



ELSEVIER

Contents lists available at ScienceDirect

Marine Policy

journal homepage: www.elsevier.com/locate/marpol

Modeling expert judgment to assess cost-effectiveness of EU Marine Strategy Framework Directive programs of measures



A. Kontogianni ^{a,*}, C. Tourkolias ^b, D. Damigos ^c, M. Skourtos ^d, B. Zanou ^e

^a University of Western Macedonia, Kozani, Greece

^b Center for Renewable Energy Sources, Pikermi, Greece

^c National Technical University, Athens, Greece

^d Agricultural University of Athens, Athens, Greece

^e Hellenic Center for Marine Research, Anavyssos, Greece

ARTICLE INFO

Article history:

Received 27 January 2015

Received in revised form

8 August 2015

Accepted 2 September 2015

Available online 29 September 2015

Keywords:

Marine Strategy Framework Directive

Cost-effectiveness

Expert judgment

Marine economics

ABSTRACT

The EU Marine Strategy Framework Directive (MSFD) requires all Member States to establish a program of measures to achieve or maintain Good Environmental Status (GES) of their marine waters, which should be justified on economic grounds. So far, however, only limited efforts exist to support, from a scientific perspective, marine policy- and decision-makers to this direction. This paper describes a first effort towards closing this gap and improving existing marine policymaking processes as regards the prioritization and selection of measures and policies towards coastal and marine resources management. More specifically, the paper presents an expert judgment-based weighting framework named 'MeTaLi'. The tool provides a cost-effectiveness ranking algorithm of alternative measures (e.g. command-and-control, economic, etc.) within the framework of MSFD by means of fuzzy and stochastic analysis. A pilot application of 'MeTaLi' in Greece for three selected MSFD descriptors is also discussed, aiming to evaluate the tool and allow drawing conclusions for real conditions. Finally, the paper concludes with a discussion of research findings and methodological challenges related to marine policy issues.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

It is widely acknowledged that the EU Marine Strategy Framework Directive (MSFD) is strongly framed within the economic logic. MSFD includes economic requirements either explicitly (i.e. areas where economic analysis is clearly demanded) or implicitly (i.e. areas where economic analysis may be beneficial in meeting the requirements although this is not a prerequisite) [1]. For instance, Art 8(1c) of the MSFD requires all Member States to undertake "an economic and social analysis of the use of those waters and of the cost of degradation of the marine environment" within their Initial Assessment reports. The economic linkages are even more intense in Article 13 of the MSFD, which expects the Member States to design and implement a program of measures to achieve good environmental status in their marine waters. These measures, implemented at different spatial levels (local/national/regional or global), may include a number of different instruments [2]:

- Traditional command and control (CAC) or regulatory instruments (e.g. regulation, norms and standards, bans) that have a

direct influence on the behavior of actors by imposing rules that limit or prescribe the actions of the target group.

- Economic instruments (e.g. fees, subsidies, liability and compensation regimes, trading systems) that modify the behavior and decisions of actors by changing the cost or price of a market good (e.g. plastic bags), service (e.g. waste collection), activity (e.g. waste dumping), input (e.g. materials), or output (e.g. pollution).
- Social instruments, which are based on voluntary aspect of actions and influence the behavior of actors and individuals indirectly. For instance, polluters are stimulated to take actions through awareness raising campaigns.
- Technical, technological and research-oriented measures, e.g. removal of man-made constructions, monitoring activities, etc. If there is an obligation to imply a certain technical measure, it should be regarded as a regulatory instrument. If the implementation of a technical measure is encouraged by subsidies, it should be regarded as an economic instrument. If an information campaign promotes the application of the technical measure, it should be regarded as a social instrument. Thus, it is sometimes difficult to categorize a measure as a technical measure or as a regulatory or economic measure.

Regardless of the measures selected for achieving or maintaining Good Environmental Status (GES) of marine waters, decision-makers

* Corresponding author.

E-mail address: akontogianni@uowm.gr (A. Kontogianni).

have to define the costs of measures and to estimate the potential benefits in order to evaluate the proposed program and the proportionality of costs by means of cost-effectiveness or cost-benefit analyses [3]. In general, cost-benefit analysis is suitable when the targets have not yet been set because it can be used to determine if the benefits of the possible targets are higher than the costs, thereby informing what the target should be [1]. Cost-effectiveness, on the other hand, is a more suitable approach to use when the objective has been established and the analysis focuses on the best way to meet the target. The latter seems to be the most relevant methodology in the context of Article 13 of the MSFD, where the objectives have already been established [1].

Considering that the MSFD sets out eleven qualitative descriptors covering broad topics, there are many challenges ahead towards achieving GES of the European marine environment. There is therefore a need for further scientific understanding for the criteria and indicators determining the descriptors. A major challenge is certainly the development of appropriate tools and approaches that would make use of best available scientific knowledge and would facilitate the implementation of the MSFD within the prescribed timeline. Among them, tools that will help Member States to draw up their Program of Measures to achieve GES are of particular importance provided that these programs are scheduled to be developed by 2015 at the latest.

To this end, this paper presents and discusses 'MeTali', a tool developed within the EU funded MERMAID project (*Marine Environmental targets linked to Regional Management schemes based on Indicators Developed for the Mediterranean*). The MERMAID aims to provide scientific understanding for assessing GES in a coherent and holistic manner. The project will develop a state-of-the-art methodology that will be tested for five selected MSFD descriptors, namely commercial fisheries and shellfish (D3), hydrology (D7), chemical pollution of the environment (D8), contaminants in fish and seafood (D9) and marine litter (D10), in three study sites of the Mediterranean Sea (i.e. the Gulf of Lions, the Saronikos Gulf and the Cilician basin). *MeTali* specifically aims at bridging the existing gap in marine policy- and decision-making related to the selection and prioritization of measures and policies towards coastal and marine ecosystem protection. For this purpose, *MeTali* provides a cost-effectiveness ranking algorithm of selected command-and-control, economic, social and technological measures using estimates that are based on expert judgment.

The rest of the paper is structured as follows: **Section 2** first describes in brief the theoretical framework, i.e. the expert judgment approach and then the methodological development of *MeTali*. **Section 3** presents the pilot application of *MeTali* in Greece and provides, for illustrative purposes, a cost-effectiveness ranking of policy measures for three selected descriptors. Finally, **Section 4** concludes and discusses policy implications.

2. Development of the *MeTali* tool

2.1. Methodological background

Expert judgment is an approach for soliciting informed opinions from experts, i.e. those who have knowledge of an issue at an appropriate level of detail and who are capable of communicating this knowledge [4], or those whose opinion might be of interest [5]. Expert judgment can provide useful insights for policy- and decision-makers when scientific research is not available or is ongoing [6,7]. Moreover, it can be useful when current research needs to be made directly useful to policy- and decision-makers but comprehensive empirical information is lacking [8]. In this sense, experts can be relied on to consolidate and synthesize new or existing qualitative and/or quantitative information into a

framework suitable for decision-making [9–11].

