*Ecological Applications*, 20(5), 2010, pp. 1402–1416


# Using expert judgment to estimate marine ecosystem vulnerability in the California Current

Sarah J. Teck,[1] Benjamin S. Halpern,[2,12] Carrie V. Kappel,[2] Fiorenza Micheli,[3] Kimberly A. Selkoe,[2,4] Caitlin M. Crain,[5] Rebecca Martone,[3] Christine Shearer,[6] Joe Arvai,[7] Baruch Fischhoff,[8] Grant Murray,[9] Rabin Neslo,[10] and Roger Cooke[10,11]

[1]*Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, California 93106 USA*
[2]*National Center for Ecological Analysis and Synthesis, 735 State Street, Santa Barbara, California 93101 USA*
[3]*Hopkins Marine Station, Stanford University, Oceanview Boulevard, Pacific Grove, California 93950 USA*
[4]*Hawai'i Institute of Marine Biology, University of Hawai'i, P.O. Box 1346, Kane'ohe, Hawaii 96744 USA*
[5]*Center for Ocean Health, Ecology and Evolutionary Biology, University of California, Santa Cruz, California 95060 USA*
[6]*Department of Sociology, University of California, Santa Barbara, California 93106 USA*
[7]*Environmental Science and Policy Program, Michigan State University, 305 Natural Resources Building, East Lansing, Michigan 48824 USA*
[8]*Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213 USA*
[9]*Institute for Coastal Research, Vancouver Island University, Nanaimo, British Columbia V9R 5S5 Canada*
[10]*Department of Mathematics, Delft University of Technology, Delft, The Netherlands*
[11]*Resources for the Future, 1616 P Street NW, Washington, D.C. 20036 USA*

*Abstract*. As resource management and conservation efforts move toward multi-sector, ecosystem-based approaches, we need methods for comparing the varying responses of ecosystems to the impacts of human activities in order to prioritize management efforts, allocate limited resources, and understand cumulative effects. Given the number and variety of human activities affecting ecosystems, relatively few empirical studies are adequately comprehensive to inform these decisions. Consequently, management often turns to expert judgment for information. Drawing on methods from decision science, we offer a method for eliciting expert judgment to (1) quantitatively estimate the relative vulnerability of ecosystems to stressors, (2) help prioritize the management of stressors across multiple ecosystems, (3) evaluate how experts give weight to different criteria to characterize vulnerability of ecosystems to anthropogenic stressors, and (4) identify key knowledge gaps. We applied this method to the California Current region in order to evaluate the relative vulnerability of 19 marine ecosystems to 53 stressors associated with human activities, based on surveys from 107 experts. When judging the relative vulnerability of ecosystems to stressors, we found that experts primarily considered two criteria: the ecosystem's resistance to the stressor and the number of species or trophic levels affected. Four intertidal ecosystems (mudflat, beach, salt marsh, and rocky intertidal) were judged most vulnerable to the suite of human activities evaluated here. The highest vulnerability rankings for coastal ecosystems were invasive species, ocean acidification, sea temperature change, sea level rise, and habitat alteration from coastal engineering, while offshore ecosystems were assessed to be most vulnerable to ocean acidification, demersal destructive fishing, and shipwrecks. These results provide a quantitative, transparent, and repeatable assessment of relative vulnerability across ecosystems to any ongoing or emerging human activity. Combining these results with data on the spatial distribution and intensity of human activities provides a systematic foundation for ecosystem-based management.

*Key words: anthropogenic impact; coastal and offshore ecosystems; ecological recovery; ecosystem-based management; ecosystem stressor; ecosystem vulnerability; human impact; resilience; threat assessment.*

## Introduction

Conservation and management efforts must prioritize where to spend resources on mitigating impacts of human activities on the environment. This need has become increasingly apparent in the California Current, a region that stretches roughly from the United States–Canada border to central Baja California, Mexico, due to both increasing human population size (and associated environmental impacts) and increased political will and funding for improving ocean management. Recent efforts to address human impacts to the marine ecosystems of this region include the West Coast Governors' Agreement on Ocean Health, California Ocean Protection Council (OPC), and Marine Life Protection Act (MLPA) Initiative, as well as Oregon's

Manuscript received 30 June 2009; revised 24 September 2009; accepted 30 September 2009. Corresponding Editor: T. E. Essington.
[12] Corresponding author. E-mail: halpern@nceas.ucsb.edu

Ocean Policy Advisory Council (OPAC), and Washington's Puget Sound Partnership, State Oceans Caucus (SOC), and Ocean Policy Advisory Group.

The process of assessing threats to species and the environment and prioritizing actions to mitigate them has a long history. Many methods have been developed by academics, agencies, and conservation NGOs; indeed entire journals and agencies are dedicated to the topic. In the United States, relevant legislation includes the Coastal Zone Management Act, National Environmental Protection Act, Endangered Species Act, Marine Mammal Protection Act, and National Marine Sanctuary Act. All require evaluating the potential negative impacts to species and ecosystems from stressors associated with human activities. Together, these methods have been the focus of thousands of research projects, analyses, and reports (e.g., Smit and Spaling 1995, Council on Environmental Quality 1997, Wilcove et al. 1998). Analogous efforts have been conducted by regional and global conservation organizations such as World Wildlife Fund's ecoregional plans (Olson and Dinerstein 1998) and Conservation International's biodiversity hotspots (Myers et al. 2000).

Nonetheless, despite mandates for comparing impacts from multiple stressors, these efforts have largely focused on specific ecosystems, species, or issues. This focus limits their ability to inform the emerging demand for cross-ecosystem, cross-sector comparisons of ecosystem–stressor interactions that are necessary for ecosystem-based management (EBM; Spaling and Smit 1993, Council on Environmental Quality 1997, U.S. Environmental Protection Agency 1999, Crowder et al. 2006). Indeed, a key challenge for such efforts is that most marine ecosystems are subjected to many different human activities, making it difficult to disentangle the unique contribution and relative importance of each, especially when each ecosystem likely responds differently to the stressors associated with each activity (Halpern et al. 2007). What is needed, then, is a method for assessing vulnerability that is at the ecosystem scale and can directly compare across multiple stressors and multiple ecosystems.

Unfortunately, the methods and metrics to quantify ecosystem vulnerability to stressors that have been developed for a single issue, such as pollution, do not provide a means to compare levels of ecosystem vulnerability to stressors across a diversity of issues or ecosystem types. For example, ecotoxicology emerged as a field primarily in order to evaluate how water quality affects species and communities, yet these tools cannot be generalized to other issues. In marine systems, comparative evaluations have tended to focus on particular species (e.g., fish stocks, marine mammals, sea turtles) or issues (e.g., water quality, fishing, habitat loss) where a common currency, such as population size, toxin load, or habitat area can be used to quantify effects. There are notable exceptions, such as the recent rezoning process on the Great Barrier Reef Marine Park (Fernandes et al. 2005) and The Nature Conservancy and World Wildlife Fund's marine ecoregional assessment processes.

There is growing consensus that ecological vulnerability is a function of exposure, sensitivity, and resilience to stressors (Metzger et al. 2005, Millennium Ecosystem Assessment 2005). This shared framework creates the opportunity to inform and guide ecosystem-based management (EBM) through the integration of specific knowledge about vulnerability into an overall assessment of how human activities affect the marine ecosystems within a region. Ideally, assessments of overall vulnerability would be based on empirical data quantifying the expected impact of each stressor on each ecosystem. However, such data are available for only a fraction of the stressor–ecosystem combinations (Halpern et al. 2007).

Because of the lack of comprehensive empirical information on ecosystem–stressor interactions, expert interpretation and synthesis are needed to make existing research directly useful to management. The complexity of these processes is a common challenge in other fields such as engineering, sociology, and economics, where expert judgment is often used to predict failure in complex machines (e.g., nuclear power reactors) and understand societies that defy controlled experiments (Morgan and Henrion 1990, Morgan et al. 2000, Fischhoff et al. 2006, O'Hagan et al. 2006). Halpern et al. (2007) presented results from applying a recently developed method for eliciting expert judgments on the vulnerability of marine ecosystems to anthropogenic stressors. In a quantitative model, experts estimated values of five components of ecosystem vulnerability: spatial scale, frequency, trophic impact, percentage change (resistance), and recovery time. The five components, called vulnerability criteria, were based on metrics of exposure and sensitivity to stressors (Table 1). Because ecological vulnerability is a fairly abstract concept, structuring the collection of expert knowledge on vulnerability into these five more concrete factors creates more consistency and transparency to the use of expert opinion. The values for the five criteria are then combined to create a single score, which expresses the relative vulnerability of each ecosystem to each stressor. These quantitative vulnerability scores can be used to rank stressors or rank ecosystems to guide management or conservation priorities in mitigating human impacts. Rather than seeking group consensus, assessments are based on the distributions of expert judgments, allowing users to see the range of opinion.

The Halpern et al. (2007) method has two important shortcomings that are now addressed in this study. First, the five vulnerability criteria were weighted equally when combined into a single score. However, it is possible that experts are more concerned with one criterion, such as recovery time, than another, such as frequency of exposure, when judging what makes an ecosystem

TABLE 1. Description of the five vulnerability criteria used to evaluate ecosystem vulnerability to each stressor.

| Vulnerability criterion | Description |
|---|---|
| Spatial scale | The spatial scale ($km^2$) at which a single act of an activity impacts the ecosystem, both directly and indirectly. |
| Frequency | The mean annual frequency (days per year) of the activity at a particular location within a given region. |
| Trophic impact | The primary extent of marine life affected by an activity within a given ecosystem and region. |
| Percentage change | The degree to which the species, trophic level(s), or entire ecosystem's "natural" state is impacted by the activity. |
| Recovery time | The mean time (in years) required for the affected species, trophic level(s), or entire community to return to its former, "natural" state following disturbance by a particular activity. |

vulnerable to a stressor. Although equal weights often approximate more complex weighting schemes (Dawes 1979, Camerer 1981, Dawes et al. 1989), the stakes are high enough in marine resource management to assess weights empirically. Second, experts assessed their uncertainty with verbal quantifiers having no clear quantitative equivalent. Without a more explicit representation of uncertainty, policy makers cannot know how much faith to place in the judgments, and scientists cannot fairly evaluate their predictions (Morgan and Henrion 1990, O'Hagan et al. 2006, Fischhoff 2009).

