RESEARCH ARTICLE

A fuzzy logic expert system for evaluating policy progress towards sustainability goals

Andrés M. Cisneros-Montemayor , Gerald G. Singh,
William W. L. Cheung

Received: 26 June 2017 / Revised: 3 November 2017 / Accepted: 22 November 2017 / Published online: 16 December 2017

Abstract Evaluating progress towards environmental sustainability goals can be difficult due to a lack of measurable benchmarks and insufficient or uncertain data. Marine settings are particularly challenging, as stakeholders and objectives tend to be less well defined and ecosystem components have high natural variability and are difficult to observe directly. Fuzzy logic expert systems are useful analytical frameworks to evaluate such systems, and we develop such a model here to formally evaluate progress towards sustainability targets based on diverse sets of indicators. Evaluation criteria include recent (since policy enactment) and historical (from earliest known state) change, type of indicators (state, benefit, pressure, response), time span and spatial scope, and the suitability of an indicator in reflecting progress toward a specific objective. A key aspect of the framework is that all assumptions are transparent and modifiable to fit different social and ecological contexts. We test the method by evaluating progress towards four Aichi Biodiversity Targets in Canadian oceans, including quantitative progress scores, information gaps, and the sensitivity of results to model and data assumptions. For Canadian marine systems, national protection plans and biodiversity awareness show good progress, but species and ecosystem states overall do not show strong improvement. Well-defined goals are vital for successful policy implementation, as ambiguity allows for conflicting potential indicators, which in natural systems increases uncertainty in progress evaluations. Importantly, our framework can be easily adapted to assess progress

towards policy goals with different themes, globally or in specific regions.

Keywords Adaptive management · Aichi Biodiversity Targets · Convention on Biological Diversity · Fuzzy logic · Policy evaluation

INTRODUCTION

Modern sustainable development goals explicitly include minimizing cross-scale human impacts on biodiversity and ecosystem services (UN 2015), and, concurrently, ecosystem management plans increasingly integrate social and economic dimensions (IPCC 2014; Steffen et al. 2015). This more adequately considers real-world contexts but presents the challenge of developing strategies that co-achieve complex social–ecological goals (Singh et al. 2017a) that further require reliable and transparent frameworks and metrics to evaluate progress towards specific objectives (Marques et al. 2014; Nilsson et al. 2016).

Contextualizing indicators into states (e.g., mangrove coverage), pressures (e.g., rate of coastal development), benefits (e.g., yearly fish recruitment, carbon sequestration), and responses (e.g., protected area coverage, alternative livelihood investment) has been useful for tracking progress towards conservation policies (Tittensor et al. 2014), evaluating linkages between objectives, and assessing effectiveness of achieving stated goals (Nilsson et al. 2016). Mechanistic processes of impacts can also be included to link human activities and other drivers of change with stressors on species and ecosystems (i.e., the Driver Pressure State Impact Response, or DPSIR, approach; Curtin and Prelezo 2010). Tittensor et al. (2014)

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13280-017-0998-3>) contains supplementary material, which is available to authorized users.

used such a framework to evaluate global progress towards the Convention on Biological Diversity (CBD) 2020 Aichi Biodiversity Targets (ABTs) (UNEP 2010) by averaging trends in various global indicators. The authors found that biodiversity conservation objectives are unlikely to be met by the target year, but mostly estimated general trends as opposed to quantifiable progress because of uncertainties in data and definitions of success.

There are three key challenges in evaluating sustainability policy progress. First, the lack of quantitative benchmarks for complex sustainability goals makes it difficult to evaluate how close we are to achieving policy success, particularly when goals are ambiguous, broad, or do not fully reflect stakeholder objectives (Tallis et al. 2010). Even sophisticated evaluations are thus often inherently limited to quantification of the direction and magnitude of change (e.g., Tittensor et al. 2014). Second, evaluations are challenged by uncertain or incomplete data (Joppa et al. 2016), requiring the use of proxy and composite indicators (Fulton et al. 2005). Third, quantitative ecology methods have difficulty incorporating qualitative expert local and traditional knowledge (LTK), thus excluding highly useful information and hindering collaboration with communities at the crux of social–ecological dynamics (Perera et al. 2012).

Here, we develop a fuzzy logic expert system—a program that emulates decision-making according to set decision rules—to address such methodological challenges in evaluating progress towards complex sustainability goals. Fuzzy logic has the advantage of formally dealing with the uncertainties, incompleteness, and interpretation of indicators to address a given problem (Zadeh 1965), and has been applied to evaluate ecosystem-based management plans (Adriaenssens et al. 2004), the vulnerability of species to climate change (Jones and Cheung 2017), global jellyfish abundance trends (Brotz 2011), usefulness of conservation strategies (Andriantiatsaholiniaina et al. 2004), and the environmental sustainability of human development (Phillis and Andriantiatsaholiniaina 2001). In a management context, fuzzy logic methods are also accepted by the International Union for the Conservation of Nature (IUCN) to establish criteria and inform the classification of endangered species (Akçakaya et al. 2000).

Fuzzy logic departs from traditional binary logic (e.g., 0/1, hot/cold), allowing a single value to simultaneously be a member of multiple overlapping states (e.g., “warm” and “hot”). This allows for uncertainty or differing perceptions when making inferences. For example, although temperature can be precisely measured, the perception of hotness/coldness at a certain temperature (say, 20 °C) varies between people. Fuzzy logic allows for the association of both “hot” and “warm” categories to the same temperature (fuzzification). Association to each category triggers

different sets of rules that result in multiple, possibly contrasting, signals, again with different degrees of association (fuzzy inferences). Finally, these signals are combined into a single metric and conclusion (defuzzification). The design and inputs to this framework (fuzzification, fuzzy inferences, and defuzzification) are based on available data and expert knowledge (Adriaenssens et al. 2004). Expert elicitation techniques can be used to establish evaluation criteria (Mackinson 2001; Burgman 2005; Martin et al. 2012) and to provide input when no other data exist, but must be rigorously implemented to avoid biases (Perera et al. 2012).