Nevertheless, whether or not expert judgment has the potential to provide accurate, reliable and uncontested data is still debatable. Some researchers argue that experts are sensitive to a number of heuristics (e.g. representativeness, availability, anchoring and adjustment, overconfidence, etc.) and may be subject to cognitive and motivational biases that impair their abilities to accurately report their true beliefs [4,11]. On the other hand, other researchers claim that experts have superior recall of information and improved abilities to abstract knowledge to new situations and, thus, they are able to think critically about data and methods in their domain [12].

Expert judgment approaches have been used extensively in marine policy and science issues. This is not surprising considering that marine science is characterized by data unavailability, large uncertainties, and costly research and monitoring, which, if combined with conflicting interests and values about governance practice, complicate marine decision-making [13]. For instance, Halpern et al. [14] devised a method for collecting expert opinion on how threats affect marine ecosystems. They surveyed 135 experts from 19 different countries who were asked to assess the functional impact, scale, and frequency of a threat to an ecosystem; the resistance and recovery time of an ecosystem to a threat; and the certainty of these estimates. Teck et al. [8] applied an expert judgment approach to the California Current region in order to evaluate the relative vulnerability of 19 marine ecosystems to 53 stressors associated with human activities, i.e. a total of 1007 stressor-by-ecosystem combinations, using surveys from 107 experts. In order to gain more understanding into the role of scientific information in marine management and policy-making, van Haastrecht and Toonen [13] asked policy makers and scientists for their expert judgment in cases where crucial information for policy decisions was missing. The experts acted both as information producers and users describing the selection process of Marine Protected Areas (MPAs) in the Dutch part of the North Sea. Thus, at times they were asked to speak from their gut, or to voice opinions that consisted of a mix of both scientific and managerial considerations. Carollo et al. [15] contacted via email more than 2000 experts working with coastal and ocean issues in the Gulf of Mexico asking them to rank 12 data categories (e.g., benthic habitats, human uses, coral reef, oyster reef, temperature, etc.) and score the relevance of four qualifiers (spatial resolution, temporal resolution, age of data, and level of detail) using a discrete choice approach. In total 348 surveys were completed and the results were used to identify data gaps and consequently identify priority areas where money should be invested for future data collection. Furthermore, Carollo et al. [16] used expert opinion at the first Gulf of Mexico Ecosystem Services Workshop (2010) in order to: (a) identify and classify the Gulf of Mexico habitat types based on the Coastal and Marine Ecological Classification Standard (CMECS); (b) link ecosystem services to the Gulf of Mexico habitat types; and (c) prioritize ecosystem services. Ban et al. [17] involved 21 experts, using the Great Barrier Reef as case study, in order to construct a framework where management options can be evaluated; to obtain estimates of outcomes associated with a variety of scenarios; and to better understand the interaction of multiple stressors and related management options where data about the effects of these interactions were incomplete. Cook et al. [18] used integrated conceptual ecosystem models of the coastal marine environment developed as part of the Marine and Estuarine Goal Setting for South Florida (MARES) project in conjunction with a modified DPSIR model, expert opinion and matrix-based analyses to explore the direct and indirect relative impact of 12 ecosystem pressures on 11 ecosystem states and 11 ecosystem services (i.e. 34 components) identified through MARES project. More specifically, among these 34 components 193 unique interactions were identified and 25 experts

were interviewed to quantify the relative interaction strength of these 193 linkages. Halpern et al. [19] used a regional application of the Ocean Health Index in the California Current in order to describe a systematic approach to eliciting opinions of experts from multiple sectors. Their approach involves two different methods to elicit preferences based on the tradeoffs that would likely emerge from management decisions within the California Current based on random utility theory and on analytical deliberation. Vacchi et al. [20] adopted a Multi-Criteria Decision Aid (MCDA) procedure in order to develop a spatial explicit tool that would be used to identify the best management strategy for the Bergoggi Island MPA. The intensity of each pressure (e.g. seafloor abrasion, organic enrichment and pollutants, traffic related pressures) created by local activities at the status of coastal ecosystem was estimated using expert elicitation under a fuzzy environment.

2.2. Overview of the MeTaLi tool

The MeTaLi tool provides a simple, transparent and straightforward method to select and rank alternative policy measures (i.e. command-and-control, economic, social and technological) to achieve or maintain GES in EU marine waters. It uses a cost-effectiveness ranking algorithm, which is based on an expert judgment framework. The MeTaLi tool has been originally developed to address five MSFD descriptors which are of interest to the MERMAID project, namely commercial fisheries and shellfish (D3), hydrology (D7), chemical pollution of the environment (D8), contaminants in fish and seafood (D9) and marine litter (D10). Nevertheless, MeTaLi is an open and expandable tool platform and, thus, it can be applied across all European countries in the context of EU MSFD for all the eleven descriptors. It could also be applied across other regional sea areas internationally to analyze and compare protection measures towards managing the marine environment.

The MeTaLi tool is organized in a spreadsheet-based format and includes three different sections and thirteen worksheets, in total. The first section collects biographical information of the expert who is interviewed. The second section provides information to the expert about the aim of the tool and the methodological approach used. The third section includes eleven worksheets and collects the information required to establish the relationship between a number of policy measures and the targets set by MSFD for the selected descriptors. For each descriptor fifteen measures are evaluated. Thirteen of the measures are pre-defined by MeTaLi (Table 1) and two measures can be determined (if necessary) by the expert.

The information requested by the expert in the third section for the descriptor(s) that she/he feels capable of answering includes:

Table 1
Pre-defined measures of MeTaLi for the achievement of GES in the selected descriptors.

Acronym	Measure
M1	Imposition of bans and activities controls
M2	Imposition of fines
M3	Taxes
M4	Subsidies
M5	Trading schemes
M6	Compensation schemes
M7	Output controls
M8	Spatial and temporal distribution controls
M9	Measures to improve the traceability
M10	Management coordination measures
M11	Mitigation and remediation tools
M12	Communication and raising public awareness
M13	Stakeholder involvement

- The gap between current status and GES for the selected descriptor(s).
- The effectiveness of selected measures, i.e. the ability of a measure to achieve the objectives set by decision-makers. The expert selects five out of the fifteen alternative measures, which are considered to be more suitable for the descriptor under investigation.
- The implementation cost of each of the selected five measures.
- The difficulty of implementing each of the selected five measures due to e.g. expected reactions from stakeholders, shortcomings related to administrative issues, weak enforcement of the measure, etc.
- The degree of certainty of the expert, i.e. a subjective assessment of certainty about her/his estimates in the effectiveness, the implementation cost, and the applicability of each of the selected measures.