Our work here takes advantage of a long history in the decision sciences of assessing how to set priorities (e.g., rank threats) when data are scarce and uncertainty exists by using the best available scientific judgments (e.g., Morgan and Henrion 1990, Fischhoff 2005, Willis et al. 2005). Basic research in human judgment has documented many ways in which unaided judgments (e.g., off-the-cuff assessments such as simply listing the rankings of threats) can produce results that do not stand up to more careful validation (Payne et al. 1992, Lichtenstein and Slovic 2006). One common approach to aid the elicitation of expert judgment is to use discrete choice tasks to help experts summarize their beliefs (Cooke and Goossens 2004). Here we apply one such approach based on ranking hypothetical scenarios of human impact to determine the relative importance of the five vulnerability criteria to expert judgment on how human activities in the California Current affect 19 different marine ecosystems. Variants on this approach have been used to assess risks in other complex uncertain situations where empirical knowledge is limited, such as complex engineered systems and disaster management (Cooke and Goossens 2004).

In this study we elicited judgments from scientific experts who study marine ecosystems within the California Current region in order to develop a deeper understanding of marine ecosystem vulnerability to a diversity of anthropogenic stressors. Even in this data-rich part of the world, critical and numerous gaps in empirical research remain, and our methods help fill these gaps in a low-cost, repeatable, and transparent manner. Our approach generates a matrix of relative vulnerability scores for every stressor-by-ecosystem combination that can be useful for management decisions and tools requiring detailed quantitative data about multiple human uses of the oceans at regional scales. Elsewhere, we use results from this study to inform a concurrent project mapping human activities across ecosystems in order to identify areas of particularly high or low cumulative impact (Halpern et al. 2009). In addition to informing management efforts dealing with these specific stressors within the California Current, we hope to demonstrate the utility of methods that can be applied elsewhere in the world.

## METHODS

Generating a matrix of vulnerability scores for all ecosystem–stressor combinations requires three components: (1) a comprehensive list of the relevant ecosystems and human stressors for the region, (2) estimated values for the five vulnerability criteria for every stressor–ecosystem combination, and (3) the weights determining how to combine the criteria values into a single score.

For the first component, ecosystem and stressor lists were based on a previous list (Halpern et al. 2007), and refined with input from two experts on California Current ecosystems (M. Beck of The Nature Conservancy and M. Ruckelshaus of NOAA Fisheries Service, *personal communications*). We included 19 distinct ecosystem types and 53 anthropogenic stressors. We acknowledge that there are additional historical stressors that humans have not imposed upon the system within the past five years, and some of these stressors have had a lasting effect on the system through to the present day. We ignored these historical stressors and focused only on present-day stressors, which the system has been exposed to within the past five years. To achieve the second and third components, we designed an expert survey focused on estimating the values for the five vulnerability criteria and eliciting their relative importance in judging vulnerability using a discrete choice approach (Appendix A). The five vulnerability criteria were developed previously (Halpern et al. 2007) in a workshop of conservation scientists and ecologists to represent whole-ecosystem vulnerability to a stressor (Table 1; see Appendix A, Part III, for more detailed criteria definitions). We use these same five criteria, but resistance is now defined as a continuous variable (i.e., percentage change). Criteria values take into account

TABLE 2. In Part III of the survey, each respondent received either a coastal or offshore version of the table below with 30 hypothetical scenarios (only a subset is presented here). All criteria values were identical between the coastal and offshore versions, but some scenario names were different.

| Coastal | Offshore | Spatial extent (km$^2$) | Frequency (no./yr)† | Trophic impact (level 1–4) | Percentage change | Recovery time (yr) |
|---|---|---|---|---|---|---|
| Aquaculture: marine plant | Aquaculture: finfish (predators) | 2 | 360 | 1 | 20 | 1 |
| Aquaculture: shellfish | Shipping: commercial, cruise, ferry | 5 | 360 | 1 | 5 | 0.1 |
| Climate change: sea level rise | | 10 000 | 1/2 | 2 | 20 | 5 |
| Climate change: sea temperature change | | 50 000 | 1 | 3 | 25 | 50 |
| Climate change: UV change | | 10 000 | 1/10 | 1 | 5 | 1 |
| Coastal engineering: habitat alteration | Ocean mining: sand, minerals, etc. | 1 | 1 | 4 | 75 | 25 |
| Direct human impact: trampling | Disease/pathogens | 150 | 360 | 2 | 35 | 25 |
| Fishing: demersal destructive | | 8 | 1/20 | 4 | 10 | 0.5 |
| Fishing: demersal nondestructive low bycatch | | 0.1 | 1/20 | 1 | 10 | 0.5 |
| Fishing: nondestructive artisanal‡ | Fishing: demersal nondestructive high bycatch | 1 | 1/20 | 1 | 50 | 1 |

*Notes:* Survey respondents were asked to rank the five hypothetical scenarios with the greatest impact based upon the vulnerability measures provided in the table. Respondents were also asked to not limit their ranking to a particular region and ecosystem, in contrast to the "stated ranks" activity, because the numbers represent a hypothetical coastal or offshore ecosystem in a hypothetical region.

† Fractions such as 1/2, 1/10, and 1/20 signify once every 2, 10, and 20 years.

‡ Artisanal refers to fishing practices such as traditional fishing on a small-scale, often subsistence or small-scale commercial fishing.

both direct effects (e.g., species mortality) and indirect ones (e.g., loss of nursery habitats). The mathematical basis for deriving the vulnerability model and the process of determining the criterion weights using a discrete choice survey are described in the next section.

### *Multi-criteria decision model*

The vulnerability model treats vulnerability as a weighted sum of the five criteria (Table 1) represented mathematically as

$$\text{Vulnerability}(\text{stressor } i, \text{ecosystem } j) = \sum_{k=1,\dots,5} W_k S^j_{i,k} \quad (1)$$

where $S^j_{i,k}$ is the value of stressor $i$ on criterion $k$ in ecosystem $j$, and $W_k$ is the weight assigned to criterion $k$, such that $W_{k\geq 0}$, $\Sigma_{k=1,\dots,5}\ W_k = 1$. The coefficients, or weights, are normalized so that they sum to unity. The weights are assumed to be the same for all ecosystems and stressors under consideration. This assumption allows for a single model to be applied to all ecosystem–stressor combinations, in turn allowing for direct comparison among them. While many mathematical models exist for combining the weights to create a single value (e.g., linear, logarithmic, polynomial), because environmental vulnerability is expected to be monotonic for all criteria (i.e., higher values denote greater impacts) it can be reasonably approximated by a simple linear model with positive coefficients.

To derive the relative weights $W_k$ of each vulnerability criterion we used a type of discrete choice task in which the expert is presented with a list of scenarios of anthropogenic stressors in a hypothetical region and ecosystem type (Table 2; see Part III of the survey instrument in Appendix A). Each scenario represents a different stressor, and hypothetical but realistic values for the five criteria are provided next to each scenario name. The expert must rank the top five scenarios that they judge to produce the largest negative human impact at the ecosystem level. The choice of five here is unrelated to the fact that there are five vulnerability factors; it is simply a large enough number to provide necessary data on the expert's decision-making process. Ranking the remaining scenarios is not only cognitively challenging but also unnecessary for the statistical analysis (Coombs 1964, Fischhoff 2005, Lichtenstein and Slovic 2006). The data on the expert's rankings are used in a statistical technique called "probabilistic inversion" (explained below) to derive the relative weights (summing to one) of the five criteria (Cooke and Goossens 2004, Fischhoff 2005, Du et al. 2006, Neslo 2008; see *Analyses* section below).

The derivation of the model weights uses a multi-criteria decision model (MCDM), a type of random utility model common to economic theory of utility. The MCDM treats the vulnerability criteria weights as random variables whose joint distribution is chosen to represent a population of experts, from which the elicited experts may be regarded as a random sample. Thus, the confidence intervals on the estimated weights reflect disagreement among the experts. To determine the joint distribution over the weights, we used probabilistic inversion, which inverts a mathematical model at a distribution or set of distributions and is analogous to maximum likelihood estimate methods.

Conceptually, the process returns values for the weights that reflect the importance of each weight in the expert's decision-making. For instance, if scenarios with large values for recovery time tend to be given high rankings, recovery time would get a large weight, and if trophic impact values show no relationship to the rankings, trophic impact would get a small weight.

Operationally and more accurately, probabilistic inversion finds a distribution for a function that maps onto the target distribution for the set of five vulnerability weights. Thus, given potential weights, we may define a function using Eq. 1 that says, in effect, "scenario 20 is ranked first, scenario 7 is ranked second," and so forth. Our expert elicitation data might in turn indicate, for example, "10% of the experts ranked scenario 20 first, 35% ranked scenario 7 second …." We search for a distribution over the weights that, when pushed through our function, realizes these probabilities. We assume that each expert's ranking is determined by Eq. 1 but with weights $W_k$ that are specific to that expert. The expert population is represented as a distribution over possible weight variables ($W_1, \ldots, W_5$). This distribution should be such that, continuing the above example, when sampled a large number of times, scenario 7 comes in second place 35% of the time. This correspondence must hold for all scenarios and all rank positions, from first ranked to fifth ranked. Scenarios with both low and high values for each criterion must be included to properly test their relationships to the rankings. Consequently, we chose criteria values for the 30 scenarios to capture the full range of possible combinations. The method used here to search for this distribution is based on the iterative proportional fitting algorithm, which finds a constrained maximum likelihood estimate of a joint distribution based on the sorts of constraints discussed previously (Csiszar 1975, Kurowicka and Cooke 2006). Analyses were conducted with a program scripted in C+ because no software currently exists for these analyses; one could use other programs separately for the MCDM and probabilistic inversion.

Random utility models allow for internal validation of the model, providing a more explicit and quantitative representation of uncertainty. Validation is based on (1) the number of inconsistencies, defined as cases where a hypothetical scenario (ecosystem–stressor combination) with lower scores on all five criteria is ranked as a greater stressor than a scenario with higher values on all five criteria, and (2) the ability of a model built with a subset of the expert discrete choice data to predict the remaining scenario rank orders.

### *The survey instrument*

In May 2007, a preliminary draft of the survey instrument was tested and revised based on input from a sample group of seven experts, none of whom participated in the final survey. The revised survey (see Appendix A) was then provided to respondents for completion by hand, phone, online, or in-person interview from June to October 2007. We asked experts to focus on one or more of six subregions, delineated to represent jurisdictional and biogeographic regions, and one or more of 19 marine ecosystem types (see Appendix A, Part I). The subregions are Washington, Oregon, northern California (San Francisco and north), central California (south of San Francisco to Point Conception), southern California (south of Point Conception), and Baja California, Mexico (north of Punta Eugenia). Respondents could expand or narrow their focal subregion(s) and ecosystem(s) in different parts of the survey.