There are multiple useful methods to monitor biodiversity or environmental state at national or subnational scales, such as the Ocean Health Index (OHI; Halpern et al. 2012), Nested Environmental Status Assessment Tool (NEAT; Uusitalo et al. 2016), and the HELCOM Biodiversity Assessment Tool (BEAT; Andersen et al. 2014), some of which are quite flexible in using multiple types of data (e.g., Backer and Leppänen 2008) and some specifically attempting to approximate the calculation of uncertainty of fuzzy logic methods (Newton 2010). The fuzzy logic system developed in this study, however, is not intended to evaluate environmental state—though it could conceivably be used to that end—but the progress status of an interdisciplinary policy goal that may include environmental state as a component of desired outcomes. Indeed, outputs from existing environmental indicator frameworks would be ideal inputs for this evaluation.

There are four key qualities that this fuzzy logic approach capitalizes on: (1) the ability to combine multidisciplinary and multiscale data, which is vital for new integrated policy goals; (2) transparency of the entire framework (assumptions, data considered, scoring mechanisms, etc.), which allows for and indeed encourages multi-stakeholder collaboration; (3) flexibility to make adaptive modifications to assumptions or data inputs, while still maintaining consistent evaluation criteria with results that can be compared over time and across regions; (4) making uncertainty and differences in perceptions explicit, which is important for integrated policies that must necessarily deal with multiple uncertainties in data, system linkages, and stakeholder preferences and objectives. To be clear, the principles above guide the development of many environmental indicator frameworks (e.g., Halpern et al. 2012; Andersen et al. 2014; Tittensor et al. 2014; Uusitalo et al. 2016); for fuzzy logic approaches, however, these principles are essential components rather than desirable traits (Akçakaya et al. 2012).

The fuzzy logic expert system developed in this study explicitly integrates different types of indicators to provide a quantitative progress score for environmental sustainability goals, in this case CBD Aichi Biodiversity Targets

(UNEP 2010) specifically within marine systems. The guidance of national and intergovernmental environmental policies is important to achieve global sustainable development, but requires formal methods for transparently evaluating policy progress and identifying research gaps and investment priorities. We examine the sensitivity of this method to data variability and assumptions of model structure, and discuss potential applications of this fuzzy logic framework to other environmental and sustainability policies, such as the UN Sustainable Development Goals (UN 2015).

MATERIALS AND METHODS

We developed a fuzzy logic expert system to score environmental policy progress based on the direction and relative magnitude of trends for diverse biological and socioeconomic indicators. The criteria and framework of this model are described in depth here, but we first provide a detailed primer on the fuzzy logic expert system key terms and evaluation framework which form the basis for our subsequent model.

Fuzzy logic expert systems

The main components of fuzzy logic expert systems include the problem (i.e., progress towards biodiversity targets), its corresponding space (i.e., potential indicator values), and criteria informing conclusions (Cox 1999; Fig. 1). Each criterion has its own space of possible indicator values and is divided into overlapping fuzzy sets (heuristic categories) that each comprise a range of values (or domain) (Fig. 1a, b). Indicator values have different degrees of membership [0,1] in corresponding sets which are determined by fuzzy membership functions. These functions are most often triangle distributions except for sets at the upper and lower bounds of the criteria or problem space, where they take on a trapezoid shape (i.e., any value beyond these limits will have a membership = 1) (Fig. 1a, b). Note that membership functions could be assigned any shape, including normal or gamma curves, or irregular user-defined functions (Wood et al. 2007) that could be more suitable depending on the application (Cox 1999).

Defuzzification and solution score

Criteria sets and associated degrees of membership are combined formally through fuzzy reasoning and a defuzzification process to provide a solution, in this case progress towards policy goals. First, a heuristic conclusion and its associated indicator value are matched to its

membership in the corresponding criteria sets, each of which has a corresponding problem set. In Fig. 1, for example, the ‘Stable’ and ‘Increasing’ sets of Criterion 1 (‘Abundance change’) correspond to the ‘Moderate’ and ‘Good’ sets of the Problem (‘Policy progress’) being evaluated (Table 1). In this example, an indicator value of 0.1 in ‘Abundance change’ has a membership of 0.65 in the ‘Stable’ fuzzy set and 0.3 in the ‘Increasing’ set, so these membership scores are transposed to the corresponding Problem sets (Table 1; Fig. 1c). In the case of set shapes with two values for one membership (i.e., two sides of the triangle), the median score for the Problem set is assigned. The resulting problem membership distribution (fuzzy region) can be displayed graphically by truncating each problem set at its maximum membership (Fig. 1d). A confidence interval for the final problem score is calculated based on fuzzy region values at a predetermined threshold (α), indicating the minimum degree of membership (or belief) acceptable to include criteria sets in the problem evaluation (Fig. 1d).

As shown in Fig. 1c, a problem solution very likely includes multiple sets; in this example, ‘Policy progress’ would be Moderate–Good. A final score (S) is then calculated by defuzzifying criteria information. The most common defuzzification technique is the centroid method (Cox 1999), a weighted average incorporating all individual criteria (i) scores (s) and memberships (m) above the α value:

$$S = \frac{\sum_i (s_i \cdot m_i \cdot w_i)}{\sum_i (m_i \cdot w_i)}, \quad (1)$$

where w is an optional weighting factor applied to each criterion, such that a higher value is placed on specific criteria (for example, if ‘Abundance change’ was judged to be more important for policy success than ‘Spatial scope’). Table 1 shows the scoring details for the example in Fig. 1, which, based on Eq. 1, would result in a final ‘Policy progress’ score of 59 (Fig. 1d).

