóczkai Á, Fabók V, Mihók B, et al. (2017) Evaluation of participatory planning: lessons from Hungarian Natura 2000

management planning processes. *Journal of Environmental Management* 204: 540–550.

Linares Quero A, Iragui Yoldi U, Gava O, Schwarz G, Povellato A, Astrain C (2022) Assessment of the Common Agricultural Policy 2014–2020 in supporting agroecological transitions: a comparative study of 15 cases across Europe. *Sustainability* 14: 9261.

Lu D-J, Kao C-W, Chao C-L (2012) Evaluating the management effectiveness of five protected areas in Taiwan using WWF's RAPPAM. *Environmental Management* 50: 272–282.

Luyet V, Schlaepfer R, Parlange MB, Buttler A (2012) A framework to implement stakeholder participation in environmental projects. *Journal of Environmental Management* 111: 213–219.

Maestro M, Chica-Ruiz JA, Pérez-Cayeiro ML (2020) Analysis of marine protected area management: the Marine Park of the Azores (Portugal). *Marine Policy* 119: 104104.

McRae L, Deinet S, Freeman R (2017) The diversity-weighted living planet index: controlling for taxonomic bias in a global biodiversity indicator. *PLoS ONE* 12: e0169156.

Meinard Y (2017) What is a legitimate conservation policy? *Biological Conservation* 213: 115–123.

Meinard Y, Barreteau O, Boschet C, Daniell KA, Ferrand N, Girard S, et al. (2021) What is policy analytics? An exploration of 5 years of environmental management applications. *Environmental Management* 67: 886–900.

Meyer C, Kreft H, Guralnick R, Jetz W (2015) Global priorities for an effective information basis of biodiversity distributions. *Nature Communications* 6: 8221.

Milla–Figueras D, Schmiing M, Amorim P, Horta e Costa B, Afonso P, Tempera F (2020) Evaluating seabed habitat representativeness across a diverse set of marine protected areas on the Mid–Atlantic Ridge. *Biodiversity and Conservation* 29: 1153–1175.

Mingarro M, Lobo JM (2018) Environmental representativeness and the role of emitter and recipient areas in the future trajectory of a protected area under climate change. *Animal Biodiversity and Conservation* 41: 333–344.

Morris RKA, Bennett T, Blyth-Skyrme R, Barham PJ, Ball A (2014) Managing Natura 2000 in the marine environment – an evaluation of the effectiveness of 'management schemes' in England. *Ocean & Coastal Management* 87: 40–51.

Mouffe C (2005) *On the Political*, 1st edition. Abingdon, UK: Routledge.

Muñoz L, Hausner V (2013) What do the IUCN categories really protect? A case study of the alpine regions in Spain. *Sustainability* 5: 2367–2388.

Osorio A, Schmitt L, Badariotti D, Meinard Y (2022) Mise en œuvre d'un processus de « participation contre-argumentative » dans la gestion et la restauration des milieux fluviaux : retour d'expérience dans une Réserve Naturelle Nationale rhénane. *Géocarrefour* 96: 19984.

Osorio A, Schmitt L, Badariotti D, Meinard Y (2023) Améliorer la gestion des espaces naturels par la participation. Une analyse du guide français pour l'élaboration des plans de gestion. *VertigO* (epub ahead of print) doi: 10.4000/vertigo.40230.

Paletto A, Hamunen K, De Meo I (2015) Social network analysis to support stakeholder analysis in participatory forest planning. *Society & Natural Resources* 28: 1108–1125.

Pe'er G, Mihoub J-B, Dislich C, Matsinos Y (2014) Towards a different attitude to uncertainty. *Nature Conservation* 8: 95–114.

Pearson DE, Clark TJ, Hahn PG (2022) Evaluating unintended consequences of intentional species introductions and eradications for improved conservation management. *Conservation Biology* 36: e13734.

Piwowarczyk J, Wróbel B (2016) Determinants of legitimate governance of marine Natura 2000 sites in a post-transition European Union country: a case study of Puck Bay, Poland. *Marine Policy* 71: 310–317.

Pranckutė R (2021) Web of Science (WoS) and Scopus: the titans of bibliographic information in today's academic world. *Publications* 9: 12.

Sanderson FJ, Pople RG, Ieronymidou C, Burfield IJ, Gregory RD, Willis SG et al. (2015). Assessing the performance of EU nature legislation in protecting target bird species in an era of climate change: impacts of EU nature legislation. *Conservation Letter*, 9, 172–180.

Santana-Medina N, Franco-Maass S, Sánchez-Vera E, Imbernon J, Nava-Bernal G (2013) Participatory generation of sustainability indicators in a natural protected area of Mexico. *Ecological Indicators* 25: 1–9.

Santos CZ, Schiavetti A (2014) Assessment of the management in Brazilian marine extractive reserves. *Ocean & Coastal Management* 93: 26–36.

Schwartz MW, Cook CN, Pressey RL, Pullin AS, Runge MC, Salafsky N, et al. (2018) Decision support frameworks and tools for conservation: decision support for conservation. *Conservation Letters* 11: e12385.

Scianna C, Niccolini F, Bianchi CN, Guidetti P (2018) Applying organization science to assess the management performance of marine protected areas: an exploratory study. *Journal of Environmental Management* 223: 175–184.

Sendzimir J, Magnuszewski P, Gunderson L (2018) Adaptive management of riverine socio-ecological systems. In S Schmutz, J Sendzimir (eds), *Riverine Ecosystem Management: Science for Governing towards a Sustainable Future* (pp. 301–324). New York, NY, USA: Springer International Publishing.

Stolton S, Dudley N, Belokurov A, Deguignet M, Burgess ND, Hockings M, et al. (2019) Lessons learned from 18 years of implementing the Management Effectiveness Tracking Tool (METT): a perspective from the METT developers and implementers. *PARKS* 25: 79–92.

Stori FT, Shinoda DC, Turra A (2019) Sewing a blue patchwork: an analysis of marine policies implementation in the southeast of Brazil. *Ocean & Coastal Management* 168: 322–339.

Sutherland WJ, Taylor NG, MacFarlane D, Amano T, Christie AP, Dicks LV, et al. (2019) Building a tool to overcome barriers in research–implementation spaces: the Conservation Evidence database. *Biological Conservation* 238: 108199.

Troudet J, Grandcolas P, Blin A, Vignes-Lebbe R, Legendre F (2017) Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports* 7: 9132.

Turnhout E, Metze T, Wyborn C, Klenk N, Louder E (2020) The politics of co-production: participation, power, and transformation. *Current Opinion in Environmental Sustainability* 42: 15–21.

Watson JEM, Dudley N, Segan DB, Hockings M (2014) The performance and potential of protected areas. *Nature* 515: 67–73.

Wyatt S, Merrill S, Natcher D (2011) Ecosystem management and forestry planning in Labrador: how does Aboriginal involvement affect management plans? *Canadian Journal of Forest Research* 41: 2247–2258.

Yang W, Ma K, Kreft H (2013) Geographical sampling bias in a large distributional database and its effects on species richness–environment models. *Journal of Biogeography* 40: 1415–1426.

Young JC, Jordan A, Searle KR, Butler A, Chapman DS, Simmons P, Watt AD (2013) Does stakeholder involvement really benefit biodiversity conservation? *Biological Conservation* 158: 359–370.