The survey had four parts. In Part I, participants provided biographical information, such as professional affiliation(s) (academic, agency, non-governmental organization, or private company), age, and years of scientific experience within each ecosystem and within each geographic subregion. These data were used to test for possible drivers (i.e., bias) of expert judgment. In Part II, participants reviewed the list of 53 stressors, divided into 22 categories, and ranked the five stressors with the greatest negative impact on their chosen ecosystem and subregion. The 53 stressors were the same on every survey, but the order of the list was randomized by category to minimize potential order biases. Respondents could add or revise stressors. These "stated rankings" were obtained so that we could assess (1) whether we had captured all important stressors, and (2) whether the rankings would come out differently when simply stating them directly (i.e., unaided judgments), with no information on vulnerability criteria values and no statistical framework, in comparison to the process of deriving rankings with the MCDM. Part III elicited expert rankings for an individual's top five hypothetical scenarios where criteria values were supplied for example stressors (Table 2), providing the information necessary for the random utility model to derive the weights in Eq. 1. Labels such as "dredging" or "recreational fishing" were provided for the 30 scenarios, even though the values were hypothetical, to provide examples. In order to test the influence of the scenario names on the ranking process, we produced two versions of Part III, one for offshore ecosystems and one for coastal ecosystems, such that eight of the 30 scenarios had different labels but identical criteria values. We used these two expert groups (i.e., offshore and coastal) to compare if weighting values differed by system.

Part IV provided participants with default vulnerability criteria estimates for each stressor affecting their chosen ecosystem, based on values from a global survey (Halpern et al. 2007) or our own judgment (when a stressor was not in the global survey). Stressors not thought to exist or to have no relevant impact in that ecosystem were assigned zero for all vulnerability criteria. Participants then used their judgment to accept or revise each estimate, or indicate that they did not know what it should be. These data were used to estimate values for $S^j_{i,k}$ in the vulnerability model.

Table 3. Number of survey responders and nonresponders per affiliation and gender category.

| Category | Affiliation: Academic | Agency | NGO | Private | Gender: Male | Female | Total |
|---|---|---|---|---|---|---|---|
| Nonresponders | 120 | 94 | 49 | 0 | 192 | 71 | 263 |
| Responders | 56 | 33 | 16 | 2 | 80 | 27 | 107 |
| Total | 176 | 127 | 65 | 2 | 272 | 98 | 370 |

### *Survey respondent pool*

For inclusion in the potential respondent pool, we identified scientific experts with personal experience in marine science, conservation, management, or policy within the California Current and affiliated with academic institutions, governmental agencies, non-governmental organizations (NGO), or private environmental consulting firms (most scientific experts fall within these four affiliations). Potential respondents were identified via web-based searches using ecosystems, stressors, and locations as key words, based on our knowledge of the field and literature, and by requesting that respondents identify other experts possibly missing from our original list. Invitations were sent to 525 people, including 27 based in organizations located outside of the California Current (in Australia, Canada, mainland Mexico, and Baja California Sur, Panama, and Spain). One hundred fifty-five invitees self-identified themselves as non-experts (i.e., inappropriate or mistaken contacts), resulting in 370 potential expert respondents (see Table 3 for expert attributes). An additional 130 of these never responded so it is unclear whether they received the invitation or were truly appropriate experts, leaving a pool of 240 confirmed potential experts.

### *Analyses*

*Producing vulnerability criteria weights*.—Prior to all analyses, scale and frequency measures were transformed (i.e., scale = ln[scale × 100] and frequency = ln[frequency × 360]) to produce positive values on roughly equivalent scales as the other three criteria. This rescaling helps avoid a single criterion driving results simply because it has higher values from which to choose. For each hypothetical scenario in Part III, we calculated the percentage of experts who ranked the scenario first, second, third, fourth, and fifth, and then used probabilistic inversion to calculate the weights that best reproduced these observed percentages. Results were compared for model runs using the first ranking, the first two rankings, the first three rankings, and the first four rankings in order to evaluate if the number of ranks used affected the weighting values. We calculated these weights for all respondents ($N = 102$; five experts did not fill out this part of the survey), and for coastal ($N = 66$) and offshore ($N = 36$) versions of the survey to evaluate if system (coastal vs. offshore) affected weight values.

To test the validity of our multidimensional vulnerability model we first assessed the degree to which our five vulnerability criteria capture the factors experts use to rank vulnerability. To do this we compared the number of inconsistencies in scenario rankings (e.g., a case where a scenario with high values for all vulnerability criteria and one with low values are both ranked highly) with the number of inconsistencies generated by a null hypothesis that experts rank scenarios randomly without regard for the criteria values. If more inconsistencies emerge than would be expected at random, either experts used criteria beyond the five provided and/or experts did not understand or correctly execute the task. This method is one way to quantify the uncertainty in expert judgment using a measure of internal validity. We also used criteria weights from the model based on the first four ranked scenarios to predict experts' fifth-ranked scenario and compared these to actual fifth ranks from expert judgment to assess how well our model captured expert judgment.

*Ecosystem vulnerability scores*.—Vulnerability criteria values from Part IV were averaged across replicates (i.e., surveys completed by participants) for each ecosystem to estimate $S^j_{i,k}$ and combined with the weights $W_k$ in Eq. 1 to produce a vulnerability score for each stressor-by-ecosystem combination. We also calculated an overall average score for each stressor from the average scores for the 19 ecosystems and an average ecosystem vulnerability score from the scores for the 53 stressors for each ecosystem and used these averages to compare among subregions and between coastal and offshore ecosystems. We were unable to rigorously test whether ecosystem vulnerability scores differed by subregion because this test requires the sample size for an ecosystem to be large in all six subregions and in no case did this occur. However, sample size was large enough for 17 ecosystem–subregion comparisons (see Table 4 for specific pairwise comparisons) to allow for a partial test of subregional differences. To compare vulnerability scores between subregions for a given ecosystem we (1) averaged stressor vulnerability scores across respondents within a single subregion (instead of lumping subregions), (2) used two-tailed paired-sample $t$ tests to test for significant differences across subregions, and (3) used correlations to measure the strength of simple linear relationships between ranking values for subregions for the full set of stressors. Stressors were

TABLE 4. Subregional comparisons of four ecosystems (kelp forest, rocky intertidal, rocky reef, and seagrass) based on linear correlations of all stressor values and two-tailed paired-sample $t$ tests.

| Ecosystem and subregional comparison | $N$ | Correlation | $t$ | $P$ |
|---|---|---|---|---|
| Kelp forest | | | | |
| CCA vs. SCA | 49 | 0.60 | −1.05 | 0.30 |
| Rocky intertidal | | | | |
| WA vs. OR | 33 | 0.90 | 7.30 | <0.0001 |
| WA vs. CCA | 23 | 0.72 | 1.94 | 0.07 |
| WA vs. SCA | 45 | 0.87 | 5.02 | <0.0001 |
| WA vs. BCA | 27 | 0.86 | 5.57 | <0.0001 |
| OR vs. CCA | 20 | 0.75 | 1.39 | 0.18 |
| OR vs. SCA | 33 | 0.93 | −1.98 | 0.06 |
| OR vs. BCA | 23 | 0.95 | 0.03 | 0.98 |
| CCA vs. SCA | 24 | 0.75 | −1.78 | 0.09 |
| CCA vs. BCA | 14 | 0.62 | −0.93 | 0.37 |
| SCA vs. BCA | 29 | 0.89 | −1.19 | 0.25 |
| Rocky reef | | | | |
| OR vs. CCA | 8 | 0.79 | −2.75 | 0.03 |
| OR vs. SCA | 17 | 0.71 | −1.02 | 0.32 |
| CCA vs. SCA | 4 | 0.96 | −0.18 | 0.87 |
| Seagrass | | | | |
| NCA vs. SCA | 19 | 0.77 | 2.03 | 0.06 |
| NCA vs. BCA | 14 | 0.36 | −3.03 | 0.01 |
| SCA vs. BCA | 39 | 0.35 | −5.18 | <0.0001 |

*Notes:* Sample sizes ($N$), correlation coefficients, $t$ ratios, and $P$ values are shown. Abbreviations for subregions: WA (Washington), OR (Oregon), NCA (northern California), CCA (central California), SCA (southern California), and BCA (Baja California, Mexico).

excluded from individual surveys when one or more of the vulnerability criteria were not provided or when a subregion had only a single response for the stressor (e.g., some respondents did not fill in values for all 53 stressors).

*Potential respondent bias.*—We used chi-square tests to evaluate potential differences between responders and nonresponders based on gender or affiliation. Within the responder group, we examined possible differences in experts' assessment of criteria values (Part IV) based on demographic information collected in Part I, using ANOVA (for affiliation), $t$ test (for gender), and least-squares regression (for years of experience). For these tests we averaged all criteria values from all stressors, transformed, as described previously, for each respondent. Seven experts did not complete this section, resulting in a sample size of 95.

*Comparing directly stated and modeled ranks.*—We also compared experts' directly stated ranks, collected in Part II, to the ranks produced by the model using Spearman's rank correlation analysis. Because ecological vulnerability to stressors is a fairly abstract concept, we expected little consistency in top rankings from Part II across experts, and substantial deviation of these rankings from those generated by the statistical model, which breaks down the abstract concept into more concrete, specific subcomponents that are each quantified separately. To rank directly stated responses, we counted how often each stressor was among experts' top five ranks regardless of ecosystem. This method was chosen over a strict average rank because it is less sensitive to unusually high rankings. We used average ranks to break ties.

## RESULTS

### *Survey pool*

Out of the 240 confirmed potential expert respondents, 107 responded (45% response rate) by completing one or more surveys ($N = 160$ surveys). Respondents were from academic institutions (52%), government agencies (31%), NGOs (15%), and private consulting firms (2%), and included 80 males (75%) and 27 females (25%; Table 3). Thirty-nine respondents (36%) filled out more than one survey, 49 surveys addressed more than one subregion (mean = 1.6 ± 0.1 subregions; maximum = 6), and nine surveys addressed more than one ecosystem (mean = 1.1 ± 0.1 ecosystems; maximum = 9). One survey was eliminated due to unclear responses. The completed surveys covered 95 of the 120 possible ecosystem-by-subregion combinations, with 1–13 surveys per combination (see Appendix B). As no expert evaluated vents/seeps, this ecosystem was excluded from all analyses. Respondents' maximum reported years of experience within the marine ecosystems or subregions averaged 18.6 ± 1.1 years. On average, respondents had 14.4 ± 0.9 years of experience within their chosen ecosystem and 13.9 ± 0.6 years of experience within their chosen subregion (see Appendix C for average years of experience per ecosystem per subregion). Additionally, offshore ecosystems tended to be evaluated by fewer experts (3.2 ± 1.0 experts) than coastal ecosystems (7.2 ± 1.4 experts). Of the 263 potential respondents who did not complete a survey, 130 never responded (after at least three reminders), 115 did not respond after initially accepting the invitation (and after at least three reminders), 12 declined but gave no reason, and six declined because they did not feel comfortable filling out the survey.