Evaluation of sustainability policy progress

Our fuzzy logic expert system builds on previous frameworks that classify indicator types into states, pressures, benefits, or responses with different desired directions and magnitudes of change (Phillis and Andriantiatsaholainaina 2001; Tittensor et al. 2014). Criteria to be informed from each indicator were selected based on their contribution to achieving sustainability goals and their flexibility in allowing for information of different types, which is vital when dealing with complex social–ecological issues (Heink and Kowarik 2010). Given a specific policy goal, appropriate indicators, and available data, policy progress (in heuristic and quantitative terms) is evaluated through

Fig. 1 Illustrative diagram of algorithms used in reasoning through fuzzy logic. **a, b** Evaluation of criteria memberships of indicator values. In **b**, the example x -axis value corresponds to both the 'Large' and 'Very Large' categories, each with a membership of 0.5. **c** Compilation of criteria information for the problem being evaluated. Memberships for each of the criteria sets intersect their corresponding problem set (see Table 1) and x -axis values provide problem scores (e.g., Criterion 1 'Increasing' set in **a** corresponds to Problem 'Good' set in **c**; Table 1). For triangle-shaped problem sets, memberships are assigned to median set x -axis value (i.e., Criterion 1 'Stable' and Criterion 2 'Large,' and Problem 'Moderate'). **d** Problem fuzzy region, final score, and confidence interval given $\alpha = 0.25$

Table 1 Illustrative example of defuzzification of results from information shown in Fig. 1

Criteria	Indicator value	Set	Membership	Problem Set	Problem Score
Abundance Change	0.1	Stable	0.65	Moderate	50
Abundance Change	0.1	Increasing	0.3	Good	65
Spatial Scope	Ocean	Large	0.50	Moderate	50
Spatial Scope	Ocean	Very Large	0.50	Good	75

the fuzzy logic model based on consistent criteria and rules (Fig. 2).

Our criteria include the following: (i) 'Recent' (change since inception of policy); (ii) 'Absolute' (change from earliest known state); (iii) 'Time' (length of indicator time series); (iv) 'Space' (spatial scope of indicator); and v)

'Suitability' (of the indicator regarding the target, used as a weighting factor). A general assumption is that longer data time series collected over larger spatial areas would more accurately reflect national-level progress relative to goal benchmarks and baseline values. Note that these criteria were developed and selected in part considering the types

Fig. 2 Framework for evaluation of progress towards sustainability policy goals, including identification of indicators, fuzzy logic evaluation, progress score and key factors, and feedback for adaptive management

of data commonly available at national scales, but there are many other potential criteria that could be used in similar approaches given specific policies and contexts (Adriaenssens et al. 2004) (see “Discussion” section).

Indicators with a decreasing trend would receive a positive score if classified as a pressure (e.g., decreasing fishing capacity), but a negative score if classified as a state or response (e.g., decreasing funding for habitat protection). In our framework, indicator trends are further differentiated and evaluated through the ‘Recent’ and ‘Absolute’ criteria to reflect the fact that new policies are intended to have a new and specific effect. The ‘Time’ and ‘Space’ criteria are similar to those proposed by Brotz et al. (2012) to determine changes in regional jellyfish abundances based on multiple data sources and types. These scope criteria result in a better overall score for indicators that provide information over a long period, and reflect trends over a large spatial area (Table 2). Note, however, that depending on the application this initial assumption may be modified. For example, special consideration might be given to indicators that are relatively new and/or not updated often but are comprehensive. Similarly, local policy evaluations would place higher value on indicators for local rather than large areas.

Finally, the ‘Suitability’ criterion is used to inform the relative appropriateness of an indicator to evaluate different policy goals. For example, a time series of the number of whale watching operators in the St. Lawrence River might be a good indicator of human economic benefits, as well as an adequate but indirect indicator of pressure on beluga populations (Blane and Jaakson 1994); in our framework, this indicator could inform both criteria, with

higher weight on the former. A detailed explanation for the criteria is presented in Table 2, and the associated fuzzy sets in Fig. 3.

Test case: CBD Aichi Biodiversity Targets

We test the fuzzy logic expert system described above using a key international policy, the Convention on Biological Diversity (CBD) Aichi Biodiversity Targets (ABTs), which include twenty Targets to be reached by 2020 (UNEP 2010). To date (2017), 196 states have signed on to the CBD as a guide for their national strategies. As the CBD was ratified in 1993, this becomes the baseline year used in the ‘Recent’ criteria (Table 2; Fig. 3). Many indicators have been proposed as reflective of overall progress and trends (Canadian Council of Resource Ministers 2014; Tittensor et al. 2014), but not all ABTs have quantitative objectives. Indeed, the current lack of a formal process for evaluating progress (Marques et al. 2014; Tittensor et al. 2014) in part prompted this study.

Drawing from an existing metadata repository of ocean-related research available for Canada (Cisneros-Montemayor et al. 2017), we compiled indicators ($n = 146$) and applied our method to evaluate advancement toward four ABTs. All indicators and the corresponding references are provided in Supplementary Materials (Table S1). The test Targets highlight our method’s strengths and limitations, and were selected to include varying (a) clarity of target success, (b) availability of indicators, and (c) agreement between available indicator trends:

Table 2 Evaluation criteria rationale, fuzzy sets, indicator values, and the corresponding progress sets

Criteria	Set	Indicator values	Corresponding progress set	Metric and rationale
Recent	Very good	$0.2 \leq \text{Ratio}$	Very High	Ratio between most recent value and earliest value since policy implementation. This represents progress after agreement to curtail unsustainable activity. An indicator of ecosystem state (e.g., species abundance) would have a better score the greater the positive change, and opposite for an indicator of ecosystem pressure (e.g., pollutant load)
	Good	$0 \leq \text{Ratio} < 0.4$	High	
	Medium	$-0.2 < \text{Ratio} < 0.2$	Medium	
	Poor	$\text{Ratio} \leq 0$	Low	
Absolute	Very good	$0.2 \leq \text{Ratio}$	Very High	Ratio between most recent value and earliest available value. The current state of an indicator relative to its earliest known state shows overall progress towards target success, e.g., recent recovery rate of a wild population may be very good, but abundance still far from previous known levels. See above for handling of ecosystem state versus pressure indicators
	Good	$0 \leq \text{Ratio} < 0.4$	High	
	Medium	$-0.2 < \text{Ratio} < 0.2$	Medium	
	Poor	$\text{Ratio} \leq 0$	Low	
Time	Very long	$12 \leq \text{Years}$	Very High	Length of data time series (years). Longer time series of indicator data allow for establishment of meaningful baselines and improved analyses and policy development. Although frequent data points along a time series are highly useful, only total length (i.e., last and first year of data) is evaluated here
	Long	$8 \leq \text{Years} < 16$	High	
	Medium	$4 < \text{Years} < 12$	Medium	
	Short	$\text{Years} \leq 8$	Low	
Space	Large	Ocean or Nation	Very High	Spatial scope of indicator. As most sustainability targets consider a national scale, indicators available for a small (e.g., city-level) spatial area can be indicative of progress, but do not carry the same weight as the ones available for much larger (e.g., ocean basin) areas. This would change depending on the spatial scope of the policy being evaluated
	Medium	Region or Province	High	
	Small	Point location or Statistical Area	Medium	
Suitability	–	–	–	Suitability weighting rank from 1 (medium) to 10 (excellent) (re-scaled in model from 0 to 1). Due to the language and broad goals of many sustainability goals, it can be difficult to select precise indicators for each target. The suitability weighting factor allows for the use of all available data while explicitly recognizing their appropriateness. Using a weighting factor (as opposed to another fuzzy criterion) facilitates using the same indicator to inform multiple targets, where the indicator can have different suitability values for each target

