PNAS
www.pnas.org

Supplementary Information for

The diversity bonus in pooling local knowledge about complex problems

Payam Aminpour[1], Steven A. Gray, Alison Singer, Steven B. Scyphers, Antonie J. Jetter, Rebecca Jordan, Robert Murphy Jr., and Jonathan H. Grabowski

[1] To whom correspondence may be addressed. Email: aminpour@msu.edu

This PDF file includes:

Supplementary text S1 to S7
Figures S1 to S18
Table S1 & S2
SI References

Supplementary Information Text

Text S1. Mental Models and Fuzzy Cognitive Maps (FCM)

Mental models (1) are simplified internal representations of reality that allow humans to perceive patterns of cause-and-effect relationships through reasoning and to make decisions. Mental models consist of beliefs and subjective knowledge that are constructed as individuals observe, interact with, and experience the world around them and concurrently develop an internal model to understand and predict how it functions (2). As such, they synthesize knowledge that is acquired through experiential, social, and formal learning. Mental models that represent causal knowledge (e.g., how social and ecological components are interconnected in a natural resource system) can be elicited through cognitive mapping (3). Cognitive maps are representations of mental models in the form of directed graphs. Nodes represent concepts that are part of the mental model and edges (arrows) are used to show the causal relationship between the concepts.

Fuzzy Cognitive Maps (FCM) (4) extend causal cognitive maps to add a dynamic component to their analysis. These are graphical models of system components (nodes) and their causal relationships (edges), with numeric parameterization of edge magnitudes, forming a weighted directed graph (Fig. S1). Relationships (edges) are characterized by a number in the interval of [-1, +1], corresponding to the strength and sign of causal relationships between nodes. They, therefore, provide a semi-quantitative system modeling technique, based on auto-associative neural networks and fuzzy set theory that make cognitive maps computable (see FCM computation section of this *SI Appendix*).

A total of 32 individuals completed the online mental modeling survey including recreational fishers, commercial fishers, and fisheries managers, each creating their own FCM (Fig. S2). Table S1 shows the number of participants from each stakeholder type. In addition, the mean and standard deviation of the number of concepts (i.e. nodes) and connections (i.e. edges) used by individuals to construct FCM representing their mental models about striped bass population dynamics are shown in Table S1.

Text S2. Online mental modeling instructions

Unlike conventional mental modeling practices that are commonly organized through workshops (e.g., ref. (5)) and interviews (e.g., ref. (6)), we demonstrated that, participants can comfortably interact and familiarize themselves with the online platform, and thus knowledge elicitation process can be automated with no influence from researchers and facilitators. This decentralized process allows individuals to freely represent their system knowledge, and therefore it increases the probability that the whole spectrum of knowledge diversity is sampled. However, given that some people do not feel comfortable doing this online (e.g., older people or those without internet access), they may opt out of participation.

The individuals who participated in online mental modeling survey were given a step-by-step instruction how to build a FCM model using the online mental modeling technology. Mental Modeler online tool is modeling software that helps individuals and communities capture their knowledge in a standardized format that can be used for scenario analysis. Based in FCM, users can develop semi-quantitative models of complex social and environmental issues by defining the important components of a system and also the relationships between these components (7). The following step-by-step direction was used to instruct participants:

“

Modeling the Striped Bass Fishery

The purpose of this activity is to create a cognitive map that represents your understanding of the striped bass fishery, and the impacts of recreational and commercial fishing on its overall health.

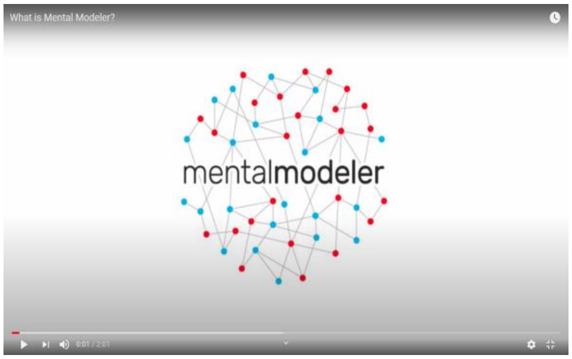
How to make a model with Mental Modeler

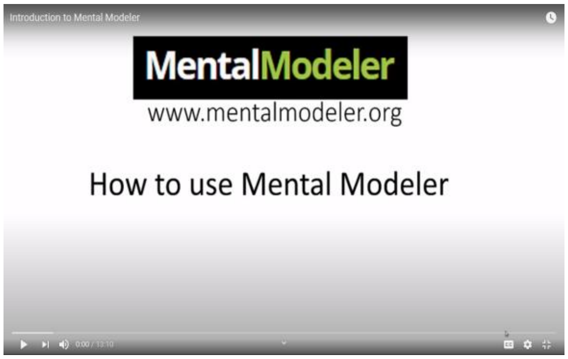
To create your model, you will use the free Fuzzy Cognitive Mapping (FCM) software Mental Modeler (link to the software found below). FCM represents knowledge by defining three characteristics of a system:

- The components (“parts) of the system
- The positive or negative relationships (e.g. influence) between the components
- The degree of influence that one component can have on another, defined using qualitative weightings (e.g. high, medium, or low influence)

Before you begin:

Please watch these two brief videos. In total the videos are 15 minutes long.





1. Access Mental Modeler at www.mentalmodeler.org. Click the “Software” tab at the top of the page, then the “Use the Online Mental Modeler” link. When prompted, enter the following information:

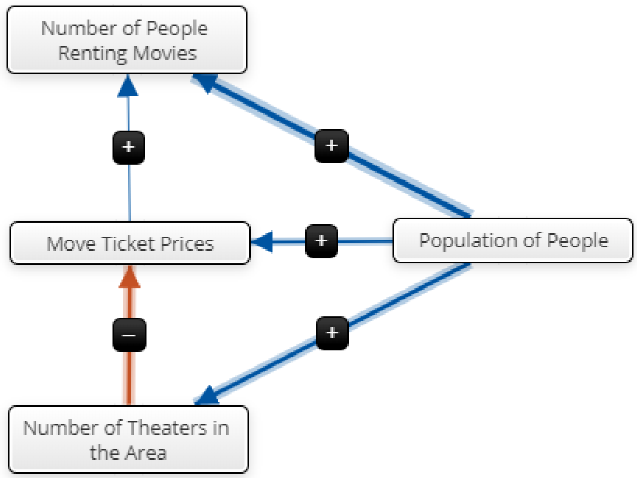
Username: mentalmodeler

Password: mentalmodeler

Once you have access to the online program, you will then build your model of the striped bass fishery by defining the following characteristics based on your understanding: (1) the components that are relevant to the system (boxes), (2) the directional relationships between these components (arrows), and (3) the degree of positive or negative influence that one component has on another (+/- on a sliding scale)

Example Model:

Here is an example of a mental model (see figure below). Let’s start with movie ticket prices as our central variable. Now we can think about factors that affect prices, such as the surrounding population or the number of theaters in the immediate area. We can also think about factors that are affected by movie ticket prices such as the number of the people that rent movies. In the model, as the population increases, it increases the number of people renting movies and increases movie ticket prices and increases the number of theaters in the area (blue arrows) Conversely as the number of theaters increase, it decreases movie ticket prices (see red arrow).



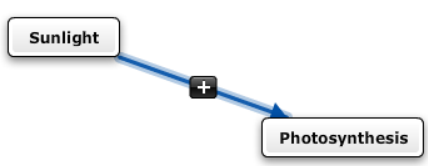
2. Add components you think are important by clicking the "add component" button at the top of the modeler window. Enter text to label a component, and drag the box to reposition it. <u>Components must be variables that can increase or decrease in quality or quantity</u>. Although the components you define in your model will be up to you based on your understanding, please begin by adding the following components:

- Striped bass population
- Recreational fishing for striped bass
- Commercial fishing for striped bass

3. Build your model by adding additional components to the three initial components, and defining the relationships between them. Think about other components that are part of the striped bass fishery (what factors influence bass abundance?), etc. When you add a component, make sure that you think about its relationship to all the other components, and define them accordingly. The following tips will help you create your model

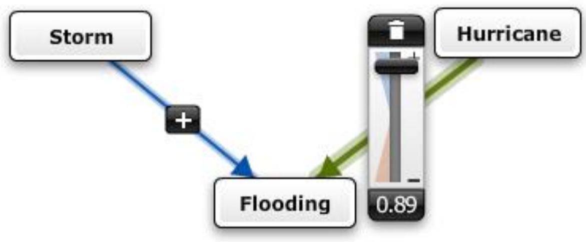
<u>Tip #1:</u> **"Defining Components"** (e.g. boxes) need to be things that can increase or decrease in quality of quantity. For example, these may be things like (a) access to healthy food, (b) food prices, (c) cropland, (d) greenhouse gas emissions. All of these things can increase or decrease. The components should not be things like (e) policy, since "policy" is not something that can increase or decrease. You can add as many components as you think is necessary to represent the system, by clicking on the "add component" button at the top of the modeling screen.