### *Potential survey bias*

Affiliation and gender did not significantly differ between invited experts who completed the survey (responders) and those who did not (nonresponders: affiliation, $\chi^2 = 6.75$, df = 3, $P = 0.08$; gender, $\chi^2 = 0.121$, df = 1, $P = 0.728$; Table 3). The marginal significance for affiliation is due to the higher rate of response from academic experts. Vulnerability criteria values $S^j_{i,k}$ (Part IV) showed no significant differences associated with affiliation (ANOVA, $F_{3,93} = 0.36$, $P = 0.78$), gender ($t$ test, $t = 1.86$, $P = 0.07$), or years of experience (bivariate linear regression, $R^2 = 0$, $P = 0.88$). The marginally significant result for gender reflected a single extreme outlier; when removed, gender showed no trend ($t = -0.058$; $P = 0.95$).

TABLE 5. Mean weighting values for vulnerability criteria based on model results from the first stressor, and the first two, three, and four stressors ranked.

| Model | No. top ranks included | Spatial extent (km$^2$) | Frequency (no./yr) | Trophic impact (level 1–4) | Change (%) | Recovery time (yr) |
|---|---|---|---|---|---|---|
| Combined | 1 | 0.033 | 0.053 | 0.201 | 0.692 | 0.020 |
| | 2 | 0.067 | 0.089 | 0.299 | 0.532 | 0.014 |
| | 3 | 0.073 | 0.044 | 0.198 | 0.672 | 0.014 |
| | 4† | 0.061 ± 0.008 | 0.046 ± 0.007 | 0.221 ± 0.022 | 0.665 ± 0.029 | 0.008 ± 0.001 |
| Coastal | 4 | 0.072 | 0.046 | 0.226 | 0.650 | 0.006 |
| Offshore | 4 | 0.055 | 0.057 | 0.250 | 0.624 | 0.015 |

*Notes:* All surveys were included in the "combined" models; results of separate model runs for "offshore" or "coastal" ecosystem surveys are shown only for the first four stressors ranked. The second column gives the number of top stressor ranks used to calculate the model. Standard error (±SE) is given for the combined model using the first four stressors ranked.

† Model used for subsequent analyses.

### *Vulnerability criteria weights and model validation*

The MCDM produced highly uneven weightings for the five vulnerability criteria. Percentage change in biomass (resistance) and trophic impact together explained 89% of the overall weights of ecosystem vulnerability (66.5% and 22.1%, respectively). Recovery time had a small contribution to the overall vulnerability score (Table 5). These weights were highly consistent regardless of the number of ranks used to develop the model (one, two, three, or four). Using the top four ranks produced good predictions of the stressor ranked fifth by the experts; the frequency of each scenario being predicted to be fifth was highly similar to the frequency of experts selecting it as their fifth-ranked scenario (mean difference 0.001 ± 0.01 SE). Furthermore, there were significantly fewer inconsistencies than expected by random. Twenty-three of the 30 scenarios could produce inconsistencies (i.e., ranking one of these scenarios higher than the other seven would be "inconsistent"). Of the 102 experts only 15 chose one of these inconsistent scenarios as rank 1. The probability of observing so few inconsistencies is extraordinarily low ($7 \times 10^{-40}$), suggesting that experts generally understood the ranking task.

### *Vulnerability scores*

Vulnerability scores for all ecosystem–stressor combinations are provided in Table 6. Sample sizes for the criteria values used to produce these scores ranged from zero to 17. Across all values that experts assigned to the five criteria, 25.5% were marked zero (i.e., stressor was not a threat to the ecosystem), 21.0% marked "don't know," 12.9% left blank, and 0.3% marked "disagree" (i.e., experts disagreed but did not provide an alternate value). If a stressor–ecosystem combination had no expert responses, we used default criteria values from previous analyses (Halpern et al. 2007). There is a significant relationship between mean sample size and mean vulnerability score per ecosystem (linear regression: $R^2 = 0.30$, $P = 0.02$), suggesting that low response rates for some ecosystems may have resulted in lower vulnerability scores. However, this relationship has a low $R^2$ and is not significant when ecosystems with a mean sample size of less than four (Table 6) are excluded ($R^2 = 0.02$, $P = 0.76$). Stressors were evaluated by 88.9 ± 0.6 experts, on average.

Ocean acidification in soft slope, hard slope, and in hard deep ecosystems had the highest vulnerability score observed (3.4) and scores for this stressor exceeded 1.2 for all ecosystems (Table 6). On average, scores were greater in coastal than offshore ecosystems, most notably higher (>1.0 difference) for sea level rise, UV change, altered flow dynamics, habitat alteration, and invasive species. Only demersal destructive fishing was notably higher in offshore ecosystems. Coastal ecosystems were judged most vulnerable to (in decreasing order) invasive species, ocean acidification, sea temperature change, sea level rise, and habitat alteration from coastal engineering; while the stressors with the highest scores for offshore ecosystems were (in decreasing order) ocean acidification, demersal destructive fishing, shipwrecks, military activity, and lost fishing gear (Table 6). On average, coastal ecosystems were judged to have some degree of vulnerability (scores > 0.0) to nearly all of the 53 stressors evaluated here (43.7 ± 2.2 stressors), while offshore ecosystems were estimated to be vulnerable to less than half of the stressors (24.6 ± 3.7 stressors).

There were over 30 additional stressors that experts felt were not appropriately captured by our 53 stressors (see Appendix D). Some of these include alteration of tributaries and watersheds, altered oceanographic regimes (e.g., wind, circulation, or upwelling) due to climate change, global temperature change (not just sea temperature change), non-toxic algal blooms, illegal harvesting (poaching or harvesting by the public), kelp harvesting, wave energy development, and oil exploration and drilling (as distinguished from oil rigs and ocean mining).

### *Subregional comparisons*

Overall, the four middle subregions (Oregon and the three California subregions) had no significant differences in vulnerability scores for the ecosystems for which comparisons could be made, except for the central California and Oregon rocky reef ecosystem comparison (Table 4). For the one ecosystem for which comparisons

TABLE 6. Vulnerability scores for 53 stressors in 19 ecosystems.