ABT 1 People are aware of the values of biodiversity and the steps they can take to conserve and use it sustainably [unclear goal, few indicators, conflicting trends].

ABT 3 Incentives, including subsidies, harmful to biodiversity are eliminated, phased out, or reformed in order to minimize or avoid negative impacts, and positive incentives for the conservation and sustainable use of biodiversity are developed and applied, consistent and in harmony with the Convention and other relevant international obligations, taking into account national socio-economic conditions [clear goal, few indicators, agreeing trends].

ABT 6 All fish and invertebrate stocks and aquatic plants are managed and harvested sustainably, legally, and applying ecosystem-based approaches, so that overfishing is avoided, recovery plans and measures are in place for all depleted species, fisheries have no significant adverse impacts on threatened species and vulnerable ecosystems, and the impacts of fisheries on stocks, species, and ecosystems are within safe ecological limits [unclear goal, many indicators, conflicting trends].

ABT 17 Each Party has developed, adopted as a policy instrument, and has commenced implementing an

effective, participatory, and updated national biodiversity strategy and action plan [unclear goal, few indicators, agreeing trends].

Sensitivity analyses

Sensitivity analyses can reveal biases in a model framework itself, as well as data gaps and uncertainties. Our sensitivity tests thus address both the number of indicators and data values and the criteria included in the evaluation. We tested the sensitivity of results to the model structure by re-calculating scores when each of the five criteria was excluded from the analysis (Phillis and Andriantiatsaholainaina 2001). This displays effects from the inclusion or exclusion of individual criteria, showing potential biases on the final score that may signal a need for further scrutiny of a criterion's attributes (Cheung et al. 2005).

Sensitivity to available data was tested using a jackknife approach omitting an increasingly large (0–80 %) random set of available indicators for each target. Scores were re-calculated using each random subset of indicators, and the exercise was repeated 1000 times for each target and subset

Fig. 3 Fuzzy sets for sustainability policy evaluation criteria and policy progress score

proportion. This analysis can be extended to account for the weight of specific indicators on target progress by sequentially omitting each single indicator for a given target and comparing the resulting progress scores with baseline values.

Finally, we re-calculated results under a wide range of α cut values, representing the threshold membership below which information is excluded from the analysis. A low α value indicates that any information, no matter how small a membership in a given set, is included in scores and confidence bounds. The effect of increasing α values on resulting confidence bounds reflects the agreement between

indicators in a given policy evaluation, highlighting the various levels of uncertainty that can be encountered when dealing with multiple information inputs (Brotz et al. 2012).

RESULTS

The fuzzy logic expert system developed here provided progress scores for test policies; the effects of both available data and the model framework itself were subsequently tested through sensitivity analyses. The selected

Fig. 4 Fuzzy regions and progress scores for sample Aichi Biodiversity Targets (as numbered on panels, top right) using available data for Canadian marine systems (Cisneros-Montemayor et al. 2017). Black arrows indicate the final progress score; gray arrow indicate the recent change score; dashed lines indicate the α value and resulting confidence intervals for final progress score

ABTs with highest progress score were ABT 17 (progress score, PS = 79), ABT 3 (PS = 72), ABT 1 (PS = 71), and ABT 6 (PS = 54). For ‘Recent change’ (relative to 1993), the highest positive change was for ABT 3 (recent score, RS = 87), ABT 17 (RS = 69), ABT 1 (RS = 56), and ABT 6 (RS = 33). These scores involve a transparent integration of multiple signals for each target that are summarized in fuzzy regions showing the uncertainty surrounding each score (Fig. 4). Note that the wide array of (often conflicting) indicators results in wide confidence bounds (range at α cut; Fig. 4).

Overall progress scores were robust to the criteria used, with the ‘Time’ criterion contributing to higher final scores (meaning that time series tended to be relatively long). The inclusion of the ‘Absolute’ and ‘Space’ criteria resulted in slightly lower scores, meaning that in the Canadian test case trends were worse relative to historical baselines, and indicators tended to reflect relatively smaller areas (Fig. 5).

The sensitivity of policy progress scores to the information used was tested by omitting an increasingly large subset of available indicators from the evaluation of each target. Targets with many available indicators had more robust results than targets with less indicators (Fig. 6). Note, however, that having many available indicators also increases the chances of conflicting information, leading to high uncertainty surrounding final progress scores (Fig. 4). As further discussed below, this uncertainty can also reflect the clarity and specificity of policies themselves.

In addition to the analysis shown in Fig. 6, we identified single indicators that have the most weight on final progress scores in terms of percentage difference from the baseline score. The more the indicators available for a target, the less weight any single indicator has on the final score (Table 3); given the same number of indicators, increased agreement between them would decrease the influence of any single indicator on the final score.