<u>Tip #2:</u> **"Defining Relationships"** between components can either be positive or negative. For example, as the amount of sunlight increases, the process of photosynthesis may also increase. Therefore, you might draw a positive arrow from the component "sunlight" to the component "photosynthesis".



<u>Tip #3:</u> **"Defining Degree of influence"** is the weightings you give to the positive or negative relationships that you define between components. For example a rain-storm may increase the

amount of flooding slightly (represented by a small positive relationship defined between these two components) but a hurricane may increase the amount of flooding a great deal (represented by a high positive relationship defined between these two components)



4. Add the final components. Once your model is complete, please add the following two components if they are not already included:

- Prey abundance
- Water temperature

Then define any relationships between these new components and the existing component.

5. Your model is complete when you feel comfortable with the components you have included and how they interact with each other.

6. Save the model and email it to the project email address. ”

Text S3. Intra group and inter-group variations

In the present paper, we assume that models (i.e. FCMs) of stakeholders are more similar to their peers of the same stakeholder type than members of other types (e.g., recreational fishers construct models that are more similar to each other than commercial fishers or fisheries managers). Therefore, the crowd may demonstrate “modular structures”—there exist subgroups of individuals within the crowd, whose mental models and knowledge are more correlated. To test this assumption, individuals’ graphical cognitive maps were converted into adjacency matrices and the distances between any possible pairs of graphs were calculated using the framework outlined in Box S1 (also see ref (8)). For each stakeholder type k we make two sets:

$$Set_1^k = \{\, \forall\, (s_i, s_j) \in\ C\,(S, 2) \big|\, Typ_i = Typ_j = k \big\}$$

$$Set_2^k = \{\, \forall\, (s_i, s_j) \in\ C\,(S, 2) \big|\, Typ_i \neq Typ_j = k \big\}$$

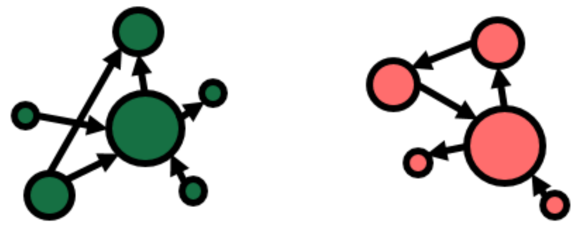
where S is the total number of individuals, (s_i, s_j) is a pair of individuals, and $C(S, 2)$ is the set of all possible paired combinations. For each stakeholder type k we call $Set1$ an *Intra-group* set with all pairs of individuals who share the same type and call $Set2$ an *Inter-group* set with all pairs of individuals from varying types. We measured the distances between pairs of matrices in both *Intra-group* and *Inter-group* sets for each group. Fig. S4 shows that individuals within each group constructs maps that are more similar to one another compared to members of other groups. For all three groups, the *Intra-group* distances are shorter than *Inter-group* distances, while independent t-tests revealed that these differences are statistically significant with $p < 0.05$ for commercial fishers and managers. Recreational fishers demonstrate the same trend but this is not significant at the level of 0.05.

Additionally, we measured the ratio of disagreement to agreement for a short-list of relationships (causal relationships that are mentioned by individuals significantly more frequently than other edges) and compared this ratio (γ) within each group (i.e., *Intra*) and across groups (i.e., *Inter*). Comparing the γ ratio intra-group and inter-group revealed that individuals within each group make more similar assumptions about the causal relationships, while inter-group variations, relatively, demonstrate higher values for γ (Table S2).

Box S1. Calculating the distance between a pair of graphs

1. Directed graphs

Each model is a directed weighted graph. We find the distances between each pair of modes.



2. Adjacency matrix

$\boldsymbol{A}$ is a square matrix used to represent a graph. The elements of the matrix indicate the edge weights between pairs of nodes that are adjacent in the graph.

$$A^1 = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \qquad A^2 = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

3. Laplacian matrix

For each adjacency matrix we find the Laplacian matrix ($\boldsymbol{L = D - A}$), where $\boldsymbol{D}$ is the degree matrix, and then we normalize it . $\boldsymbol{L^{sym}}$ is the normalized symmetric laplacian which is positive-semidefinite and thus ($\lambda_i \geq 0 \; for \; all \; i$), where $\boldsymbol{\lambda_i}$ is the $\boldsymbol{i^{th}}$ largest eigenvalue.