| Stressors | Coastal ecosystems | | | | | | | | | Offshore ecosystems | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | KF | RR | SG | Shl | SR | BE | MF | RI | SM | SSh | SSl | SD | HSh |
| Aquaculture: finfish (herbivores) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Aquaculture: finfish (predators) | 0.2 | 1.0 | 0.3 | 0 | 0 | 0 | 0 | 0.2 | 0 | 0.9 | 0.5 | 0 | 0.7 |
| Aquaculture: marine plant | 0 | 0 | 0.4 | 0 | 0 | 0 | 0.3 | 0.8 | 0.4 | 0 | 0 | 0 | 0 |
| Aquaculture: shellfish | 0.4 | 0.5 | 1.6 | 0.5 | 1.5 | 0 | 1.1 | 1.0 | 0.9 | 0.2 | 0 | 0 | 0 |
| Benthic structures (e.g., oil rigs) | 1.6 | 1.7 | 1.6 | 1.4 | 2.0 | 1.4 | 2.4 | 0.9 | 1.8 | 2.2 | 1.4 | 0.4 | 2.4 |
| Climate change: ocean acidific. | 2.0 | 2.2 | 2.1 | 1.2 | 2.5 | 1.8 | 2.4 | 3.1 | 2.4 | 2.6 | 3.4 | 2.5 | 2.7 |
| Climate change: sea level rise | 1.8 | 1.5 | 1.9 | 0 | 2.2 | 1.7 | 1.9 | 2.7 | 2.5 | 0 | 0 | 0 | 0 |
| Climate change: sea temp. change | 2.9 | 2.2 | 1.9 | 0 | 2.2 | 1.7 | 1.8 | 2.7 | 1.8 | 1.7 | 0.6 | 0.5 | 1.9 |
| Climate change: UV change | 1.6 | 1.7 | 1.5 | 0 | 1.8 | 1.8 | 1.7 | 2.3 | 1.9 | 0 | 0 | 0 | 0 |
| Coastal engineer.: alt. flow dynam. | 1.2 | 0.7 | 1.1 | 0.6 | 2.4 | 1.3 | 2.0 | 1.5 | 2.5 | 0.2 | 0 | 0 | 0 |
| Coastal engineer.: habitat alteration | 1.4 | 1.1 | 1.6 | 0.6 | 2.4 | 1.3 | 2.1 | 1.7 | 2.7 | 0.2 | 0 | 0 | 0 |
| Direct human impact: trampling | 0.1 | 0.2 | 0.8 | 0.3 | 0 | 1.7 | 0.3 | 1.6 | 1.0 | 0.1 | 0 | 0 | 0 |
| Disease/pathogens | 1.0 | 1.0 | 0.9 | 0 | 1.5 | 1.1 | 1.1 | 1.1 | 0.8 | 0.9 | 0 | 0 | 1.1 |
| Dredging | 0.1 | 0.2 | 1.7 | 0.5 | 2.2 | 1.6 | 1.7 | 0.2 | 1.3 | 0.6 | 0 | 0 | 0 |
| Fishing: aquarium | 0.7 | 0.7 | 0.1 | 0 | 0 | 0 | 0.1 | 0.6 | 0.1 | 0.2 | 0 | 0 | 0 |
| Fishing: demersal destructive | 0.3 | 1.2 | 0.2 | 1.2 | 0 | 0.9 | 1.1 | 0.7 | 0.3 | 2.0 | 2.3 | 2.2 | 1.6 |
| Fishing: demrs. non-des. high byc. | 1.2 | 1.3 | 0.6 | 0.6 | 0.8 | 1.3 | 0.9 | 0.4 | 1.0 | 1.3 | 1.3 | 1.3 | 1.3 |
| Fishing: demrs. non-des. low byc. | 1.2 | 1.2 | 0.5 | 0.4 | 1.6 | 0.8 | 0.7 | 0.4 | 0.8 | 0.8 | 0.9 | 0.8 | 1.1 |
| Fishing: destructive artisanal | 0.2 | 0.1 | 0.2 | 0 | 1.5 | 0.8 | 0 | 0.7 | 0.8 | 0 | 0 | 0 | 0 |
| Fishing: non-destructive artisanal | 0.2 | 0.3 | 0.4 | 0.2 | 0.9 | 0.7 | 0.6 | 0.8 | 0.6 | 0.2 | 0 | 0 | 0 |
| Fishing: pelagic high bycatch | 0.3 | 0.9 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0.6 | 0.2 | 0 | 0.1 | 0 |
| Fishing: pelagic low bycatch | 0.2 | 0.8 | 0 | 0 | 0 | 0 | 0 | 0 | 0.4 | 0.3 | 0.3 | 0 | 0 |
| Fishing: recreational | 1.5 | 1.4 | 0.9 | 0.8 | 0.9 | 1.0 | 1.1 | 1.2 | 1.0 | 0.4 | 0 | 0.8 | 1.2 |
| Freshwater input: decrease | 0.2 | 0.1 | 0.6 | 0 | 0 | 0.6 | 1.1 | 0.8 | 1.5 | 0 | 0 | 0 | 0 |
| Freshwater input: increase | 0.8 | 1.0 | 0.8 | 0 | 0.8 | 0.9 | 1.2 | 1.1 | 1.4 | 0 | 0 | 0 | 0.8 |
| Invasive species | 2.4 | 1.8 | 1.7 | 1.3 | 2.1 | 3.2 | 3.0 | 2.6 | 2.0 | 0.7 | 0 | 1.1 | 1.5 |
| Marine component of forestry | 0 | 0 | 0.5 | 0 | 1.1 | 0.6 | 0.7 | 0.5 | 0.9 | 0.3 | 0 | 0 | 0 |
| Military activity | 1.0 | 0.3 | 0.8 | 0.6 | 0 | 1.4 | 0.8 | 0.4 | 0 | 1.3 | 1.2 | 1.2 | 1.3 |
| Nutrient input: causing HABs | 1.1 | 1.3 | 1.1 | 0.5 | 2.1 | 1.7 | 1.7 | 1.1 | 1.4 | 1.1 | 0 | 0 | 1.4 |
| Nutrient input: causing hyp. zones | 1.0 | 1.2 | 0.9 | 0.4 | 2.2 | 1.5 | 1.6 | 0.8 | 1.4 | 1.5 | 0 | 0 | 1.8 |
| Nutrient input: into eutrophic water | 0.9 | 1.0 | 0.9 | 0.3 | 2.2 | 1.5 | 1.0 | 0.9 | 0.9 | 1.0 | 0 | 0 | 1.2 |
| Nutrient input: into meso. water | 1.2 | 0.9 | 0.9 | 0.3 | 2.2 | 1.1 | 1.2 | 1.0 | 1.5 | 0.8 | 0 | 0 | 1.3 |
| Ocean dumping: lost fishing gear | 1.1 | 1.1 | 0.9 | 0.4 | 0.5 | 1.5 | 1.4 | 1.0 | 1.2 | 1.2 | 1.3 | 1.2 | 1.2 |
| Ocean dumping: marine debris | 0.8 | 0.9 | 0.6 | 0.4 | 0.6 | 1.0 | 1.0 | 0.9 | 1.0 | 0.8 | 0.9 | 0.8 | 1.0 |
| Ocean dumping: ship wrecks | 1.7 | 2.1 | 1.5 | 1.0 | 2.3 | 1.6 | 2.3 | 1.5 | 0 | 1.8 | 1.6 | 1.3 | 2.4 |
| Ocean dumping: toxic materials | 0.9 | 1.0 | 1.1 | 0.5 | 1.7 | 1.4 | 1.5 | 0.9 | 1.4 | 1.1 | 1.1 | 1.1 | 1.1 |
| Ocean mining (sand, minerals, etc.) | 0.1 | 0 | 0.5 | 0 | 0 | 1.1 | 0 | 0.1 | 0 | 1.3 | 1.6 | 0 | 0 |
| Ocean pollution (from ships/ports) | 0.9 | 1.0 | 0.9 | 0.4 | 0 | 0.8 | 1.3 | 1.3 | 1.1 | 0.8 | 0 | 0 | 1.0 |
| Pollution input: atmospheric | 1.0 | 1.1 | 1.0 | 0 | 1.3 | 0.9 | 1.3 | 1.1 | 1.4 | 1.1 | 0 | 0 | 1.2 |
| Pollution input: inorganic | 1.3 | 1.5 | 0.9 | 0.8 | 1.4 | 1.3 | 1.3 | 1.2 | 1.4 | 1.6 | 2.0 | 1.9 | 1.1 |
| Pollution input: light/noise | 0.1 | 0 | 0.6 | 0.3 | 1.3 | 1.2 | 1.2 | 0.9 | 1.2 | 0.2 | 0 | 0 | 0 |
| Pollution input: organic | 1.3 | 1.5 | 1.1 | 1.4 | 1.5 | 1.9 | 1.9 | 1.3 | 1.4 | 1.5 | 2.0 | 2.5 | 1.3 |
| Pollution input: trash, urban runoff | 0.7 | 1.2 | 0.9 | 1.0 | 1.0 | 1.2 | 1.2 | 1.1 | 1.1 | 0.9 | 1.2 | 1.9 | 0.4 |
| Power, desalination plants | 0.8 | 1.1 | 0.5 | 0.5 | 0 | 1.7 | 1.3 | 1.5 | 1.2 | 0.2 | 0 | 0 | 0 |
| Scientific research: collecting | 0.7 | 0.7 | 0.6 | 0.1 | 0 | 0.7 | 0.7 | 0.8 | 0.9 | 0.9 | 1.3 | 0.7 | 0.7 |
| Scientific research: expts./surveys | 0.7 | 0.8 | 0.8 | 0.8 | 0 | 0.7 | 0.7 | 0.9 | 0.8 | 1.0 | 1.3 | 0.8 | 0.8 |
| Sediment input: decrease | 0.1 | 0 | 0.8 | 0.6 | 1.6 | 1.2 | 2.3 | 0.8 | 2.4 | 0.3 | 0 | 0.4 | 0 |
| Sediment input: increase | 1.4 | 1.1 | 1.3 | 0.8 | 1.4 | 1.8 | 1.4 | 1.5 | 1.4 | 0 | 0 | 1.2 | 0 |
| Shipping (commercial, cruise, etc.) | 0 | 0.3 | 0.3 | 0 | 0 | 1.4 | 0.4 | 0.2 | 0 | 0.3 | 0 | 0 | 0 |
| Tourism: kayaking | 0.6 | 0.5 | 0.1 | 0 | 0 | 0.4 | 0.4 | 0.4 | 0.2 | 0.1 | 0 | 0 | 0 |
| Tourism: recreational boating | 0.9 | 1.0 | 1.0 | 0 | 0 | 0.2 | 0.8 | 0 | 0.2 | 0.2 | 0 | 0 | 0.4 |
| Tourism: scuba diving | 1.0 | 0.9 | 0.1 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 | 0 | 0 | 0.2 |
| Tourism: surfing | 0 | 0 | 0.1 | 0 | 0 | 0.4 | 0 | 0.6 | 0 | 0 | 0 | 0 | 0 |
| Score mean | 0.8 | 0.9 | 0.8 | 0.4 | 1.0 | 1.1 | 1.1 | 1.0 | 1.0 | 0.7 | 0.5 | 0.5 | 0.7 |
| Score SE | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| Sample size mean | 11.5 | 9.3 | 9.5 | 3.8 | 0.2 | 3.8 | 5.2 | 13.1 | 8.0 | 6.1 | 1.7 | 2.1 | 3.1 |
| Sample size SE | 0.1 | 0.1 | 0.1 | 0.1 | 0.0 | 0.1 | 0.1 | 0.2 | 0.1 | 0.1 | 0.0 | 0.0 | 0.1 |

*Notes:* Mean scores for each stressor across all ecosystems, and for each ecosystem across all stressors, are reported in the "score mean" column and row, respectively. Ecosystem abbreviations: KF (kelp forest), RR (rocky reef), SG (seagrass), Shl (shallow soft), SR (suspension-feeding reefs), BE (beach), MF (mud flats), RI (rocky intertidal), SM (salt marsh), SSh (soft shelf), SSl (soft slope), SD (soft deep), HSh (hard shelf), HSl (hard slope), HD (hard deep), Cyn (canyons), SMt (seamounts), Surf (surface waters), Deep (deep waters). Score identities are based on four equal divisions of the range of values: 0.0–0.8 (normal black font), 0.9–1.7 (light blue), 1.8–2.6 (purple), and 2.7–3.4 (red). Sample size means (±SE) were calculated based on the average sample size per vulnerability criteria in Part IV (i.e., when a respondent evaluated a stressor, all vulnerability criteria values were not always provided). Stressor abbreviations: acidific. (acidification), alt. (altered), byc. (bycatch), demrs. (demersal), des. (destructive), dynam. (dynamics), engineer. (engineering), expts. (experiments), HABs (harmful algal blooms), hyp. (hypoxic), meso. (mesotrophic), temp. (temperature).

TABLE 6. Extended.

| Offshore ecosystems | | | | | | |
|---|---|---|---|---|---|---|
| HSl | HD | Cyn | SMt | Surf | Deep | Score mean |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1.2 | 0 | 0.3 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.1 |
| 0 | 0 | 0 | 0 | 0.2 | 0 | 0.4 |
| 2.3 | 0 | 2.3 | 0 | 0.4 | 0 | 1.4 |
| 2.3 | 3.4 | 2.6 | 2.6 | 3.2 | 2.7 | 2.6 |
| 0 | 0 | 0 | 0 | 0 | 0.6 | 0.9 |
| 1.2 | 0 | 1.7 | 0 | 2.5 | 1.9 | 1.5 |
| 0 | 0 | 0 | 0 | 2.5 | 0.8 | 0.9 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.7 |
| 0 | 0 | 0 | 0 | 0.2 | 0 | 0.8 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.3 |
| 1.2 | 0 | 1.1 | 0 | 0.5 | 0 | 0.7 |
| 0 | 0 | 0 | 0 | 0.1 | 0 | 0.5 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.1 |
| 2.3 | 2.8 | 2.5 | 2.7 | 0.3 | 0 | 1.3 |
| 1.8 | 1.3 | 1.3 | 1.3 | 0.4 | 0 | 1.0 |
| 1.1 | 0.9 | 0.8 | 0.9 | 0.3 | 0 | 0.8 |
| 0 | 0 | 0 | 0 | 0.1 | 0 | 0.2 |
| 0 | 0 | 0 | 0 | 0.1 | 0 | 0.3 |
| 0 | 0 | 0 | 0 | 1.6 | 1.6 | 0.3 |
| 0 | 0 | 0 | 0 | 1.1 | 1.5 | 0.2 |
| 1.6 | 0.7 | 0 | 0.7 | 1.1 | 0 | 0.9 |
| 0 | 0 | 0 | 0 | 0.1 | 0 | 0.3 |
| 0 | 0 | 0 | 0 | 1.0 | 0 | 0.5 |
| 0 | 0 | 0 | 0 | 0.3 | 0 | 1.2 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.2 |
| 1.3 | 1.3 | 1.3 | 1.2 | 1.4 | 1.4 | 1.0 |
| 1.3 | 0 | 0 | 0 | 1.5 | 1.8 | 1.0 |
| 1.0 | 0 | 0 | 0 | 1.8 | 2.0 | 1.0 |
| 0 | 0 | 1.0 | 0 | 1.1 | 0.9 | 0.8 |
| 0 | 0 | 1.3 | 0 | 1.4 | 2.1 | 0.9 |
| 1.2 | 1.3 | 1.2 | 1.3 | 1.3 | 1.3 | 1.1 |
| 1.0 | 0.9 | 0.6 | 0.7 | 1.0 | 0.8 | 0.8 |
| 2.5 | 2.5 | 2.3 | 2.3 | 0 | 0 | 1.6 |
| 1.4 | 1.1 | 1.3 | 1.1 | 1.1 | 1.3 | 1.2 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.2 |
| 1.3 | 0 | 0.9 | 0 | 1.4 | 0 | 0.7 |
| 0.7 | 0 | 1.3 | 0 | 1.6 | 0.9 | 0.8 |
| 0 | 0 | 1.3 | 0 | 1.6 | 1.3 | 1.2 |
| 0 | 0 | 0 | 0 | 0.4 | 0.8 | 0.4 |
| 0 | 0 | 1.2 | 0 | 1.5 | 1.3 | 1.3 |
| 0.9 | 0 | 0.8 | 0 | 1.1 | 0.8 | 0.9 |
| 0 | 0 | 0 | 0 | 1.5 | 0.2 | 0.6 |
| 0 | 0.7 | 0.7 | 0.7 | 0.7 | 0.4 | 0.6 |
| 0 | 0.9 | 0.9 | 0.9 | 1.0 | 0.5 | 0.8 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.6 |
| 1.0 | 0 | 1.4 | 0 | 0.2 | 0 | 0.8 |
| 0 | 0 | 0 | 0 | 1.5 | 0 | 0.2 |
| 0 | 0 | 0 | 0 | 0.9 | 0 | 0.2 |
| 0 | 0 | 0 | 0 | 1.1 | 0 | 0.3 |
| 0 | 0 | 0 | 0 | 0.2 | 0 | 0.1 |
| 0 | 0 | 0 | 0 | 0.2 | 0 | 0.1 |
| 0.5 | 0.3 | 0.6 | 0.3 | 0.8 | 0.5 | 0.7 |
| 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.07 |
| 2.2 | 1.0 | 0.7 | 1.9 | 11.6 | 1.8 | 19.0 |
| 0.1 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.00 |