The gap between the current status and the GES is defined in linguistic form using a five-point Likert scale (Very low, Low, Medium, High and Very high). The same five-point linguistic scale is used to detect the effectiveness, the implementation cost and the implementation difficulty (applicability) of the measures evaluated, as well as the degree of certainty of the expert interviewed.

In order to tackle features of uncertainty, the estimates are provided in the form of fuzzy numbers based on fuzzy set theory [21,22]. This step accordingly implements a fuzzy scoring approach where the five-point scale demonstrates the conversion of triangular fuzzy numbers into crisp scores. A triangular fuzzy number $T=(a, b, c)$ is defined as a convex and normalized fuzzy set determined by the minimum, the central (i.e. most plausible) and the maximum value with membership function $\mu_A(x)$:

$$T = \begin{cases} x - a/b - a & a \leq x \leq b \\ x - c/b - c & b \leq x \leq c \\ 0 & \text{otherwise} \end{cases}$$

where $[a,c]$ is the supporting interval and the point $(b,1)$ is the peak.

The vagueness of each linguistic term used in the five-point Likert scale is listed in Table 2 and is presented in Fig. 1.

Finally, the expert defines the certainty of her/his estimates using a similar five-point Likert scale. In this case, however, the degree of certainty of the expert acts as an overall weighting factor of the estimates. The conversion of the certainty linguistic terms in numerical values is given in Table 3.

2.3. Cost-effectiveness ranking of measures

The collected data are processed in order to calculate the cost-effectiveness indicator of measures for each descriptor separately according to the approach described hereinafter. The ranking of the selected measures is performed through the calculation of an overall cost-effectiveness indicator, which reflects the effectiveness, the implementation cost and the applicability of the selected measures for the specified descriptor. In addition, this indicator takes into account two additional indices, namely: (a) the frequency of selection of each measure and (b) the weighting of the

Table 2
Representation of linguistic terms by triangular fuzzy numbers.

Linguistic term	Very low	Low	Medium	High	Very high
Fuzzy number	(0, 0, 0.3)	(0.1, 0.3, 0.5)	(0.3, 0.5, 0.7)	(0.5, 0.7, 0.9)	(0.7, 1, 1)

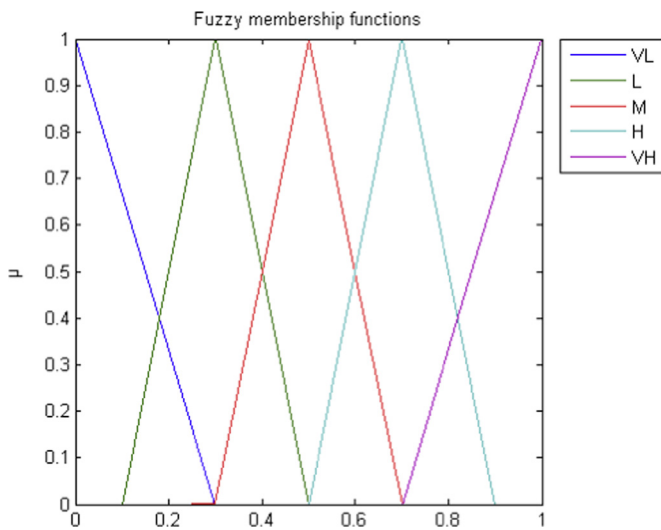


Fig. 1. The set of the five linguistic terms with their semantics.

Table 3
Representation of the certainty linguistic variables by their numerical values.

Linguistic term	Very low	Low	Medium	High	Very high
Numerical values	1.1	1.2	1.3	1.4	1.5

degree of certainty expressed by the expert. The first index is estimated as follow:

$$FS_i = \frac{\sum_{j=1}^n M_i}{n}$$

where FS_i is the frequency of selection of measure i , M_i is the total number of times the measure, i is selected by the experts j and n is the total number of experts that responded for the descriptor under examination.

The weighting of the degree of certainty uses the subjective opinion of the experts involved in the survey. As mentioned, expert E_i expresses her/his certainty as to the accuracy of her/his estimates in linguistic terms for each of the three parameters (i.e. effectiveness, implementation cost and applicability). The overall ‘certainty’ weight of expert E_i is estimated as the geometric mean of her/his degrees of certainty:

$$DC_j = \left(\prod_{k=1}^3 dc_{jk} \right)^{1/3}$$

where DC_j is the degree of certainty of expert j and dc_{jk} is the degree of certainty of expert j for parameter k .

Hence, the cost-effectiveness indicator of each measure i for a specific descriptor is estimated as follow:

$$CEff_{ij} = \frac{Eff_{ij}}{Cost_{ij}} \times (1 - App_{ij}) \times FS_i \times DC_j$$

where $CSEff_{ij}$ is the cost-effectiveness indicator of measure i as estimated by expert j , Eff_{ij} is the effectiveness of measure i estimated by expert j , $Cost_{ij}$ is the implementation cost of measure i estimated by expert j , App_{ij} is the applicability (i.e. implementation difficulty) of measure i estimated by expert j , FS_i is the frequency of selection of measure i by all experts and DC_j is the degree of certainty of expert j .

In the above-mentioned equation, the estimates of the parameters Eff , $Cost$ and App are provided in the form of triangular fuzzy numbers. In general, there are three main approaches when faced with differing expert opinions: (i) propagate each expert’s

distribution separately; (ii) require the experts to create a single consensus distribution; (iii) combine the expert opinions in some way. There is no consensus on how to deal with this issue in the context of integrated assessment models [23]. This method adopts the third approach in order to calculate experts’ values with respect to the effectiveness, cost of implementation and applicability of selected measures towards achieving the MSFD targets for the selected descriptors.

Given that typical t-norm (intersection) and t-conorm (union) operators result in a very restricted representation of the wide range of experts’ beliefs and may inadvertently propagate uncertainty and limit the ability of policymakers to make strategic hedges against risky outlier events [24], the Maximum Entropy approach was used. This approach is proposed for the wider range of ‘judged’ uncertainty elicited by the experts and it is a common measure of information in modern communications theory [25]. The idea behind Maximum Entropy is to formulate a distribution for the data such that the distribution maximizes the uncertainty in the data, subject to known constraints [4]. As Gay and Estrada [26] note, Maximum Entropy Principle is ‘...a useful tool for constructing probabilistic climate change scenarios that are the least biased estimates possible, consistent with the information at hand (including expert or decision-maker judgment) and that maximize what is not known’.

This definition of entropy, introduced by Shannon [27], resembles a formula for a thermodynamic notion of entropy. For a continuous probability density function $p(x)$ on an interval I , its entropy is defined as

$$h(p) = - \int_I p(x) \ln p(x) dx$$

Using Shannon’s entropy measure, Jaynes [28] showed that the maximum entropy estimate is the least biased estimate possible with the information at hand since it maximizes the uncertainty subject to the partial information available. This means that the choice of any other distribution will require making additional assumptions unsupported by the given constraints [29]. A direct derivation of the maximum entropy distribution involves solving a system of nonlinear equations, the solution of which involves variational calculus using the Lagrange multiplier method. The maximum entropy distribution can help assign probability distributions given certain constraints. For instance, when only the lower and upper bounds for an uncertain parameter are known, the principle of maximum entropy would indicate a uniform distribution. When the minimum, maximum and mode values are given, the beta distribution that maximizes the entropy is chosen ([30], quoted in [31]).