$$L = D - A \qquad L^{sym} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$$

$$L_1^{sym} \qquad L_2^{sym}$$

4. Vector of eigenvalues

We find all eigenvalues for each Normalized symmetric Laplacian matrix and put them in a one-dimensional array (i.e. a vector).

$$v_e^1 = \begin{pmatrix} \lambda_1^1 \\ \lambda_2^1 \\ \vdots \\ \lambda_m^1 \end{pmatrix} \qquad v_e^2 = \begin{pmatrix} \lambda_1^2 \\ \lambda_2^2 \\ \vdots \\ \lambda_m^2 \end{pmatrix}$$

5. Distance between vectors

We find the smallest $\boldsymbol{k}$ such that the sum of the $\boldsymbol{k}$ largest eigenvalues constitutes at least 90% of the sum of all of the eigenvalues. If the values of $\boldsymbol{k}$ are different between the two graphs, we use the smaller one ($\boldsymbol{k^*}$). $\boldsymbol{ED}$ represents the distance between two graphs

$$\min_k \{ \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{m} \lambda_i} > \frac{90}{100} \} \qquad k^* = \min_j \{k^j\}$$

$$ED = \sum_{i=1}^{k^*} \left(\lambda_i^1 - \lambda_i^2\right)^2$$

Text S4. Mental models aggregation

Given the type of problem or the task in a collective intelligence experimental design, various aggregation mechanisms could be used to reach a collective solution. General aggregation mechanisms are *average rule* (i.e., using a central tendency measure like mean); *addition rule* (i.e., adding or crowdsourcing fragmented information from multiple sources to build a whole thing); *majority rule* (i.e., using voting mechanisms); and *convergence rule* (i.e., reaching a consensus by deliberation or convergence of opinions).

In the present study, and based on theoretical assumptions about the negative impact of social influence on the wisdom of crowds (WOC) effect (see ref. (9)), participants were not allowed to socially interact with each other during the mental modeling process, and this was required to meet the *independence* condition. Therefore, convergence rule (i.e., reaching a consensus about different parts of the model through social interactions and deliberation) could not be an appropriate aggregation mechanism due to our experiment's objectives. On the other hand, the majority rule could only determine whether or not individuals agree on parts of the models (cause-and-effect relationships) and could not be used to aggregate the magnitude of causal relationships (i.e., the edge weights). The most appropriate aggregation mechanisms were the *average* and *addition* rules and their synthetization to not only merge unique information from multiple individuals (add parts that are unique to only one individual), but also average the overlapping information presented in the maps of more than one individual and thus operationalize the WOC effect.

Stakeholder-specific models (homogenous groups)

Individual mental models represented as FCMs can be aggregated mathematically using matrix algebra operations on their adjacency matrices. These aggregated models—also referred to as "community maps"—can be used to represent the knowledge and perception of a group of participants and thus provide a tool for knowledge-pooling (10). To combine mental models of a homogeneous group with individuals from a specific stakeholder type we summed their maps (*addition rule*) and used the arithmetic mean (i.e., simple average) of edge weights that are shared among multiple FCMs (*average rule*):

$$A_{ij}^{FCM_g} = \sum_{p=1}^{N} A_{ij}^{FCM_p} \Big/ \sum_{p=1}^{N} (1 \mid A_{ij}^{FCM_p} \neq 0) \quad (\boldsymbol{S1})$$

where A^{FCM_p} is the adjacency matrix of the FCM of participant *p*, *N* is the total number of participants in a group, and $A_{ij}^{FCM_p}$ indicates the element of this matrix with the value equals to the weight of the edge between node *i* and *j*. FCM_g represents the aggregated FCM of a group with the corresponding adjacency matrix A^{FCM_g}. We used the above aggregation method to create stakeholder-specific (homogeneous) models of recreational fishers (Fig. S5), commercial fishers (Fig. S6), and fisheries managers (Fig. S7).

Crowd model (diverse group)

Our sample includes individuals from three stakeholder types, and we showed that each type constructs models that are more similar to one another than to members of another group (Fig S4). Therefore, the crowd demonstrates "modular structures"—there exist subgroups of individuals within the crowd, whose mental models and knowledge are more correlated. In such cases, prior theoretical (e.g., ref. (11)) and empirical evidence (e.g., ref. (9)) demonstrated that the collective response can be enhanced if the aggregation takes place in two levels: aggregating responses within the modules, followed by an aggregation across the modules. Therefore, to build an aggregated mental model of diverse stakeholders (i.e. the crowd model), we used a multi-level aggregation technique (Fig. S3). The first level of aggregation was achieved by averaging the mental models of individuals from the same stakeholder type (see Eq. S1). At the second level, we aggregated the averaged models across groups using the median of their adjacency matrices:

$$A_{ij}^{FCM_{crowd}} = Median\left(A_{ij}^{FCM_{g1}}, A_{ij}^{FCM_{g2}}, \dots \ , A_{ij}^{FCM_{gn}}\right) \quad (\boldsymbol{S2})$$

where $A_{ij}^{FCM_{crowd}}$ indicates the element of the adjacency matrix of the crowd model with the value equals to the weight of the edge between node *i* and *j*. In our case, there are three types of stakeholders: recreational fishers, commercial fishers and fisheries managers. Thus, we used the median of edge-weights across three averaged stakeholder-specific maps (i.e. FCM_{g1}, FCM_{g2}, and FCM_{g3}) to build the diverse crowd model (Fig. S8).

How does median outperform other mechanisms?

At the second level of aggregation, we could use the arithmetic mean of averaged maps to aggregate across the stakeholder types; however, prior to aggregate across types, we formed stakeholder-specific models that consist of same-type individuals and this could likely amplify the accumulation of stakeholder-specific biases. Thus, the distribution of models across the stakeholder types could most likely be skewed to some degree. In such cases, median has been shown to outperform the arithmetic mean in likely skewed distributions (see ref. (12)). Yet, voting mechanism (the majority rule) could be used to specify whether or not groups agree on causal interdependences; however, as stated earlier, the voting mechanism could not aggregate causal relationships' strength or the magnitude of links (i.e. edge weights). Therefore, we decided to use the *median* of group maps to address these issues. To more comprehensibly demonstrate the supremacy of *median* over *mean* and *majority* we used made-up examples and compared the performance of different aggregation mechanisms (see Box S2).

Box S2. Made-up examples for comparing different aggregation methods

	Group 1	Group 2	Group 3	Aggregation methods		
				Median	Mean	Majority
Ex. 1	$A \xrightarrow{0.5} B$	$A \xrightarrow{0.25} B$	$A \xrightarrow{0.75} B$	$A \xrightarrow{0.5} B$	$A \xrightarrow{0.5} B$	+
Ex. 2	$A \xrightarrow{0.5} B$	$A \xrightarrow{0.25} B$	x	$A \xrightarrow{0.25} B$	$A \xrightarrow{0.25} B$	+
Ex. 3	$A \xrightarrow{0.5} B$	$A \xrightarrow{-0.25} B$	x	0	$A \xrightarrow{0.083} B$	0
Ex. 4	$A \xrightarrow{0.5} B$	x	x	0	$A \xrightarrow{0.167} B$	0
Ex. 5	$A \xrightarrow{-0.25} B$	$A \xrightarrow{-0.75} B$	$A \xrightarrow{1.0} B$	$A \xrightarrow{-0.75} B$	0	–
Ex. 6	$A \xrightarrow{-0.25} B$	$A \xrightarrow{-0.25} B$	$A \xrightarrow{1.0} B$	$A \xrightarrow{-0.25} B$	$A \xrightarrow{0.167} B$	–

Example 1 represents a case where all three groups have a positive link from node A to node B and the distribution of edge weights across groups is symmetric. In this case, the majority agrees on a positive relationship between A and B and both mean and median produce the same results. *Example 2* represents a case where two groups have a positive link from node A to node B, while one group does not have such a link, and the distribution of edge weights across groups is

symmetric. Similarly, in this case, the majority agrees on a positive relationship between A and B and both mean and median produce the same results.

However, in many cases, the distribution of edge weights across groups is likely skewed (not symmetric) and/or groups do not agree on parts of the models. *Example 3* represents a case where one group has a positive link, one group has a negative link, and one group does not have a link from node A to node B. In this case, there is no majority agreement on the existence of a link between A and B. Among mean and median, the latter aligns with this lack of agreement while the former still maintains a link in the aggregated map.

Example 4 represents a case where only one group has a positive link from node A to node B while the other two groups have no link. Similarly, the majority agrees on no link between A and B in this case. Again, the median aligns with this agreement while the mean still maintains a link in the aggregated map, which does not properly represent the majority opinion.

Example 5 represents a case where two groups have a negative link, while the other group has a positive link from node A to node B, and the sum of negative edge weights equals the positive edge weight. In this case, the majority agrees on a negative link between A and B. Among mean and median, the latter aligns with this agreement while the former causes the loss of a link in the aggregated map because the mean of edge weights is zero.