Fig. 5 Sensitivity of policy progress score to evaluation criteria over all sample Aichi Biodiversity Targets after omitting each indicated criterion. ‘Baseline’ shows score including all criteria. Horizontal dashed line indicates the median progress score over all tests

Fig. 6 Sensitivity of final progress score to the number of available indicators ($n = 4$ is ABT 1; $n = 125$ is ABT 6). For each target, an increasingly large, random subset of indicators was omitted from the analysis. Boxes show resulting scores over 1000 iterations of random subsets. Horizontal dashed line indicates the median progress score over all tests

DISCUSSION

The fuzzy logic expert system developed here allows for evaluating progress towards sustainability policy goals,

highlighting the key factors in progress, and prioritizing research and collection of information. Fuzzy logic is already being used in environmental management to establish impacts and threat levels to species (Akçakaya

Table 3 Percent change from baseline target progress score if indicator is included. We show a selected subset of indicators with largest positive and negative effects on each Aichi Biodiversity Target (ABT)

ABT (indicators)	Indicator	Δ Progress score (%)	Δ Recent change score (%)
1 (4)	Canadians participating in wildlife or nature-related activities	− 11	− 21
	Whale watching tour participants	+ 13	+ 27
	Recreational fishers (active)	− 9	− 34
	Ocean Wise partners	+ 8	+ 32
3 (3)	Ratio of beneficial to capacity-enhancing subsidies	− 6	− 11
6 (125)	Number of listed marine Species At Risk	+ 0.7	+ 2
	Atlantic cod biomass (NAFO Subdivision 3PS)	− 0.3	− 0.9
7 (3)	Fisheries and Oceans Canada operating budget	− 18	− 27
	Number of listed marine Species At Risk	+ 11	+ 19

et al. 2000; Keith et al. 2013) and habitats (Wood et al. 2007), but the method proposed here is intended to provide insights into the formulation and progress of a policy itself.

The evaluation of complex policies must always consider qualitative aspects of objectives and intent, indicator selection, and contextual factors of implementation (Tallis et al. 2010), yet a final quantitative score is highly useful for conveying outcomes to stakeholders (Andriantatsaholainaina et al. 2004). In our test case using four Aichi Biodiversity Targets (ABTs) in Canadian oceans, the highest progress scores were for ABT 17 ('Policy Implementation'; Progress Score, PS = 79), followed by ABT 3 ('Eliminate Subsidies,' PS = 72), ABT 1 ('Biodiversity awareness,' PS = 71), and ABT 6 ('Sustainable Stocks,' PS = 54) (Fig. 4). Notably, scores for recent (since the inception of the CBD in 1993) indicator change were worse for three of the four targets evaluated (Fig. 4). The uncertainty associated with complex environmental policy targets is reflected in wide confidence bounds, though ABT 17 shows potential for high agreement even using few indicators (Fig. 4). Note that the criteria developed for this test case, although informed by scientific evidence, would in practice be further revised through stakeholder and expert consultation for application to a specific policy or region. Fuzzy logic is not a substitute for high-quality information, but an avenue to facilitate its use and integration.

Coastal communities and resource users interact most closely with marine ecosystems, and Local and Traditional Knowledge (LTK) can be a major contributor to biodiversity assessments. A major benefit of fuzzy logic systems—particularly useful for management processes related to multi-stakeholder environmental issues—is their ability to work with a variety of data types (Zadeh 1983), which facilitates the collaboration between multiple stakeholders. Expert elicitation is a valuable supporting tool for such collaboration using fuzzy logic (Mackinson 2001) and is increasingly utilized in environmental science and management (Burgman 2005; Morgan 2014). There are a variety of specific elicitation techniques, some specifically for fuzzy logic systems (Gaines and Shaw 1986; Cornelissen et al. 2003), but in general these help select indicators and define membership functions (Cornelissen et al. 2003). For example, a simple approach would elicit from experts the values of the bounds and the highest degree of membership in a heuristic category. Another method could ask experts to define the qualitative categories for specific indicator values and subsequently define membership functions based on the range of responses. In all cases, elicitation techniques based on reducing overconfidence and combating cognitive biases should be followed (for reviews on this topic, see Ayyub 2001, Burgman 2005, and Morgan 2014), requiring technical and procedural knowledge of both elicitation and fuzzy logic for effective evaluation.

Quantitative scores are highly useful for benchmarking progress and achieving policy targets, though do not eliminate the need for a careful qualitative consideration of available information, its appropriateness as indicators for given policies, the intent of policy targets, and priorities and limitations for management (Dale and Beyeler 2001; Tallis et al. 2010). The integration of the DPSIR approach reflects systems-thinking aligned with the multi-objective approaches of sustainability goals and allows for considering concurrent negative and positive impacts from a single source (Atkins et al. 2011). In our test case, results highlight lagging progress on achieving sustainable marine populations compared to the other evaluated targets, overall and since the inception of the CBD (ABT 6; Fig. 4). Results also specifically point out the positive contributions of factors like whale watching and market-based incentives to achieving targets, and detrimental gaps in specific funding (Table 3). These are quantitative results from the fuzzy logic model that can prompt qualitative discussions on the suitability of particular indicators for different policies, and can be paired with more detailed analyses of existing management and information states (Ricard et al. 2012; Cisneros-Montemayor et al. 2017).

Just as important as available information is the construction of the fuzzy evaluation framework itself,

including criteria and their corresponding rules (Adriaenssens et al. 2004) (Table 2). Our method extends previous fuzzy logic frameworks used to evaluate environment-related trends (Mackinson 2000; Brotz et al. 2012), together with similar efforts (not using fuzzy logic) to gauge progress towards environmental policies (Tittensor et al. 2014). A key novel aspect is the explicit inclusion of quantitative policy progress scores as result of the evaluation (Fig. 4), which allows for formal comparisons but also stresses the proper determination of criteria and membership sets, and policies themselves (Fig. 3). As noted above, expert elicitation can be highly useful for informing fuzzy logic applications (Mackinson 2001; Andriantatsaholainaina et al. 2004), though must be performed carefully both in framing analyses (Morgan 2014) and for establishing specific rules (Wood et al. 2007). Although very sophisticated methods are available for eliciting such information, in terms of final results there does not seem to be much added value in creating highly complex fuzzy logic frameworks, which would require more training in order to avoid misleading input stemming from misunderstood problem structures (Cornelissen et al. 2003; Martin et al. 2012; McBride and Burgman 2012), so simplicity of application should be a priority.