In order to better represent the divergence of opinions and the uncertainty involved in the estimates, the minimum and maximum values provided by the experts are combined with equal weight. Thus, expert judgments are aggregated to construct a triangular distribution, using the minimum, maximum and mode values of all experts:

$$T(a, b, c) = [\min a_L^i, \text{mean } a_M^i, \max a_U^i]$$

where $\min a_L^i$ is the minimum of the minimum values elicited by the experts, $\text{mean } a_M^i$ is the average of the mode values elicited by the experts and $\max a_U^i$ is the maximum of the maximum values elicited by the experts

The fuzzy numbers are then defuzzified using fuzzy sets and integration theory [e.g. [32,33] by means of the centroid ($C_{\tilde{B}}$) of the triangular $\tilde{B} = (a, b, c)$], as follows:

$$C_{\tilde{B}} = \frac{a + b + c}{3}$$

At the final step, the crisp values of the cost-effectiveness indicators obtained after defuzzification from the fuzzy average estimates of the experts are normalized in a scale from 1 to 100, forming the final cost-effectiveness ranking of alternative measures for the descriptor under examination.

Furthermore, an uncertainty analysis is performed so as to assess the robustness of the obtained results. To this direction the Monte Carlo stochastic process was selected, implementing a random sampling of values according to appropriately established probability distribution for each uncertain input parameter for numerous scenarios (iterations). Monte Carlo simulation evaluates iteratively the specified output using different sets of random numbers as inputs and, thus, it can be viewed as inclusive of all deterministic estimates of the experts with a finite probability of occurrence. In this context, triangular distributions defined according to Maximum Entropy approach were utilized in order to depict the uncertainties in the effectiveness, the implementation cost and the applicability of the examined measures for all the descriptors.

3. Testing the *MeTaLi* tool

3.1. Objectives and administration of the pilot study

A pilot application of the tool was conducted in Greece, in order to: (a) test the methodological approach; (b) evaluate its capabilities; (c) identify potential drawbacks; (d) make any necessary changes; and (e) draw conclusions and prepare guidelines for its implementation in real cases. To this end, three descriptors, i.e. chemical pollution of the environment (D8), contaminants in fish and seafood (D9) and marine litter (D10), were initially selected and examined. A strategic sample of experts was selected involving partners of MERMAID project, and various other experts with relevant scientific background and experience within the three examined descriptors. The expert panel's composition is a crucial parameter for the successful implementation of any expert judgment survey. Thus, a well-composed and balanced sample was created consisting of 28 policy makers, physical scientists and practitioners, who were divided into three groups according to their expertise (i.e. one group per each descriptor), as shown in Table 4.

Following Morgan et al. [7], a structured elicitation of each expert's judgment was selected provided that neither consensus between experts nor a mechanism for iterative communication between experts was required. This approach also ensured that the expert judgments provided were free of interactions, since the reactions of other experts present in interactive groups can provoke the so-called 'social pressure' bias [4]. The survey was carried out through the conduction of face-to-face interviews in order to facilitate the expert judgment process in a most appropriate and efficient way.

3.2. Selection and ranking of measures

Firstly, the experts of each of the three groups (Table 4) were asked to evaluate the gap between the current situation and the

fulfillment of a GES target set by MSFD for the examined descriptors. According to the results illustrated in Fig. 2, the gap between the current situation and the fulfillment of the GES target seems to be lower for descriptor D9 (43% of the participants assessed the gap as low). For the case of descriptor D8, 55% of the experts evaluated the gap either as low or medium; 80% of the respondents specified that the gap for descriptor D10 is either high or very high, highlighting the need for more concerted efforts to achieve the MSFD targets in this case.

Then, the experts of each group were asked to select the five most effective measures (Table 1) to fulfill the MSFD targets (Fig. 3) for the corresponding descriptors. Taxes (M3) and trading schemes (M5) were not selected for any of the examined descriptors, and compensation schemes (M6) were selected only once (for D10). Furthermore, two additional measures were identified for D8, namely licensing procedures with standards and common framework (M14) and education of the personnel (M15).

Regarding the most popular measures, the imposition of bans and activities controls (M1) was selected from at least 86% of the experts for all the examined descriptors, and output controls (M7) and communication and raising public awareness (M12) were selected from at least 64% of the experts, accordingly.

The implementation of the described methodology resulted in the final cost-effectiveness ranking (Table 5) and the normalized cost-effectiveness indicators (Fig. 4) of the examined measures.

According to the final cost-effectiveness ranking list, the imposition of bans and activities controls (M1) is the most cost-effective measure for descriptor D8, while the measures of communication and raising public awareness (M12) and output controls (M7) are the second and third most cost-effective. For descriptor D9, the main cost-effective measures consist of measures to improve the traceability (M9), output controls (M7) and the imposition of bans and activities controls (M1). Finally, the imposition of bans and activities controls (M1) seems to be the most cost-effective for the descriptor D10 followed by the measures of communication and raising public awareness (M12) and stakeholder involvement (M13).

The evaluation of the previous results should be performed taking into consideration the relative differences among the normalized cost-effectiveness indicators due to the fact that in some descriptors the most cost-effective measures are more competitive than in other descriptors.

According to the results of Fig. 4, the imposition of bans and activities controls (M1) has as indicator the value of 100 for descriptor D10, while the communication and raising public awareness actions (M12) has the second highest indicator equal to 53.

As regards the other two descriptors, there are no significant differences among the examined measures. Indicatively, the imposition of bans and activities controls (M1) has as indicator of 100 for descriptor D8, while the indicator of the measure of communication and raising public awareness (M12), which is the second most cost-effective, amounts to 76. Similarly, the indicators of the measures to improve the traceability (M9) and output controls equal 100 and 80 respectively for descriptor D9.

3.3. Stochastic analysis

The results of the Monte Carlo simulation process are presented in Tables 6–8 for each descriptor separately. The results consist of various statistical measures such as mean, median, min, max and standard deviation, and the percentiles of the estimated normalized cost-effectiveness indicators of measures in order to achieve the GES.

The results of the stochastic analysis are almost identical with the results of the basic (i.e. deterministic) analysis confirming the robustness of the obtained estimates. The same conclusion can be

Table 4
Number of participants in the pilot application per descriptor examined.