Example 6 is the same as the previous example, but in this case the sum of negative edge weights is smaller than the positive edge weight. Again, the majority agrees on a negative link between A and B; however, only median aligns with this agreement while using the mean yields a positive link from A to B in the aggregated map, which is opposed to the majority opinion.

In all of these examples, *median* produces aggregated results which agree with the majority opinion regarding the presence, absence, or the sign of the link, while also mimics an average-like mechanism to aggregate edge weights. However, in four out of six cases, the mean produces results that do not correctly represent the majority opinion. These examples clearly demonstrate how *median* outperforms the *mean* in aggregating group models across stakeholder types.

Text S5. Fuzzy cognitive maps computation

FCM models are semi-quantitative simulation models (13) that can be used to assess the perceived dynamic behavior of the system they represent (10, 14, 15). Here, we used FCM computational analysis to demonstrate how stakeholders, based on their collective perceptions and knowledge, predict the changes in the state of system's elements (e.g., striped bass population) given an initial change in one or combination of concepts called scenario inputs (e.g., decrease water quality or increase water temperature) (also see refs. (16, 17) for details about scenario analysis). An increase (or a decrease) in a concept affects all concepts that are causally dependent upon it and thus triggers a cascade of changes to other system concepts. Subsequently, the newly activated concepts affect other concepts they have causal interdependency with, and this iterative propagation of the initial change continues until the system converges into a new so-called "system state" (18). By comparing the system states (i.e. the activation of concepts) before and after initiation of a change, FCM can be used to implement "what if" scenario analysis, and therefore represent the perceived dynamic behavior of the system (in this case, striped bass fisheries).

To run a scenario, the activation value of one or more concepts (i.e., scenario nodes) in a FCM was manipulated and forced to stay at either +1 (an increase) or -1 (a decrease). This initial change passes through the network of nodes and connections including feedback loops until the system reaches a new state. The consequent alterations in the state of other system concepts were calculated by subtracting their initial values from their values after the scenario was introduced and system converged into a new state. The initial value of each concept—also known as the "steady state"—is calculated using the following formula:

$$c_i^{(k+1)} = f\left(c_i^{(k)} + \sum_j c_j^{(k)}.A_{ji}\right) \quad (\boldsymbol{S3})$$

where $c_i^{(k+1)}$ is the value of concept C_i at iteration step $k+1$, $c_i^{(k)}$ is the value of concept C_i at iteration step k, $c_j^{(k)}$ is the value of concept C_j at iteration step k, and A_{ji} is the weight of the edge relationship between C_j and C_i. Function $f(x)$ is the threshold function (aka "activation function"), that was used to squash the concept values at each step to a normalized interval between -1 and 1. Two popular kinds of activation functions which are commonly used in FCM studies are *sigmoid* and *hyperbolic tangent* (19–22). In this study, we used a hyperbolic tangent function (see ref. (20) for more details about hyperbolic tangent function and FCM dynamics):

$$f(x) = \ Tanh\ (\lambda x) = \ \frac{e^{\lambda x} - e^{-\lambda x}}{e^{\lambda x} + \ e^{-\lambda x}} \quad (\boldsymbol{S4})$$

where λ is a real positive number (in our case $\lambda = 0.5$) which determines the steepness of the function f. The value of each concept under a scenario was computed using the same formula (Eq. S3), but here, scenario nodes were forced to take fixed values (either +1 or -1 to simulate an increase or a decrease respectively). The scenario outcomes were then calculated as the differences between the values of the system's concepts when the system was self-administered and when it was forced by fixed manipulations in the state of scenario concepts (10, 18). For each concept C_i the change in its value as a result of running a scenario is:

$$D_i^{sc} = c_i^{ss} - c_i^{sc} \qquad (\boldsymbol{S5})$$

where D_i^{sc} is the change in the value of concept C_i, c_i^{ss} is the value of concept C_i in the steady state, and c_i^{sc} is the value of concept C_i after converging into a new state while scenario concepts are clamped on fixed values.

Fig. S10 shows the results of 6 scenarios that were used in model evaluations. Since we compare the system states to an initial baseline ("steady state"), scenario outcomes represent *relative changes* in concepts' activation, and therefore the normalized results (i.e., normalized patterns of relative changes in system states) are independent from the choice of activation function. To statistically test this independency and check the robustness of our findings, we used both sigmoid and tangent hyperbolic functions and compared the FCM dynamic outcomes for 6 scenarios that were used in model evaluations. Fig. S11 shows the correlation between results of sigmoid and hyperbolic tangent activation functions. For all four models, Pearson Correlation (r) is significant at the 0.01 level (2-tailed), suggesting that the patterns of relative changes in the system states are independent from the choice of activation function.
In addition, there are no scenarios where the crowd model produces outcomes that differ from all three groups (i.e., recreational fishers, commercial fishers, and managers). This is shown in Fig. S12 using simple correlation matrix of scenario outcomes between the four models. Among 6 scenarios, outcomes of four models are less correlated for *Increase water temperature*, *Decrease water quality*, and *Increase demand*.