could be made to Washington (rocky intertidal), the Washington subregion differed significantly from all other subregions except central California, but had highly correlated values in all of these cases ($R^2 > 0.72$). Baja California was significantly different from northern and southern California in vulnerability scores for seagrass ecosystems (as well as Washington in rocky intertidal ecosystems), and also had low correlation values. Vulnerability scores for rocky intertidal ecosystems in Baja did not differ significantly from central and southern California.

### *Comparing directly stated and modeled ranks*

Spearman's rank correlation between directly stated ranks (Part II) and model-derived mean vulnerability scores (mean values across ecosystems from Part IV) is significant ($P = 0.001$) but relatively low ($\rho = 0.44$). The five most commonly directly stated top five stressors were sea temperature change (42% of respondents ranked it in their top five), recreational fishing (33%), habitat alteration from coastal engineering (32%), increasing sediment loads (22%), and invasive species (22%); yet of these, only sea temperature change and invasive species were among the top five modeled vulnerability ranks (Table 7), and recreational fishing and sediment increase were not among the top 10 modeled ranks. Ocean acidification received the highest modeled vulnerability score, yet was included in only 11% of respondents' stated top five stressors. Additionally, UV change, sea level rise, benthic structures, shipwrecks, and hypoxic zones caused by nutrient input were only included in $\leq$7% of stated top five stressors, yet all appeared among the top 10 modeled ranks. Remarkably, all but three of the 53 stressors were ranked by at least one expert in their top-five stated stressors (across all ecosystems).

## DISCUSSION

### *Decision theory approach to assessing ecosystem vulnerability*

Our approach moves beyond previous methods for assessing environmental risk in several key ways. The decision rules (criteria) and relative importance of those criteria (weights) are explicit and quantified, rather than implicit and qualitative as is the case for most Delphi processes. The structured approach to assessing these criteria and weights compels experts to take an ecosystem-level perspective when evaluating the importance of stressors rather than, for example, focusing only on the species they study and to explicitly consider (and quantify) exposure and sensitivity aspects of vulnerability. Vulnerability is an abstract concept and defining it at an ecosystem-level scale adds further complexity to the concept. This complexity challenges an individual's cognitive ability to compare the vulnerability of ecosystems to various stressors in a consistent and fair manner without the aid of a model built from concrete subcomponents. Indeed, experts' directly stated top stressors showed little correlation with the modeled top stressors. When experts simply list key stressors, there is no way to know why they chose those stressors, with responses potentially subject to biases that cannot be tested (Payne et al. 1992, Lichtenstein and Slovic 2006). Using a mathematical model, however, requires

TABLE 7. Scores and rank orders for directly stated top stressors (Part II) and the multi-criteria decision model (MCDM) (based on Parts III and IV).

| | Directly stated | | MCDM for all ecosystems | | Coastal MCDM | | Offshore MCDM | |
|---|---|---|---|---|---|---|---|---|
| Stressor | Score | Rank order | Score | Rank order | Score | Rank order | Score | Rank order |
| Aquaculture: finfish (herbivores) | 0.01 | 40 | 0.00 | 53 | 0.00 | 53 | 0.00 | 52 |
| Aquaculture: finfish (predators) | 0.01 | 42 | 0.27 | 40 | 0.19 | 50 | 0.33 | 28 |
| Aquaculture: marine plant | 0.00 | 51 | 0.10 | 51 | 0.21 | 48 | 0.0 | 53 |
| Aquaculture: shellfish | 0.10 | 19 | 0.41 | 36 | 0.83 | 30 | 0.04 | 42 |
| Benthic structures: e.g., oil rigs | 0.03 | 33 | 1.37 | 4 | 1.64 | 6 | 1.14 | 8 |
| Climate change: ocean acidification | 0.11 | 18 | 2.57 | 1 | 2.19 | 2 | 2.91 | 1 |
| Climate change: sea level rise | 0.06 | 25 | 0.88 | 18 | 1.80 | 4 | 0.06 | 40 |
| Climate change: sea temperature change | 0.42 | 1 | 1.54 | 3 | 1.91 | 3 | 1.20 | 6 |
| Climate change: UV change | 0.01 | 43 | 0.93 | 15 | 1.59 | 7 | 0.33 | 29 |
| Coastal engineering: altered flow dynamics | 0.19 | 11 | 0.71 | 27 | 1.48 | 9 | 0.02 | 46 |
| Coastal engineering: habitat alteration | 0.32 | 3 | 0.81 | 23 | 1.66 | 5 | 0.04 | 43 |
| Direct human impact: trampling | 0.15 | 13 | 0.33 | 37 | 0.67 | 34 | 0.01 | 49 |
| Disease/pathogens | 0.12 | 17 | 0.70 | 28 | 0.94 | 25 | 0.48 | 24 |
| Dredging | 0.08 | 22 | 0.54 | 33 | 1.06 | 20 | 0.07 | 38 |
| Fishing: aquarium | 0.01 | 41 | 0.13 | 50 | 0.26 | 45 | 0.02 | 47 |
| Fishing: demersal destructive | 0.20 | 6 | 1.30 | 5 | 0.66 | 35 | 1.87 | 2 |
| Fishing: demersal nondestructive high bycatch | 0.19 | 10 | 1.01 | 11 | 0.90 | 26 | 1.13 | 9 |
| Fishing: demersal nondestructive low bycatch | 0.20 | 7 | 0.80 | 24 | 0.84 | 29 | 0.76 | 16 |
| Fishing: destructive artisanal | 0.04 | 29 | 0.24 | 46 | 0.48 | 41 | 0.01 | 50 |
| Fishing: nondestructive artisanal | 0.05 | 26 | 0.27 | 41 | 0.52 | 39 | 0.03 | 44 |
| Fishing: pelagic high bycatch | 0.12 | 15 | 0.29 | 39 | 0.22 | 46 | 0.35 | 27 |
| Fishing: pelagic low bycatch | 0.07 | 23 | 0.25 | 43 | 0.16 | 51 | 0.32 | 30 |
| Fishing: recreational | 0.33 | 2 | 0.85 | 19 | 1.09 | 17 | 0.65 | 21 |
| Freshwater input: decrease | 0.01 | 46 | 0.26 | 42 | 0.54 | 38 | 0.01 | 51 |
| Freshwater input: increase | 0.02 | 36 | 0.51 | 34 | 0.89 | 27 | 0.18 | 33 |
| Invasive species: from ballast, etc. | 0.22 | 5 | 1.25 | 7 | 2.23 | 1 | 0.36 | 26 |
| Marine component of forestry operations (log booms) | 0.03 | 34 | 0.25 | 45 | 0.48 | 40 | 0.03 | 45 |
| Military activity | 0.01 | 50 | 0.96 | 14 | 0.59 | 36 | 1.29 | 4 |
| Nutrient input: causing harmful algal blooms | 0.14 | 14 | 1.00 | 12 | 1.33 | 12 | 0.71 | 17 |
| Nutrient input: causing hypoxic zones | 0.07 | 24 | 1.00 | 13 | 1.22 | 14 | 0.81 | 14 |
| Nutrient input: into eutrophic (upwelled) waters | 0.09 | 21 | 0.78 | 25 | 1.07 | 19 | 0.52 | 23 |
| Nutrient input: into mesotrophic (non-upwelled) waters | 0.05 | 27 | 0.91 | 17 | 1.14 | 16 | 0.69 | 18 |
| Ocean dumping: lost fishing gear | 0.04 | 32 | 1.13 | 10 | 1.01 | 23 | 1.25 | 5 |
| Ocean dumping: marine debris, trash, etc. | 0.04 | 31 | 0.82 | 22 | 0.80 | 31 | 0.85 | 12 |
| Ocean dumping: ship wrecks | 0.00 | 52 | 1.61 | 2 | 1.56 | 8 | 1.67 | 3 |
| Ocean dumping: toxic materials | 0.02 | 37 | 1.15 | 8 | 1.16 | 15 | 1.17 | 7 |
| Ocean mining: sand, minerals, etc. | 0.00 | 53 | 0.25 | 44 | 0.20 | 49 | 0.29 | 31 |
| Ocean pollution: chemicals from ships, ports, spills | 0.16 | 12 | 0.69 | 29 | 0.86 | 28 | 0.54 | 22 |
| Pollution input: atmospheric | 0.01 | 44 | 0.84 | 20 | 1.01 | 22 | 0.68 | 20 |
| Pollution input: inorganic | 0.12 | 16 | 1.15 | 9 | 1.23 | 13 | 1.08 | 11 |
| Pollution input: light/noise | 0.02 | 39 | 0.43 | 35 | 0.76 | 32 | 0.14 | 36 |
| Pollution input: organic | 0.20 | 9 | 1.29 | 6 | 1.48 | 10 | 1.13 | 10 |
| Pollution input: trash, etc. (i.e., urban runoff) | 0.20 | 8 | 0.91 | 16 | 1.04 | 21 | 0.80 | 15 |
| Power, desalination plants | 0.05 | 28 | 0.55 | 32 | 0.96 | 24 | 0.19 | 32 |
| Scientific research: collecting | 0.01 | 47 | 0.63 | 30 | 0.58 | 37 | 0.68 | 19 |
| Scientific research: experiments/surveys | 0.02 | 38 | 0.75 | 26 | 0.69 | 33 | 0.81 | 13 |
| Sediment input: decrease | 0.09 | 20 | 0.56 | 31 | 1.09 | 18 | 0.07 | 39 |
| Sediment input: increase | 0.22 | 4 | 0.83 | 21 | 1.34 | 11 | 0.38 | 25 |
| Shipping: commercial, cruise, ferry | 0.01 | 45 | 0.22 | 47 | 0.29 | 43 | 0.18 | 34 |
| Tourism: kayaking | 0.01 | 48 | 0.20 | 48 | 0.29 | 44 | 0.10 | 37 |
| Tourism: recreational boating | 0.04 | 30 | 0.30 | 38 | 0.46 | 42 | 0.17 | 35 |
| Tourism: scuba diving | 0.03 | 35 | 0.13 | 49 | 0.22 | 47 | 0.05 | 41 |
| Tourism: surfing | 0.01 | 49 | 0.07 | 52 | 0.12 | 52 | 0.02 | 48 |