Descriptor	Number of participants
Contaminants within water	11
Contaminants within seafood	7
Marine litter	10

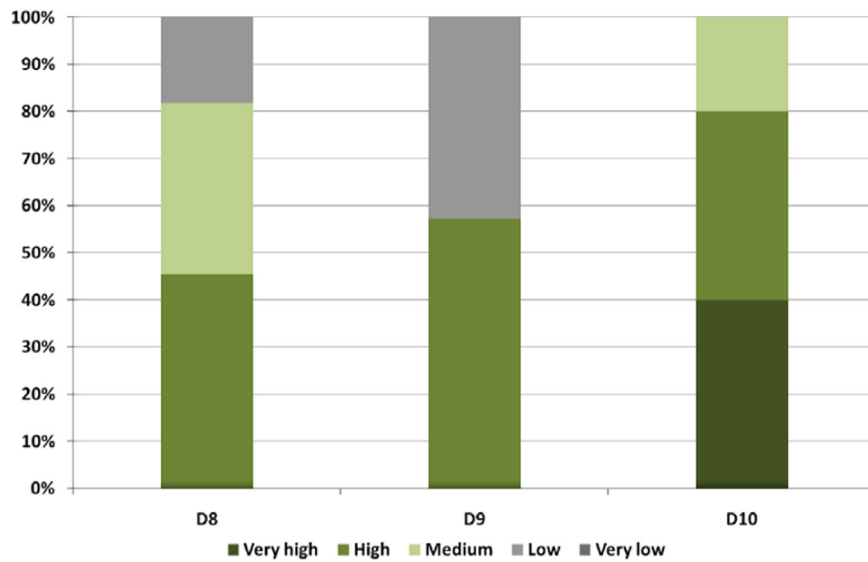


Fig. 2. Estimated gap between the current situation and the GES target set by MSFD 2008/56/EC for the examined descriptors.

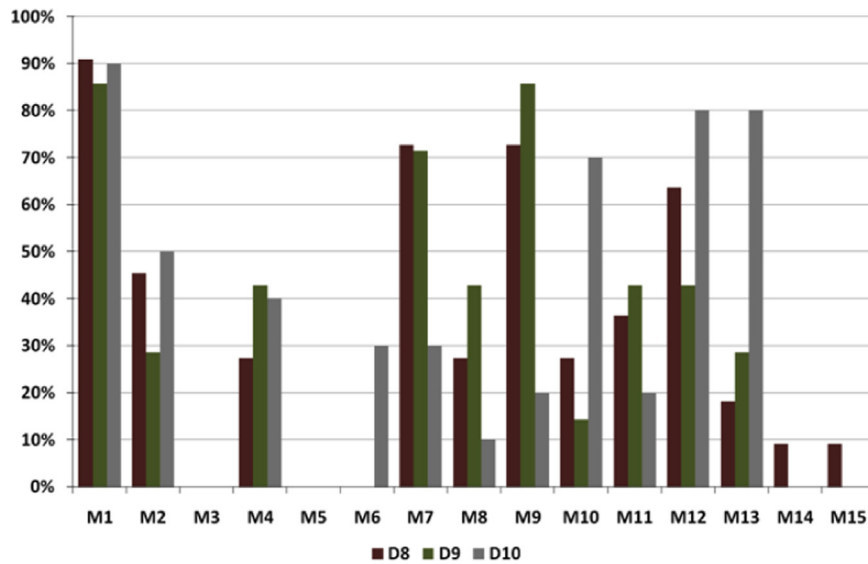


Fig. 3. Popularity of selected measures for the examined descriptors towards achieving GES.

Table 5
Final cost-effectiveness ranking of measures for the examined descriptors towards achieving GES.

	D8	D9	D10
M1	1	3	1
M2	5	7	4
M3	13	11	12
M4	9	9	7
M5	13	11	12
M6	13	11	8
M7	3	2	6
M8	7	4	11
M9	4	1	8
M10	11	10	5
M11	8	5	10
M12	2	6	2
M13	6	8	3
M14	11	11	12
M15	9	11	12

derived from Table 9 were the final cost-effectiveness ranking of measures for achieving GES estimated by the mean estimates of Monte Carlo analysis is presented. Specifically, there are no

differences among the most cost-effective measures for the examined descriptors, while only a few minor differentiations can be observed in the final ranking of the rest of the measures.

3.4. Discussion of results

According to the pilot-test findings, the gap between the current situation and the fulfillment of the GES target seems to be lower for descriptors D9 and D8 in comparison with D10. The imposition of bans and activities controls (M1) and the output controls (M7), which belong to traditional command and control instruments and the communication and raising public awareness (M12), which are categorized as social instruments, were selected from the vast majority of the experts for the examined descriptors. On the contrary, economic instruments, such as taxes (M3), trading schemes (M5) and compensation schemes (M6) were the less favored. Although the number of experts participating in the three groups of the pilot-test cannot lead to absolute conclusions, the findings coincide, at least in part, with those of other works in the field of marine policy. For instance, in a recent guidance [2] that collected evidence to support EU Member States to compile a set of measures for the implementation of the

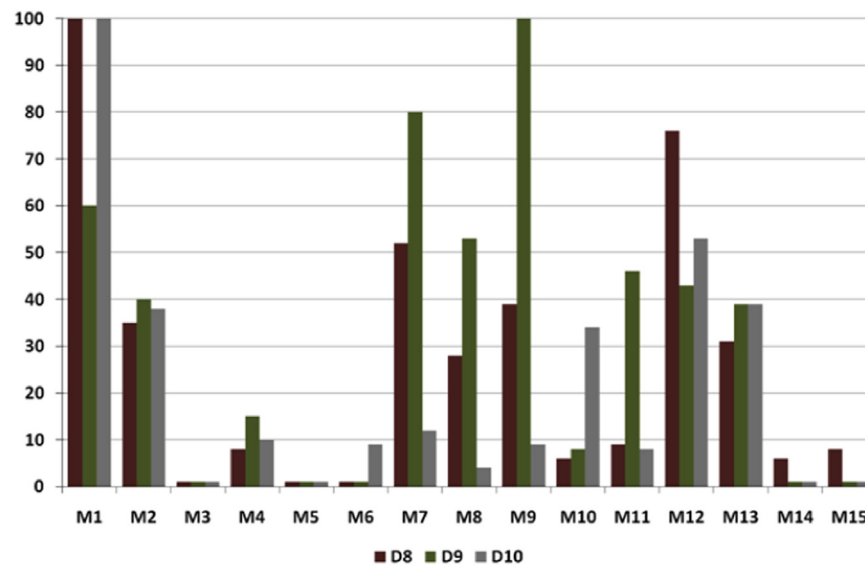


Fig. 4. Normalized cost-effectiveness indicators of selected measures for the examined descriptors towards achieving GES.

Table 6

Results of Monte Carlo analysis for descriptor D8.