Text S6. Striped Bass Fishery and Ecosystem-Based Fishery Management Model

Striped Bass

Striped bass (*Morone saxatilis*) are an anadromous fish native to the east coast of the United States and can reach sizes well over one meter. Today, most the key spawning grounds are located in the mid-Atlantic region of the United States, such as the Chesapeake Bay and Delaware River (23). After spawning, a large contingent of striped bass make an annual migration north during the spring and summer into New England rivers, estuaries, and coastal waters, where they forage for a diversity of fish and invertebrate prey, including the American lobster (*Homarus americanus*), Atlantic menhaden (*Brevoortia tyrannus*), and American sand lance (*Ammodytes americanus*) (24, 25). Often traveling in large schools, Striped Bass can consume sizeable portions of local prey populations and as such, play a critical role in coastal ecosystems (26).

Striped bass are also integrally connected to people throughout the east coast, where they hold both historic and present-day importance; striped bass were a valuable food source for the early European colonizers of New England and offer opportunities in numerous states for commercial fishers and for a diversity of recreational fishers with different motivations for fishing (27). However, they have undergone a series of large population fluctuations, including a historic low in the late twentieth century, which spurred substantial regulatory changes across their range (28). Since then, striped bass populations have enjoyed a full recovery, but more recent decreases have again resulted in tighter fishing restrictions from state and regional management agencies.

Ecosystem-Based Fishery Management

Ecosystem-Based Fishery Management (EBFM) has been identified as the most efficient and effective model for the National Oceanic and Atmospheric Administration (NOAA) Fisheries (www.fisheries.noaa.gov). EBFM can help maintain interconnected ecosystems and human societies in a healthy, productive, and resilient condition, while considering the full range of trade-offs and complex social-ecological interactions. Here we used a striped bass EBFM model with four critical ecosystem considerations (i.e., sub-models): *Habitat suitability, Stock dynamics, Food-webs,* and *Socioeconomics* (29). This model is developed by EBFM Striped Bass Species Team at Maryland Sea Grant to build a sustainable ecosystem for fisheries within Chesapeake Bay. However, the identified critical sub-models are generic, and this EBFM model can be applied to Atlantic striped bass fisheries in Massachusetts (MA). Also, Chesapeake Bay is a primary spawning grounds for striped bass, and they migrate to MA in later stages of their lifecycle, and therefore these ecosystems are interconnected (24, 25).

We therefore developed a casual model of social-ecological relationships based on what has been identified in ref. (29) as the critical ecosystem components and causal interdependences. We used *Mental Modeler* online tool and built a directed graph to represent cause-and-effect relationships between components (Fig. S17). This graphical model was then exported to an adjacency matrix that was used as a reference point against which the performance (i.e., success) of stakeholder-driven models were measured (see the next section).

Text S7. Monte Carlo Analysis (MCA): The framework

To conduct the MCA we used the framework shown in Box S2. We made virtual communities of agents with randomly generated cognitive maps in two ways:

First, to mimic our real-world experiment, wherein approximately 10 individuals participated from three stakeholder types, we generated virtual communities of size 10 for each stakeholder type. In each replicate, and for each stakeholder type (i.e., recreational fishers, commercial fishers, and managers), 10 virtual agents with random cognitive maps were generated (in sum, 30 virtual agents, which closely approximates our experiment). Aggregated models of homogeneous communities were generated by averaging individual maps within each group. Finally, the diverse crowd model was generated by aggregating maps across groups. This process was repeated for 10,000 replicates where the performance of three homogeneous groups and the diverse group were compared.

Second, to further demonstrate the impact of "*diversity*" on group performance, we generated virtual communities with different levels of diversity. To that end, in each replicate, and for each stakeholder type (i.e., recreational fishers, commercial fishers, and managers), a random number of virtual agents were generated such that the total number of all virtual agents is equal to 30. This MCA closely approximates the size of the crowd in our experiment, but explores different combinations of three stakeholder types which may lead to different levels of diversity. To measure each group's diversity, we used Shannon's entropy formula as described in ref. (30). This diversity index takes into account both the richness (i.e., how many unique stakeholder types exist in a group) and the evenness (i.e., how even are the proportions of stakeholder types) of groups:

$$H_g = -\sum_i p_i \times Ln(p_i) \quad and \quad for\ \forall\ i \in \{U\},\ \ p_i = \frac{n_i}{n_g} \qquad (\boldsymbol{S6})$$

Where H_g is the diversity index for group g, U is the set of unique stakeholder types, n_i is the number of individuals within the group whose type is i, and n_g is the total number of individuals in the group, and in this case, $n_g = 30$ for all virtual groups. This process was repeated for 10,000 replicates.