*Notes:* Directly stated scores were calculated based on the frequency of each stressor occurring among the top five stated stressors across all ecosystems; average ranks were used to break ties. The multi-criteria decision model score was calculated from Eq. 1 for each ecosystem and then averaged across all ecosystems, and for coastal and offshore ecosystems separately.

knowing which subcomponents to use in building the model and how to combine subcomponents in a way that matches an expert's decision-making process. The subcomponents (i.e., vulnerability criteria) come from a long history of research on the topic; the multi-criteria decision model (MCDM) fills the latter role of combining subcomponents. The MCDM revealed that experts primarily used percentage change (i.e., resistance) and trophic impact when evaluating ecosystem vulnerability to stressors, despite that vulnerability is thought to also be a function of exposure, not just measures of sensitivity (Metzger et al. 2005, Millennium Ecosystem Assessment 2005).

The MCDM also allowed us to test how consistently experts used the vulnerability criteria in their assessments (i.e., internal model validity) by comparing results from two versions of the survey based on different systems (coastal vs. offshore), calculating model weights using different numbers of scenario rankings, and using the model to predict the next-ranked stressor. These comparisons do not allow us to test the uncertainty of individual experts but do provide several methods for testing and quantifying variability (i.e., uncertainty) among experts, a key improvement over our previous approach (Halpern et al. 2007). We found high model validity in all cases: model weights were consistent between systems and with different numbers of ranks used to build the model, and the ecosystem vulnerability model predicted well the next-ranked stressor. This ability to evaluate model validity is rare among methods for eliciting expert judgment. The robustness of the model suggests that the vulnerability model can be used with the same values for the criteria weights to evaluate new stressors and ecosystems not included here. Thus, the model provides a rapid way to assess additional and emerging ocean uses, such as wind and wave farms or liquefied natural gas (LNG) terminals, and quickly "slide" them into the appropriate rank order once their vulnerability scores are estimated.

The vulnerability model solves the "apples to oranges" problem of making comparisons between very different types of systems, and the use of expert judgment allows the filling of knowledge gaps where empirical data do not exist. The vulnerability model, in particular, differentiates our work from other efforts to rank stressors to ecosystems (Olson and Dinerstein 1998, Myers et al. 2000, Pew Oceans Commission 2003, Metzger et al. 2005, Millennium Ecosystem Assessment 2005). In this structured framework, judgment is a means to efficiently incorporate knowledge and understand the world. Thus, the approach presented here benefits from greater transparency and repeatability than most other expert judgment elicitation procedures.

*Ecosystem vulnerability in the California Current*

Ecosystem-based approaches to resource management require knowledge of how each ecosystem responds to the stressors associated with human uses of the ocean, yet empirical information on such responses is limited. Using a decision theory method for eliciting expert judgments, we have evaluated the vulnerability of 19 marine ecosystems within the California Current region to 53 different stressors, a total of 1007 stressor-by-ecosystem combinations. There are both expected and unanticipated aspects to the vulnerability assessments for the California Current. Averaged across all ecosystems, stressors with high vulnerability scores were associated with climate change, invasive species, habitat destruction (benthic structures, coastal engineering), and pollution; all of which have been previously highlighted as key issues (Vitousek et al. 1997). Coastal ecosystems were assessed to be more vulnerable to human stressors, and to a higher number of stressors, than offshore systems. At the ecosystem level, rank order of stressors by vulnerability scores varies greatly with ecosystem type, as is expected. More unexpectedly, ocean acidification topped the rankings for many ecosystems. This result highlights the urgent need to develop strategies for addressing this climate stressor. However, very few experts listed ocean acidification, UV change, and sea level rise among their top five stated ranks, yet these all fell within the top 10 modeled ranks, indicating the need for greater awareness of these climate stressors. On the other hand, experts do seem to be aware of the importance of sea temperature change and invasive species, as these stressors ranked high for both ranking methods. Modeled results ranked commercial fishing as a top stressor in most offshore ecosystems (Tables 6 and 7), as has been found by many others (e.g., Pauly et al. 1998, Myers and Worm 2003, Worm et al. 2006); but across all ecosystems, the five types of commercial fishing showed lower vulnerability scores than many other stressors. This is because experts judged pelagic fishing to have very little or no impact on many ecosystems and land-based sources of stress to have larger impacts on a suite of coastal ecosystems. Fishing may have ranked lower as well because our approach focuses on present-day stressors and therefore ignores the historical, accumulated stress of fishing (in particular overfishing) on ecosystems.

Although many of these results on the top stressors or most vulnerable systems may seem expected or known, it is extremely valuable to test those assumptions with a rigorous scientific approach and to assess the level of consensus on rankings among experts. Results from a rigorous survey provide strong rationale and justification for management decisions even when top stressors are believed to be already well-known; the value of this supporting role should not be underestimated given the politically charged environment in which these decisions are often made. The combined input of a large number of experts should carry considerably more weight than that of one or few managers or scientific advisors asserting their beliefs about top stressors.

Although these relative stressor rankings are valuable for aiding conservation and management prioritization efforts, another useful result is the matrix of quantitative vulnerability scores that is produced (Table 6). These scores not only give a quantitative, relative estimate of vulnerability of an ecosystem to each stressor (e.g., kelp forests are judged to be five times as vulnerable to ocean acidification as they are to shellfish aquaculture), but also allow direct and quantitative comparisons of stressor vulnerability among ecosystem types (e.g., rocky reefs are judged to be 30% more vulnerable to recreational fishing than seagrass beds are to organic pollution). This ability to compare very different entities in a quantitative manner has broad potential application and relevance to various cost–benefit analyses of how and where to prioritize management, mitigation, and conservation effort.

A key challenge for any effort to evaluate cross-ecosystem, cross-sector vulnerability is to decide how much to lump or split categories of stressors and ecosystems. Fishing can be considered as a single stressor, as five categories of stress (as we have done here), or as many categories in which each species and gear type is evaluated separately. Similarly, habitats can be classified according to any number of physical and biological attributes (e.g., sediment grain size or type, tidal flux, depth, relief, wave exposure, upwelling characteristics, temperature, salinity, species composition, and diversity; Carlton 2007), which can lead to few or many habitat types depending on these decisions. For example, one could choose to lump all salt marshes together as a single ecosystem type or split them into estuarine and coastal salt marshes. Here we strove to focus on a level of habitat classification that was general enough to likely be addressed by management efforts in the California Current but fine enough to capture important differences, and a level of stressor classification that captures important differences in potential impact to ecosystems from subdivisions of a stressor class but is general enough to match typical management focus. Additionally, we have assumed that experts take into account the temporal dynamics of oceanographic and climatic processes (e.g., El Niño Southern Oscillation cycle, the Metonic cycle, the Pacific decadal oscillation; Halpin et al. 2004) when assessing the influence of a particular stressor on an ecosystem. However, our survey focused on assessing the present day (within the past five years), so longer temporal dynamics could be the focus of future studies. In summary, our method for assessing ecosystem vulnerability can easily be adapted to assess a different classification scheme, spatial scale/extent, or time period, and directly compared to our output here.

Our assessment of the differential vulnerability of ecosystems does not account for potential synergistic effects among stressors, where some combinations of stressors may lead to greater impacts than our estimates here, resulting in higher scores. These synergisms are currently poorly understood (Crain et al. 2008, Darling and Cote 2008), so it is difficult to account for them in the vulnerability model. Also, the default vulnerability criteria values provided in Part IV may have influenced experts, or experts may have been reluctant to modify defaulted values unless they felt them to be radically wrong. An alternative would have been to leave these values blank, but experts tend to skip blank values (Halpern et al. 2007).

Ultimately, the accuracy of the vulnerability scores depends on the quality of expert judgment. We were careful to include only experts with empirical knowledge and experience in marine ecosystems within the California Current, but we recognize that this does not ensure accuracy. Future studies could incorporate additional assessments; for example, including experts with backgrounds in each of the stressors. In addition, carefully controlled experiments that clearly show the relative vulnerabilities of ecosystems to different stressors are the gold standard for environmental risk assessment, but the day is far off when such data exist for the numerous ecosystem–stressor combinations. Until then, expert judgment elicitation can provide some guidance to management efforts.

### *Management implications*

Our approach and results can be used in a number of ways to inform and aid management efforts and particularly address the fundamental question of how and where to prioritize stressor and ecosystem management. Our results alone cannot answer that question, as there are many dimensions (socioeconomic, opportunities, etc.) that drive such decisions, but our quantitative vulnerability scores can provide a key piece of the answer. The matrix of vulnerability scores based on expert judgment informs which stressors are likely most important to address, which ecosystems are likely most vulnerable, and which factors (i.e., criteria) likely drive that vulnerability. Even if these results are believed to be known, having a quantitative and transparent method for assessing vulnerability is of enormous value to anyone or any organization that must explain and defend their management decisions.