D8		M1	M2	M4	M7	M8	M9	M10	M11	M12	M13	M14	M15	
Statistics	Mean	68	47	11	64	36	54	10	14	95	48	9	12	
	Median	67	45	10	61	34	51	10	13	100	44	8	11	
	St. dev.	21	16	4	20	13	17	4	5	11	19	4	4	
	Min	18	14	3	21	11	18	4	4	4	41	17	3	5
	Max	100	100	30	100	92	100	31	33	33	100	100	32	28
Percentiles	0%	18	14	3	21	11	18	4	4	41	17	3	5	
	10%	40	29	6	39	21	34	6	8	78	28	5	7	
	20%	48	33	7	45	24	39	7	9	92	33	6	9	
	30%	55	37	8	51	28	43	8	11	100	37	7	10	
	40%	61	41	9	57	31	47	9	12	100	40	8	10	
	50%	67	45	10	61	34	51	10	13	100	44	8	11	
	60%	74	49	11	66	36	56	11	14	100	48	9	12	
	70%	81	54	12	73	40	61	12	16	100	53	11	13	
	80%	90	59	14	82	45	68	13	18	100	61	12	14	
	90%	100	68	16	95	54	78	15	20	100	75	15	17	
	100%	100	100	30	100	92	100	31	33	100	100	32	28	

MSFD a database consisting of nearly 140 measures was developed. The inventory was based on a two-track approach. First, relevant sources, organizations and experts were detected through literature review and, then, 24 experts were contacted to provide their experiences through telephone interviews. According to the findings of the study, the majority of measures are allocated to command-and-control instruments (36%) and a broad category of technical, technological and research-oriented measures (32%), followed by economic instruments (23%) and social measures (9%). Regarding specifically descriptor D10, Reinhard et al. [34] assessed various measures in order to fulfill the requirements of MSFD using cost-effectiveness analysis. They concluded that the most cost-effective measures seem to be awareness campaigns, harmonization of the fees of port reception facilities and the control of garbage. A more recent report analyzed the contribution of 72 LIFE projects between 2005 and 2012 to all stages of MSFD implementation schedule [35]. As regards the type of measures implemented, it was found out that the majority of the projects make up contribution to command-and-control mechanisms (input, output and spatial and temporal distribution controls) and social instruments (management coordination measures, communication, stakeholder involvement and public awareness campaigns). Economic incentives and technical measures (e.g. measures to improve the traceability of

marine pollution, 'clean up' operations, etc.) were by far fewer. Hence, it seems that despite the increasing effort in recent years towards implementing market-based instruments (e.g. the "Fishing for Litter" initiatives) in support of the "polluter pays" principle, the "user/beneficiary pays" principle and the "full-cost recovery" principle [36], the use of command-and-control measures still dominates marine policy-making and thinking.

According to the final cost-effectiveness ranking list, the imposition of bans and activities controls (M1) is declared as the most cost-effective policy measure for descriptors D8 and D10 and the third most cost-effective for descriptor D9. The measure of output controls (M7) was the second and third most cost-effective for descriptors D9 and D8 correspondingly, and the measure of communication and raising public awareness (M12) was the second most cost-effective for both of the descriptors D8 and D10.

Generally, the establishment of actions in order to quantify the potential cumulative and in-combination environmental effects triggered by the implemented measures can be considered as priority [37,38]. Moreover, Berg et al. [39] have suggested the re-arrangement and elimination of redundant criteria and attributes in order to avoid 'double counting' in specific descriptors. To the same direction, the STAGES project targeted to analyze and model

Table 7
Results of Monte Carlo analysis for descriptor D9.

D9		M1	M2	M4	M7	M8	M9	M10	M11	M12	M13
Statistics	Mean	54	11	13	70	41	76	8	45	43	32
	Median	49	7	12	72	36	86	7	39	40	28
	St. dev.	29	12	8	27	25	26	3	26	19	17
	Min	2	1	1	7	3	7	2	4	6	3
	Max	100	100	62	100	100	100	26	100	100	100
Percentiles	0%	2	1	1	7	3	7	2	4	6	3
	10%	18	3	5	31	15	36	4	16	20	14
	20%	26	4	7	41	20	48	5	22	26	18
	30%	33	5	8	51	25	60	5	28	31	21
	40%	41	6	10	61	30	72	6	34	35	25
	50%	49	7	12	72	36	86	7	39	40	28
	60%	58	9	13	83	41	100	8	46	45	32
	70%	70	11	15	98	48	100	9	54	50	39
	80%	87	15	18	100	60	100	10	65	57	46
	90%	100	24	23	100	81	100	12	89	68	55
	100%	100	100	62	100	100	100	26	100	100	100

the synergistic and cumulative effects of multiple pressures for the case of fisheries and eutrophication [40]. Hence, in cases that implementation of specific measures per descriptor is not feasible, decision makers might seek to find solutions that would generally attain the goals of MSFD for as low a cost as possible. Nevertheless, no specific and homogeneous methodology exists, so far, for the effective assessment of these effects in the examined descriptors within the framework of MSFD. To examine this issue, the overall cost-effectiveness of the examined measures was estimated taking into consideration the ranking of the measures for each examined descriptor (Table 10). The obtained results confirmed the above-mentioned remarks; it seems that even though no specific note was given to the experts, the overall effectiveness of each measure was taken into consideration, at least partially, during the provision of their estimates.

Monte Carlo simulation allows the analyst to determine estimates that would be difficult to determine using deterministic analysis and, thus, can help increase the understanding of data and tackle the uncertainty involved in the estimates [4,41]. In this case, the estimated overall rank derived by the Monte Carlo analysis confirmed the robustness of the results since only minor differences exist among the ranking of the measures between the basic and the stochastic process.

Table 8
Results of Monte Carlo analysis for descriptor D10.

D10		M1	M2	M4	M6	M7	M8	M9	M10	M11	M12	M13
Statistics	Mean	73	42	15	15	18	5	12	51	13	59	57
	Median	84	36	13	13	16	5	10	45	11	56	52
	St. dev.	29	25	9	9	9	2	7	30	7	30	31
	Min	2	2	1	2	2	1	1	3	2	2	4
	Max	100	100	66	64	68	15	52	100	48	100	100
Percentiles	0%	2	2	1	2	2	1	1	3	2	2	4
	10%	27	13	5	6	8	3	5	15	5	20	18
	20%	41	19	7	8	11	4	6	24	7	29	27
	30%	54	25	9	10	13	4	8	30	9	38	34
	40%	66	30	11	11	15	5	9	37	10	47	42
	50%	83	36	13	13	16	5	10	45	11	56	52
	60%	99	42	15	15	19	6	12	56	13	68	64
	70%	100	51	17	18	21	6	14	67	15	81	80
	80%	100	62	21	21	24	7	16	84	18	100	100
	90%	100	81	26	28	29	8	20	100	23	100	100
	100%	100	100	66	64	68	15	52	100	48	100	100

Table 9

Final cost-effectiveness ranking of measures for the descriptors examined based on Monte Carlo analysis.