Sensitivity analyses show that policy progress scores were robust to various model structures, with the ‘Time’ criterion resulting in slightly higher progress scores for the test ABTs (Fig. 5). Note that we did not account for potential indirect relationships between criteria, e.g., a longer time series would result in a higher progress score, but may also represent a higher baseline value for ‘Absolute’ change. The potential effects of such relationships can be readily tested, but the important point is that such linkages and the criteria used must be considered and developed to fit specific policies and social–ecological contexts (Adriaenssens et al. 2004). For example, one might include criteria related to the type of information being used, who produced it or how, yearly variability, degree of support for the use of an indicator (for example, through expert consultation or literature reviews), the relationships between multiple indicators (for example, interacting environmental drivers), etc. These region- and goal-specific refinements are crucial, as most actions in global policies (e.g., UN SDGs) are intended to be taken at national or subnational scales. Therefore, the inclusion and development of new indicators and criteria, as well as the determination of what are deemed “good enough” data, will necessarily be context specific (Andriantatsaholainaina et al. 2004; Svarstad et al. 2008). Nevertheless, it is possible to arrive at some general indicators that are reasonable to collect, while still being indicative of much more complicated processes relevant to policy success (Pereira et al. 2013).

The results of this study reflect three key factors in formally evaluating sustainability policy progress: availability of appropriate indicator information, and the clarity and complexity of the policy goal. Unsurprisingly, a policy target with many available indicators results in a final progress score that is more robust than one with few available indicators, so long as the indicators’ trends are relatively consistent (Fig. 6). However, when a goal is ambiguous, a large number of potentially acceptable but less specific indicators may provide conflicting evidence (Dale and Beyeler 2001). In the case of ABT 1 (‘Biodiversity Awareness’), conflicting trends likely reflect uncertainty and lack of clarity in the intent of the target itself when many types of indicators could credibly be included. (For example, do more seafood labeling programs reflect awareness in the public? If people are more “aware” of the value of biodiversity but choose not to protect it, would this constitute progress?) For ABT 6 (‘Sustainable Stocks’), however, conflicting trends rather reflect complexity. There are (sometimes competing) definitions of whether a stock is “sustainable,” but *all* stocks within an ecosystem could never be expected to share the same abundance trends, regardless of how these are evaluated.

These inherent barriers to “success as stated” must first be addressed through modifications of policies themselves, ideally following ‘SMART’ criteria to create goals that are Specific, Measurable, Assignable, Realistic, and Time-bound (Doran 1981). This enables the use of composite indicators that reduce complexity (Heink and Kowarik 2010) and reflect system status rather than solely that of individual species or stocks (Dale and Beyeler 2001; Froese 2004; Fulton et al. 2005; Shin et al. 2010). Fuzzy logic frameworks can be used to flag such cases, promoting the adoption of clear and achievable goals and facilitating participatory modeling to identify the indicators (and weighting values) that most impact policy success.

CONCLUSION

Strategic goals and associated targets—such as the CBD Aichi Targets (UNEP 2010), Sustainable Development Goals (UN 2015), and the Millennium Development Goals before them—are critical to modern sustainable development and to avoid global catastrophes, but actions and progress must be formally evaluated for policies to be truly meaningful (Heink and Kowarik 2010). In this study, we address the general lack of standardized approaches to evaluate complex sustainability policies and develop a fuzzy logic expert system to assess and score progress. Fuzzy logic is a useful framework, particularly in facilitating the transparent integration of stakeholder perceptions

and multidisciplinary information in evaluation criteria (Wood et al. 2007). The growing demand for interdisciplinary sustainability goals, paired with rapid global changes, requires analytical tools that are flexible, easily updatable, and can incorporate broad (often piecemeal) data and expert input (Singh et al. 2017b). A key message that is furthermore highlighted by the results of this study is that quantitative evaluation of policy progress can only be as effective as the goals themselves are clear and achievable (Turnhout et al. 2007), and the iterative and adaptive process of policy development and formal evaluation is thus key to regional and global success.

Acknowledgements This is a product of the OceanCanada Partnership, funded by the Social Sciences and Humanities Research Council of Canada, and the Nippon Foundation Nereus Program, a collaborative initiative by the Nippon Foundation and partners including the University of British Columbia.