Our analyses provide results most appropriate for state-level and federal-level management or conservation organizations focused on large biogeographic regions or the California Current as a planning unit. At this scale, the high vulnerability scores of most ecosystems for climate change stressors point to the immediate need for local, state, federal, and international action to address this key stressor for nearly all ecosystems. Two of the high-scoring stressors revealed by our analysis (invasive species and coastal engineering) highlight management challenges that might be most successfully addressed at different spatial scales. Although removal of existing invasive species may be possible by local action, it is generally very difficult, and the risk of new species invasions can only be reduced by state, federal, and even international regulations that control the movement of species (i.e., vectors such as ballast water, hull fouling, aquaculture, and aquarium trade; Bax et al. 2001, Ruiz and Carlton 2003). Given the difficulty of eradicating invasive species and reversing their impacts on local ecological communities, prioritizing the reduction of invasive species risks at the regional level may have a high ecosystem-wide payoff. Habitat alteration due to coastal engineering also had high scores in several coastal ecosystems. Although it is difficult to fully reverse, it can be regulated and managed locally at the scale at which it occurs, and there are some options for local habitat restoration. For local-scale management, vulnerability rankings could be different. Fortunately, our framework is fully scalable, with the model weights expected to be consistent across scales and locations, and the output allows for quantitative, relative vulnerability assessments that are often not intuitive or known. The model also provides a rapid method for assessing the potential impact of new stressors relative to existing stressors, and in theory the same stressors in new locations where one would simply need to gather new criteria scores (Part IV of the survey).

These results provide a critical piece of information for moving toward marine ecosystem-based management (EBM) and ocean zoning, but they are clearly not all that is needed for effective management. Among other things, EBM requires consideration of spatial patterns of cumulative impacts of human activities on ecosystems (McLeod et al. 2005, Halpern et al. 2008*a*), and in order to map cumulative impacts, one needs information on the relative vulnerability of ecosystems to those stressors, along with information on the intensity of each stressor (Crowder et al. 2006, Halpern et al. 2008*b*, Halpern et al. 2009). Such mapping also allows one to assess the realized impact of each stressor on each ecosystem, rather than the expected vulnerability as is captured here. Ultimately, effective management and conservation also require assessments of the costs and benefits of any management action, recognition of logistical and financial constraints, compromises for political feasibility, and the flexibility to manage adaptively as new information becomes available. Without knowledge of relative ecosystem vulnerability to different human activities, however, ecosystem-based management will be difficult if not impossible to achieve.

## Acknowledgments

We are very grateful to the 107 experts who participated in the survey and the seven participants in the pre-survey. We thank J. T. Carlton and S. D. Gaines for their valuable comments on an earlier version of this manuscript. This work was generously supported by a grant from the Gordon and Betty Moore Foundation (to B. S. Halpern, K. A. Selkoe, C. V. Kappel, C. M. Crain, and F. Micheli) and the National Center for Ecological Analysis and Synthesis (B. S. Halpern, K. A., Selkoe, and C. V. Kappel).

## Literature Cited

Bax, N., J. T. Carlton, A. Mathews-Amos, R. L. Haedrich, F. G. Howarth, J. E. Purcell, A. Rieser, and A. Gray. 2001. The control of biological invasions in the world's oceans. Conservation Biology 15:1234–1246.

Camerer, C. 1981. General conditions for the success of bootstrapping models. Organizational Behavior and Human Performance 27:411–422.

Carlton, J. T., editor. 2007. The Light and Smith manual: intertidal invertebrates from Central California to Oregon. Fourth edition. University of California Press, Berkeley, California, USA.

Cooke, R. M., and L. H. J. Goossens. 2004. Expert judgment elicitation for risk assessments of critical infrastructures. Journal of Risk Research 7:643–656.

Coombs, C. H. 1964. A theory of data. John Wiley and Sons, New York, New York, USA.

Council on Environmental Quality. 1997. The National Environmental Policy Act: a study of its effectiveness after twenty-five years.

Crain, C. M., K. Kroeker, and B. S. Halpern. 2008. Interactive and cumulative effects of multiple human stressors in marine systems. Ecology Letters 11:1304–1315.

Crowder, L. B., et al. 2006. Resolving mismatches in U.S. ocean governance. Science 313:617–618.

Csiszar, I. 1975. Divergence geometry of probability distributions and minimization problems. Annals of Probability 3: 146–158.

Darling, E. S., and I. M. Cote. 2008. Quantifying the evidence for ecological synergies. Ecology Letters 11:1278–1286.

Dawes, R. M. 1979. The robust beauty of improper linear models in decision making. American Psychologist 34:571–582.

Dawes, R. M., D. Faust, and P. E. Meehl. 1989. Clinical versus actuarial judgment. Science 243:1668–1674.

Du, C., D. Kurowicka, and R. M. Cooke. 2006. Techniques for generic probabilistic inversion. Computational Statistics and Data Analysis 50:1164–1187.

Fernandes, L., et al. 2005. Establishing representative no-take areas in the Great Barrier Reef: large-scale implementation of theory on marine protected areas. Conservation Biology 19: 1733–1744.

Fischhoff, B. 2005. Cognitive processes in stated preference methods. Pages 937–968 *in* K. Mäler and J. Vincent, editors. Handbook of environmental economics. Elsevier, Amsterdam, The Netherlands.

Fischhoff, B. 2009. Risk perception and communication. Section 8.8 *in* R. Detels, R. Beaglehole, M. A. Lansang, and M. Gulliford, editors. Oxford textbook of public health. Oxford University Press, Oxford, UK.

Fischhoff, B., W. Bruine de Bruin, U. Buvenc, D. Caruso, and L. Brilliant. 2006. Analyzing disaster risks and plans: an avian flu example. Journal of Risk and Uncertainty 33:133–151.

Halpern, B. S., C. V. Kappel, K. A. Selkoe, F. Micheli, C. Ebert, C. Kontgis, C. M. Crain, R. Martone, C. Shearer, and S. J. Teck. 2009. Mapping cumulative human impacts to California Current marine ecosystems. Conservation Letters 2:138–148.

Halpern, B. S., K. L. McLeod, A. A. Rosenberg, and L. B. Crowder. 2008*a*. Managing for cumulative impacts in ecosystem-based management through ocean zoning. Ocean and Coastal Management 51:203–211.

Halpern, B. S., K. A. Selkoe, F. Micheli, and C. V. Kappel. 2007. Evaluating and ranking the vulnerability of global marine ecosystems to anthropogenic threats. Conservation Biology 21:1301–1315.

Halpern, B. S., et al. 2008*b*. A global map of human impact on marine ecosystems. Science 319:948–952.

Halpin, P. M., P. Strub, W. T. Peterson, and T. Baumgartner. 2004. An overview of interactions among oceanography, marine ecosystems, climatic and human disruptions along the eastern margins of the Pacific Ocean. Revista Chilena de Historia Natural 77:371–409.

Kurowicka, D., and R. G. Cooke. 2006. Uncertainty analysis with high dimensional dependence modelling. Wiley, Washington, D.C., USA.

Lichtenstein, S., and P. Slovic. 2006. The construction of preference. Cambridge University Press, Cambridge, UK.

McLeod, K. L., J. Lubchenco, S. R. Palumbi, and A. A. Rosenberg. 2005. Scientific consensus statement on marine ecosystem-based management. Communication Partnership for Science and the Sea. ⟨http://www.compassonline.org/pdf_files/EBM_Consensus_Statement_v12.pdf⟩

Metzger, M., R. Leemans, and D. Schröter. 2005. A multidisciplinary multi-scale framework for assessing vulnerability to global change. International Journal of Applied Earth Observation and Geoinformation 7:253–267.

Morgan, M. G., K. Florig, M. DeKay, and P. Fischbeck. 2000. Categorizing risks for risk ranking. Risk Analysis 20:49–58.

Morgan, M. G., and M. Henrion. 1990. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge, UK.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.

Myers, R. A., and B. Worm. 2003. Rapid worldwide depletion of predatory fish communities. Nature 423:280–283.

Neslo, R., F. Micheli, C. V. Kappel, K. A. Selkoe, B. S. Halpern, and R. M. Cooke. 2008. Modeling stakeholder preferences with probabilistic inversion: application to prioritizing marine ecosystem vulnerabilities. Pages 265–285 *in* I. Linkov, E. Ferguson, and V. Magar, editors. Real time and deliberative decision making. Springer, Amsterdam, The Netherlands.

O'Hagan, A., C. E. Buck, A. Daneshkhah, J. E. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley, and T. Rakow. 2006. Uncertain judgments: eliciting expert probabilities. Wiley and Sons, Chichester, UK.

Olson, D. M., and E. Dinerstein. 1998. The global 200: a representation approach in conserving the Earth's most biologically valuable ecoregions. Conservation Biology 12: 502–515.

Pauly, D., V. Christensen, J. Dalsgaard, R. Froese, and F. Torres. 1998. Fishing down marine food webs. Science 279: 860–863.

Payne, J., J. Bettman, and E. Johnson. 1992. Behavioral decision research: a constructive processing perspective. Annual Review of Psychology 43:87–132.

Pew Oceans Commission. 2003. America's living oceans: charting a course for sea change. A report to the nation. May 2003. Pew Oceans Commission, Arlington, Virginia, USA.

Ruiz, G. M., and J. T. Carlton. 2003. Invasive species: vectors and management strategies. Island Press, Washington, D.C., USA.

Smit, B., and H. Spaling. 1995. Methods for cumulative effects assessment. Environmental Impact Assessment Review 15: 81–106.

Spaling, H., and B. Smit. 1993. Cumulative environmental change: conceptual frameworks, evaluation approaches, and institutional perspectives. Environmental Management 17: 587–600.

U.S. Environmental Protection Agency. 1999. Consideration of cumulative impacts in EPA review of NEPA documents. Office of Federal Activities (2252A), Washington, D.C., USA.

Vitousek, P. M., H. A. Mooney, J. Lubchenco, and J. M. Melillo. 1997. Human domination of Earth's ecosystems. Science 277:494–499.

Wilcove, D. S., D. Rothstein, J. Dubow, A. Phillips, and E. Losos. 1998. Quantifying threats to imperiled species in the United States. Bioscience 48:607–615.

Willis, H. H., M. L. DeKay, B. Fischhoff, and M. G. Morgan. 2005. Aggregate, disaggregate, and hybrid analyses of ecological risk perceptions. Risk Analysis 25:405–428.

Worm, B., et al. 2006. Impacts of biodiversity loss on ocean ecosystem services. Science 314:787–790.

## APPENDIX A

Sample survey for a kelp forest ecosystem is provided, except that Part III for both coastal and offshore ecosystems is included to show where scenario names were changed between these two expert groups (*Ecological Archives* A020-049-A1).

## APPENDIX B

Number of surveys completed per ecosystem per region and subtotals per ecosystem and region (*Ecological Archives* A020-049-A2).

## APPENDIX C

Respondents' years of experience per chosen ecosystem and per chosen subregion (*Ecological Archives* A020-049-A3).

## APPENDIX D

Threats listed by participants in the survey that were not on our stressor list (*Ecological Archives* A020-049-A4).