	D8	D9	D10
M1	2	3	1
M2	4	9	5
M3	13	11	12
M4	9	8	8
M5	13	11	12
M6	13	11	7
M7	3	2	6
M8	7	6	11
M9	5	1	10
M10	11	10	4
M11	8	4	9
M12	1	5	2
M13	6	7	3
M14	12	11	12
M15	10	11	12

4. Policy implications and conclusions

The EU MSFD sets quite demanding requirements with reference to the evaluation of the proposed program of measures using cost-benefit or cost-effectiveness analyses. So far, however,

Table 10
Total ranking of the examined measures for the deterministic and Monte Carlo analyses.

Measure	Overall rank – basic analysis	Overall rank – stochastic analysis
M1	1	1
M12	2	2
M7	3	3
M9	4	4
M2	5	6
M13	6	4
M8	7	8
M11	8	7
M4	9	9
M10	10	9
M6	11	11
M15	11	12
M14	13	13
M3	14	14
M5	14	14

there are only limited efforts to develop suitable tools towards facilitating this task within the prescribed timeline. Aiming at bridging this gap in marine policy- and decision-making, a cost-effectiveness ranking algorithm, named *MeTaLi*, has been developed. It provides a simple, transparent and straightforward framework to evaluate alternative policy measures for achieving or maintaining GES in EU marine waters under the MSFD. Originally, the tool is oriented to five specific MSFD descriptors, namely commercial fisheries and shellfish (D3), hydrology (D7), chemical pollution of the environment (D8), contaminants in fish and sea-food (D9) and marine litter (D10). Yet, its open architecture makes it extremely adaptable and ‘transferable’; it can be applied to a wider spectrum of marine policy applications with minimum modifications across EU Member States for all the eleven descriptors of the MSFD, as well as across other areas around the world for managing marine policy problems, in general.

MeTaLi is built upon expert judgment since the available scientific knowledge and information (e.g. detailed cost data of measures) in the field of marine policy is far from being adequate, at least under the logic of the EU MSFD. This also means that the methodology is sensitive to a number of heuristics and biases that affect expert judgment approaches cf. [4,42]. As a means to tackle with these issues, at least partially, the methodology involves fuzzy numbers relating to the language-based representations of probability of the estimates and stochastic analysis to reveal the parameters that are most sensitive to experts’ opinions via Monte Carlo simulations. In general, the methodological approach follows the guidelines and the conclusion drawn by a recent study [42] as regards the harmonization of economic and social analysis when preparing programs of measures. This study highlighted the importance of the integration of semi-quantitative or qualitative assessments, the elicitation of expert opinions in combination with bibliographical review, the involvement of stakeholders for the collection of information and judgments and the conduction of uncertainty analysis.

The findings from the pilot-test of the tool are promising. For example, besides establishing a priority ranking of measures just for one selected descriptor, a synthesis of the results for different descriptors may be used to help marine policy- and decision-makers determine the most ‘overall’ cost-effective measures. This ‘overall’ ranking takes could take into consideration potential synergies of policy measures between different descriptors and could prove valuable in cases that time and/or budget constraints forbid the implementation of descriptor-specific measures.

At the same time, however, the pilot application reveals the need for further improvements and considerations. For instance, it is evident that there exist ambiguities in the identification of the

effectiveness, the implementation cost and the implementation difficulties of the measures. These are associated, among others, with information gaps, as well as the different backgrounds of the experts. Thus, in real-scale applications of the methodology, e.g. towards establishing “best-practices” for programs of measures at local, national, regional and European levels, a diverse and sufficient number experts should be involved considering the multi-disciplinary nature of the problem. Moreover, future research is needed for measuring the impacts of the proposed measures in physical and economic terms. Some measures may be considered to be cost-effective; yet, they may fail to achieve environmental goals or may not be feasible from a social point of view (e.g. in cases where the cost of the measure exceeds the monetized social benefits).

Concluding it should be emphasized that the results of *MeTaLi* alone do not provide a final answer regarding the least-cost achievement of GES in EU marine waters. There are many dimensions (e.g. scientific, socioeconomic, etc.) that drive such decisions and should be thoroughly investigated. However, having a quantitative and transparent cost-effectiveness method could provide a key piece of the answer to the fundamental question of how to prioritize measures in the context of the MSFD when scientific evidence is lacking and – from this point of view – *MeTaLi* could be of enormous value to Member States that need to justify and defend their marine policy decisions on economic grounds.

Acknowledgments

The research leading to these results has received funding from the European Commission FP7 (ERANET/SEAS-ERA) and the General Secretary of Research and Technology (GSRT), Greece, in the framework of the project “Marine Environmental targets linked to Regional Management schemes based on Indicators Developed for the Mediterranean – MERMAID” (<http://www.mermaid-era.eu/home/>). The authors would like to thank Dr. E. Kamberi for her contribution to many productive discussions related to the content of this paper. The authors are also grateful to all scientists who participated in the elicitation process and kindly offered their valuable input.

References

- [1] European Commission, Scoping study on the requirements for economic assessment in the Marine Strategy Framework Directive, COWI, Report for DG Environment – European Commission, 2010.
- [2] Arcadis Belgium, Economic assessment of policy measures for the implementation of the Marine Strategy Framework Directive, European Commission – DG Environment, Project number 11601, Final Report, 2012.
- [3] Working Group on Economic and Social Assessment (WG-ESA), Economic and social analysis for the Initial Assessment for the Marine Strategy Framework Directive: a guidance document, 21 December 2010, Directorate-General Environment, European Commission, 2010.
- [4] M.A. Meyer, J.M. Booker, *Eliciting and Analyzing Expert Judgment – A Practical Guide*, Statistical Science Group, Los Alamos National Laboratory, Los Alamos, New Mexico, 2001.
- [5] T. Krueger, et al., The role of expert opinion in environmental modelling, *Environ. Model. Softw.* 36 (2012) 4–18.
- [6] M.G. Morgan, L.F. Pitelka, E. Shevliakova, Elicitation of expert judgments of climate change impacts on forest ecosystems, *Clim. Change* 49 (2001) 279–307.
- [7] M.G. Morgan, P.J. Adams, D.W. Keith, Elicitation of expert judgments of aerosol forcing, *Clim. Change* 75 (2006) 195–214.
- [8] S.J. Teck, B.S. Halpern, C.V. Kappel, F. Micheli, K.A. Selkoe, C.M. Crain, R. Martone, C. Shearer, J. Arvai, B. Fischhoff, et al., Using expert judgment to estimate marine ecosystem vulnerability in the California Current, *Ecol. Appl.* 20 (2010) 1402–1416.
- [9] I. Fazey, J.A. Fazey, J.G. Salisbury, D.B. Lindenmayer, S. Dovers, The nature and role of experiential knowledge for environmental conservation, *Environ. Conserv.* 33 (1) (2006) 1–10.
- [10] P.M. Kuhnert, T.G. Martin, S.P. Griffiths, A guide to eliciting and using expert