REFERENCES

- Adriaenssens, V., B.D. Baets, P.L.M. Goethals, and N.D. Pauw. 2004. Fuzzy rule-based models for decision support in ecosystem management. *Science of the Total Environment* 319: 1–12. [https://doi.org/10.1016/S0048-9697\(03\)00433-9](https://doi.org/10.1016/S0048-9697(03)00433-9).
- Akçakaya, H.R., S. Ferson, M.A. Burgman, D.A. Keith, G.M. Mace, and C.R. Todd. 2000. Making consistent IUCN classifications under uncertainty. *Conservation Biology* 14: 1001–1013.
- Akçakaya, H.R., S. Ferson, M.A. Burgman, D.A. Keith, G.M. Mace, and C.R. Todd. 2012. Commentary: IUCN classifications under uncertainty. *Environmental Modelling & Software* 38: 119–121. <https://doi.org/10.1016/j.envsoft.2012.05.009>.
- Andersen, J.H., K. Dahl, C. Göke, M. Hartvig, C. Murray, A. Rindorf, H. Skov, M. Vinther, et al. 2014. Integrated assessment of marine biodiversity status using a prototype indicator-based assessment tool. *Frontiers in Marine Science*. <https://doi.org/10.3389/fmars.2014.00055>.
- Andriantiatsaholainaina, L.A., V.S. Kouikoglou, and Y.A. Phillis. 2004. Evaluating strategies for sustainable development: Fuzzy logic reasoning and sensitivity analysis. *Ecological Economics* 48: 149–172. <https://doi.org/10.1016/j.ecolecon.2003.08.009>.
- Atkins, J.P., D. Burdon, M. Elliott, and A.J. Gregory. 2011. Management of the marine environment: Integrating ecosystem services and societal benefits with the DPSIR framework in a systems approach. *Marine Pollution Bulletin* 62: 215–226. <https://doi.org/10.1016/j.marpolbul.2010.12.012>.
- Ayyub, B.M. 2001. *Elicitation of expert opinions for uncertainty and risks*. Boca Raton: CRC Press.
- Backer, H., and J.-M. Leppänen. 2008. The HELCOM system of a vision, strategic goals and ecological objectives: Implementing an ecosystem approach to the management of human activities in the Baltic Sea. *Aquatic Conservation: Marine and Freshwater Ecosystems* 18: 321–334. <https://doi.org/10.1002/aqc.851>.
- Blane, J.M., and R. Jaakson. 1994. The Impact of Ecotourism Boats on the St Lawrence Beluga Whales. *Environmental Conservation* 21: 267–269.
- Brotz, L. 2011. *Changing jellyfish populations: Trends in Large Marine Ecosystems*. Fisheries Centre Research Reports 19. The University of British Columbia.
- Brotz, L., W.W.L. Cheung, K. Kleisner, E. Pakhomov, and D. Pauly. 2012. Increasing jellyfish populations: Trends in large marine ecosystems. *Hydrobiologia* 690: 3–20. <https://doi.org/10.1007/s10750-012-1039-7>.
- Burgman, M. 2005. *Risks and decisions for conservation and environmental management*. New York: Cambridge University Press.
- Canadian Council of Resource Ministers. 2014. Fifth National Report to the United Nations Convention on Biological Diversity. Government of Canada.
- Cheung, W.W.L., T.J. Pitcher, and D. Pauly. 2005. A fuzzy logic expert system to estimate intrinsic extinction vulnerabilities of marine fishes to fishing. *Biological Conservation* 124: 97–111. <https://doi.org/10.1016/j.biocon.2005.01.017>.
- Cisneros-Montemayor, A.M., W.W.L. Cheung, K. Bodtker, L. Teh, N. Steiner, M. Bailey, C. Hoover, and U.R. Sumaila. 2017. Towards an integrated database on Canadian ocean resources: Benefits, current states, and research gaps. *Canadian Journal of Fisheries and Aquatic Sciences* 74: 65–74. <https://doi.org/10.1139/cjfas-2015-0573>.
- Cornelissen, A., J. van den Berg, W. Koops, and U. Kaymak. 2003. Elicitation of expert knowledge for fuzzy evaluation of agricultural production systems. *Agriculture, Ecosystems & Environment* 95: 1–18.
- Cox, E. 1999. *The fuzzy systems handbook: A practitioners guide to building, using, and maintaining fuzzy systems*. San Diego: AP Professional.
- Curtin, R., and R. Prellezo. 2010. Understanding marine ecosystem based management: A literature review. *Marine Policy* 34: 821–830.
- Dale, V.H., and S.C. Beyeler. 2001. Challenges in the development and use of ecological indicators. *Ecological Indicators* 1: 3–10.
- Doran, G.T. 1981. There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review* 70: 35–36.
- Froese, R. 2004. Keep it simple: Three indicators to deal with overfishing. *Fish and Fisheries* 5: 86–91.
- Fulton, E., A. Smith, and A. Punt. 2005. Which ecological indicators can robustly detect effects of fishing? *ICES Journal of Marine Science* 62: 540–551. <https://doi.org/10.1016/j.icesjms.2004.12.012>.
- Gaines, B.R., and M.L. Shaw. 1986. Induction of inference rules for expert systems. *Fuzzy Sets and Systems* 18: 315–328.
- Halpern, B.S., C. Longo, D. Hardy, K.L. McLeod, J.F. Samhour, S.K. Katona, K. Kleisner, S.E. Lester, et al. 2012. An index to assess the health and benefits of the global ocean. *Nature* 488: 615–620. <https://doi.org/10.1038/nature11397>.
- Heink, U., and I. Kowarik. 2010. What are indicators? On the definition of indicators in ecology and environmental planning. *Ecological Indicators* 10: 584–593. <https://doi.org/10.1016/j.ecolind.2009.09.009>.
- IPCC. 2014. *Climate change 2014: Synthesis report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva: IPCC.
- Jones, M.C., and W.W.L. Cheung. 2017. Using fuzzy logic to determine the vulnerability of marine species to climate change. *Global Change Biology*. <https://doi.org/10.1111/gcb.13869>.
- Joppa, L.N., B. O'Connor, P. Visconti, C. Smith, J. Geldmann, M. Hoffmann, J.E.M. Watson, S.H.M. Butchart, et al. 2016. Filling in biodiversity threat gaps. *Science* 352: 416–418.
- Keith, D.A., J.P. Rodríguez, K.M. Rodríguez-Clark, E. Nicholson, K. Aapala, A. Alonso, M. Asmussen, S. Bachman, et al. 2013. Scientific foundations for an IUCN Red List of Ecosystems. Edited by Matteo Convertino. *PLoS ONE* 8: e62111. <https://doi.org/10.1371/journal.pone.0062111>.
- Mackinson, S. 2000. An adaptive fuzzy expert system for predicting structure, dynamics and distribution of herring shoals. *Ecological Modelling* 126: 155–178.