- knowledge in Bayesian ecological models, *Ecol. Lett.* 7 (2010) 900–914.
- [11] M.F. McBride, F. Fidler, M.A. Burgman, Evaluating the accuracy and calibration of expert predictions under uncertainty: predicting the outcomes of ecological research, *Divers. Distrib.* 18 (2012) 782–794.
- [12] M. Burgman, F. Fidler, M. McBride, T. Walshe, B. Wintle, Eliciting expert judgments: literature review, First project report, Australian Centre of Excellence for Risk Analysis, ACERA Project 0611, University of Melbourne, 2006.
- [13] E.K. Van Haastrecht, H.M. Toonen, Science-policy interactions in MPA site selection in the Dutch part of the North Sea, *Environ. Manag.* 47 (4) (2011) 656–670.
- [14] B.S. Halpern, K.A. Selkoe, F. Micheli, C.V. Kappel, Evaluating and ranking the vulnerability of global marine ecosystems to anthropogenic threats, *Conserv. Biol.* 21 (5) (2007) 1301–1315.
- [15] C. Carollo, A. McKenzie, E.H. Leone, Quantitative gap analysis for the West Coast of Florida based on experts opinion, *J. Coast. Conserv.* 16 (2012) 281–316.
- [16] C. Carollo, R.J. Allee, D.W. Yoskowitz, Linking the Coastal and Marine Ecological Classification Standard (CMECS) to ecosystem services: an application to the US Gulf of Mexico, *Int. J. Biodivers. Sci. Ecosyst. Serv. Manag.* 9 (2013) 249–256.
- [17] S.S. Ban, R.L. Pressey, N.A.J. Graham, Assessing interactions of multiple stressors when data are limited: a Bayesian belief network applied to coral reefs, *Global Environ. Change* 27 (2014) 64–72.
- [18] G.S. Cook, P.J. Fletcher, C.R. Kelble, Towards marine ecosystem based management in South Florida: investigating the connections among ecosystem pressures, states, and services in a complex coastal system, *Ecol. Indic.* 44 (2014) 26–39.
- [19] B.S. Halpern, C. Longo, K.L. McLeod, R. Cooke, B. Fischhoff, J.F. Samhour, C. Scarborough, Elicited preferences for components of ocean health in the California Current, *Mar. Policy* 42 (2013) 68–73.
- [20] M. Vacchi, M. Montefalcone, V. Parravicini, A. Rovere, P. Vassallo, M. Ferrari, C. Morri, C.N. Bianchi, Spatial models to support the management of coastal marine ecosystems: a short review of best practices in Liguria, Italy, *Mediterr. Mar. Sci.* 15 (1) (2014) 189–197.
- [21] L. Zadeh, Fuzzy sets, *Inform. Control* 8 (1965) 338–353.
- [22] L. Zadeh, *Fuzzy Sets and Applications: Selected Papers*, Wiley, New York, 1987.
- [23] D. Webster, P. Sokolov, A methodology for quantifying uncertainty in climate projections, *Clim. Change* 46 (4) (2000) 417–446.
- [24] S.H. Schneider, K. Kuntz-Duriseti, Uncertainty and climate change policy, in: S. H. Schneider, A. Rosencranz, J.O. Niles (Eds.), *Climate Change Policy: A Survey*, Island Press, Washington, 2002, pp. 53–88.
- [25] G. Baecher, J. Christian, *Reliability and Statistics in Geotechnical Engineering*, Wiley, England (2003), p. 618.
- [26] C. Gay, F. Estrada, Objective probabilities about future climate are a matter of opinion, *Clim. Change* 99 (2010) 27–46.
- [27] C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* 27 (379–423) (1948) 623–656.
- [28] E.T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.* 106 (4) (1957) 620–630.
- [29] A. Duracz, *Derivation of Probability Distributions for Risk Assessment*, Master's Thesis in Computer Science, Stockholm University, Sweden, 2006.
- [30] M.E. Harr, *Reliability-Based Design in Civil Engineering*, McGraw-Hill, New York, 1987.
- [31] S. Mishra, Assigning probability distributions to input parameters of performance assessment models, SKB Technical Report TR-02-11, 2002.
- [32] Y. Liu, B. Liu, Expected value of fuzzy variable and fuzzy expected value models, *IEEE Trans.* 10 (4) (2002) 445–450.
- [33] G. Bojadziev, M. Bojadziev, *Fuzzy Logic for Business, Finance, and Management*, 2nd ed, World scientific Publishing, Singapore (2007), p. 232.
- [34] S. Reinhard, A. de Blaeij, M.-J. Bogaardt, A. Gaaff, M. Leopold, M. Scholl, D. Slijkerman, W.-J. Strietman, P. van der Wielen, Cost-effectiveness and cost-benefit analysis for the MSFD, Framework for the Netherlands, LEI report 2011-036, 2012.
- [35] L. Barratt, J. Houston, C. Rose, D. Mitchell, The future of Europe's seas – contribution of the LIFE programme to protecting and improving the marine environment, Marine Thematic Report, 2014. Available at: (http://ec.europa.eu/environment/life/publications/lifepublications/generalpublications/documents/marine_report.pdf).
- [36] P. Ten Brink, I. Lutchman, S. Bassi, S. Speck, S. Sheavly, K. Register, C. Woolaway, Guidelines on the Use of Market-based Instruments to Address the Problem of Marine Litter, Institute for European Environmental Policy (IEEP), Brussels, Belgium, and Sheavly Consultants, Virginia Beach, Virginia, USA, 2009, 60 pp.
- [37] Defra, Marine Strategy Framework Directive consultation: UK Initial Assessment and Proposals for Good Environmental Status, Impact Assessment, HM Government, 2012.
- [38] Wildlife and Countryside Link, Scottish Environment LINK, Wales Environment Link and the Northern Ireland Marine Task Force, Marine Strategy Framework Directive consultation: Programme of Measures, 2015.
- [39] T. Berg, K. Fürhaupter, H. Teixeira, L. Uusitalo, N. Zampoukas, *The Marine Strategy Framework Directive and the ecosystem-based approach – pitfalls and solutions*, *Mar. Pollut. Bull.* 96 (2015) 18–28.
- [40] ICES, Towards management of multiple stressors in the context of the marine strategy framework directive: a Baltic Sea pilot study, STAGES Project, Marine Research Centre, Finnish Environment Institute, International Council for the Exploration of the Sea, NIVA Denmark Water Research, Stockholm Resilience Centre, University of Stockholm, 2014.
- [41] A. Kontogianni, C. Tourkoulas, D. Damigos, M. Skourtos, Assessing sea level rise costs and adaptation benefits under uncertainty in Greece, *Environ. Sci. Policy* 37 (2014) 61–78.
- [42] AKTiiVSLtd, Harmonising the Economic and Social Analysis for the MSFD Programmes of measures in the project's countries – Recommendations' Report, Project GES-REG, 2013.