- Mackinson, S. 2001. Integrating local and scientific knowledge: An example in fisheries science. *Environmental Management* 27: 533–545. <https://doi.org/10.1007/s0026702366>.
- Marques, A., H.M. Pereira, C. Krug, P.W. Leadley, P. Visconti, S.R. Januchowski-Hartley, R.M. Krug, R. Alkemade, et al. 2014. A framework to identify enabling and urgent actions for the 2020 Aichi Targets. *Basic and Applied Ecology* 15: 633–638. <https://doi.org/10.1016/j.baec.2014.09.004>.
- Martin, T.G., M.A. Burgman, F. Fidler, P.M. Kuhnert, S. Low-Choy, M. McBride, and K. Mengersen. 2012. Eliciting expert knowledge in conservation science. *Conservation Biology* 26: 29–38. <https://doi.org/10.1111/j.1523-1739.2011.01806.x>.
- McBride, M., and M. Burgman. 2012. What is expert knowledge, how is such knowledge gathered, and how do we use it to address questions in landscape ecology? In *Expert knowledge and its application in landscape ecology*, ed. A.H. Perera, C.A. Drew, and C.J. Johnson, 11–38. New York: Springer. <https://doi.org/10.1007/978-1-4614-1034-8>.
- Morgan, M.G. 2014. Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences* 111: 7176–7184.
- Newton, A.C. 2010. Use of a Bayesian network for Red Listing under uncertainty. *Environmental Modelling & Software* 25: 15–23. <https://doi.org/10.1016/j.envsoft.2009.07.016>.
- Nilsson, M., H. Griggs, and M. Visbeck. 2016. Map the interactions between sustainable development goals. *Nature* 534: 320–322.
- Pereira, H.M., S. Ferrier, M. Walters, G.N. Geller, R.H.G. Jongman, R.J. Scholes, M.W. Bruford, N. Brummitt, et al. 2013. Essential biodiversity variables. *Science* 339: 277–278. <https://doi.org/10.1126/science.1229931>.
- Perera, A.H., C.A. Drew, and C.J. Johnson (eds.). 2012. *Expert knowledge and its application in landscape ecology*. New York, NY: Springer. <https://doi.org/10.1007/978-1-4614-1034-8>.
- Phillis, Y.A., and L.A. Andriantiatsaholainaina. 2001. Sustainability: An ill-defined concept and its assessment using fuzzy logic. *Ecological Economics* 37: 435–456.
- Ricard, D., C. Minto, O.P. Jensen, and J.K. Baum. 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. *Fish and Fisheries* 13: 380–398. <https://doi.org/10.1111/j.1467-2979.2011.00435.x>.
- Shin, Y.-J., A. Bundy, L.J. Shannon, M. Simier, M. Coll, E.A. Fulton, J.S. Link, D. Jouffre, et al. 2010. Can simple be useful and reliable? Using ecological indicators to represent and compare the states of marine ecosystems. *ICES Journal of Marine Science* 67: 717–731.
- Singh, G.G., A.M. Cisneros-Montemayor, W. Swartz, W. Cheung, J.A. Guy, T.-A. Kenny, C.J. McOwen, R. Asch, et al. 2017a. A rapid assessment of co-benefits and trade-offs among sustainable development goals. *Marine Policy*. <https://doi.org/10.1016/j.marpol.2017.05.030>.
- Singh, G.G., J. Sinner, J. Ellis, M. Kandlikar, B.S. Halpern, T. Satterfield, and K.M. Chan. 2017b. Mechanisms and risk of cumulative impacts to coastal ecosystem services: An expert elicitation approach. *Journal of Environmental Management* 199: 229–241.
- Steffen, W., K. Richardson, J. Rockstrom, S.E. Cornell, I. Fetzer, E.M. Bennett, R. Biggs, S.R. Carpenter, et al. 2015. Planetary boundaries: Guiding human development on a changing planet. *Science* 347: 1259855. <https://doi.org/10.1126/science.1259855>.
- Svarstad, H., L.K. Petersen, D. Rothman, H. Siepel, and F. Wätzold. 2008. Discursive biases of the environmental research framework DPSIR. *Land Use Policy* 25: 116–125. <https://doi.org/10.1016/j.landusepol.2007.03.005>.
- Tallis, H., P.S. Levin, M. Ruckelshaus, S.E. Lester, K.L. McLeod, D.L. Fluharty, and B.S. Halpern. 2010. The many faces of ecosystem-based management: Making the process work today in real places. *Marine Policy* 34: 340–348. <https://doi.org/10.1016/j.marpol.2009.08.003>.
- Tittensor, D.P., M. Walpole, S.L.L. Hill, D.G. Boyce, G.L. Britten, N.D. Burgess, S.H.M. Butchart, P.W. Leadley, et al. 2014. A mid-term analysis of progress toward international biodiversity targets. *Science* 346: 241–244. <https://doi.org/10.1126/science.1257484>.
- Turnhout, E., M. Hisschemöller, and H. Eijsackers. 2007. Ecological indicators: Between the two fires of science and policy. *Ecological Indicators* 7: 215–228. <https://doi.org/10.1016/j.ecolind.2005.12.003>.
- UN. 2015. *Transforming our world: The 2030 Agenda for Sustainable Development*. A/Res/70/1. UN General Assembly. New York: United Nations.
- UNEP. 2010. *The strategic plan for biodiversity 2011–2020 and the Aichi Biodiversity Targets*. Conference of the Parties to the Convention on Biological Diversity UNEP/CBD/COP/DEC/XI/2. Nagoya, Japan.
- Uusitalo, L., H. Blanchet, J.H. Andersen, O. Beauchard, T. Berg, S. Bianchelli, A. Cantafaro, J. Carstensen, et al. 2016. Indicator-based assessment of marine biological diversity—lessons from 10 case studies across the European Seas. *Frontiers in Marine Science*. <https://doi.org/10.3389/fmars.2016.00159>.
- Wood, G., A. Rodriguez-Bachiller, and J. Becker. 2007. Fuzzy sets and simulated environmental change: Evaluating and communicating impact significance in environmental impact assessment. *Environment and Planning A* 39: 810–829. <https://doi.org/10.1068/a3878>.
- Zadeh, L. 1965. *Fuzzy logic and its applications*. New York: Academic Press.
- Zadeh, L.A. 1983. The role of fuzzy logic in the management of uncertainty in expert systems. *Fuzzy Sets and Systems* 11: 199–227.

AUTHOR BIOGRAPHIES

Andrés M. Cisneros-Montemayor (✉) is a Research Associate at the University of British Columbia and Program Manager of the Nippon Foundation Nereus Program. His research focuses on the economics of ocean resource use, including fisheries, ecotourism, and sustainable development policy.
Address: Institute for the Oceans and Fisheries, The University of British Columbia, 2202, Main Mall, Vancouver, BC V6T 1Z4, Canada.
e-mail: a.cisneros@oceans.ubc.ca

Gerald G. Singh is a Senior Fellow with the Nippon Foundation Nereus Program based at the University of British Columbia. He specializes in network analyses and expert elicitation methods applied to ecosystem services and natural resource use.
Address: Institute for the Oceans and Fisheries, The University of British Columbia, 2202, Main Mall, Vancouver, BC V6T 1Z4, Canada.
e-mail: g.singh@oceans.ubc.ca

William W. L. Cheung is an Associate Professor at the University of British Columbia, Co-Principal Investigator of the Ocean Canada Partnership, and Science Director of the Nippon Foundation Nereus Program. He specializes in climate change impacts on marine ecosystems and potential mitigation and adaptation measures.
Address: Institute for the Oceans and Fisheries, The University of British Columbia, 2202, Main Mall, Vancouver, BC V6T 1Z4, Canada.
e-mail: w.cheung@oceans.ubc.ca