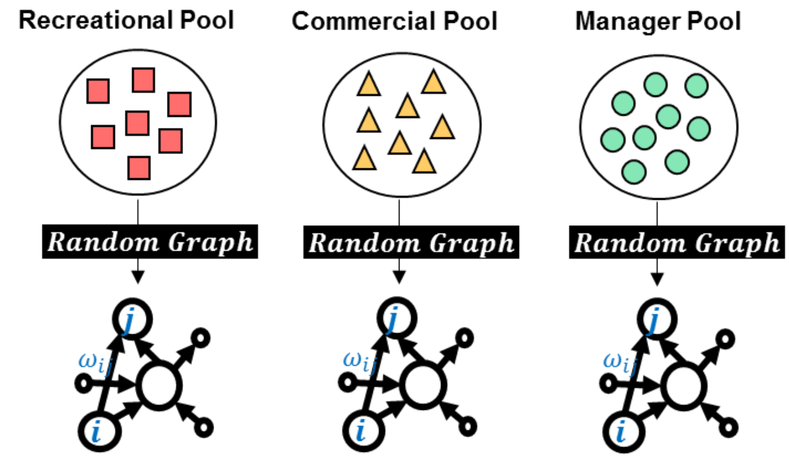
In addition, we used MCA to further demonstrate the supremacy of using *median* over using *mean* in aggregating models across groups while increasing the group size. To that end, we generated virtual crowds of random sizes between 3 and 99 with all three stakeholder types, aggregated group models across types using *mean* and *median*, and explored the impact of crowd size on the crowd performance for both approaches. In each replicate, the number of virtual agents from each stakeholder type was the same, but each time, the total crowd size was different. Because in each group there are the same number of agents from each stakeholder type, the diversity level is maximum (i.e., $H_g = -\sum_{i=1}^{3} \frac{1}{3} \times \ln(\frac{1}{3})$) for all virtual groups, while their sizes are different. Our results revealed that, once the *median* is used, the crowd performance may improve as the crowd size increases. However, once the *mean* is used, the crowd performance deteriorates in larger crowds (Fig. S18).

In all of these MCAs, EBFM was used as a reference point to measure group success (i.e., performance). That is, the distance of each group's aggregated model to EBFM is measured via a bilateral (graph-spectral graph) matrix similarity index that takes into account 1) the overlap of their edges and 2) the distance between the eigenvalues of their normalizes Laplacian matrices.

Box S3. Monte Carlo Analysis framework

1. Random model generation

We create virtual models from each group (i.e. stakeholder type) where an edge exists proportional to the number of people in the group who included that edge (P_{ij}). The edge weight (ω_{ij}) is randomly drawn from a normal distribution with a mean (μ) and standard deviation (σ) representing the distribution of weights assigned to the edge between nodes i *and* j by all individuals in the group who included that edge in their cognitive maps.

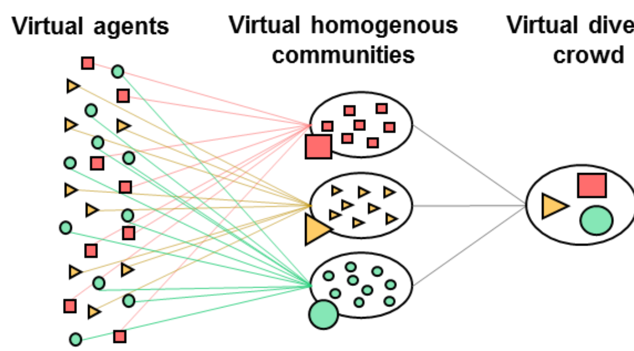


$$for\ \forall\ i,j\ \in Nodes \quad P_{ij} = \frac{1}{n}\sum_{k=1}^{n}\theta_k, \quad \theta_k = \begin{cases} 0 & a_{ij} = 0 \\ 1 & a_{ij} \neq 0 \end{cases}$$

$$\omega_{ij} \sim N(\mu, \sigma^2)$$

3. Model aggregation

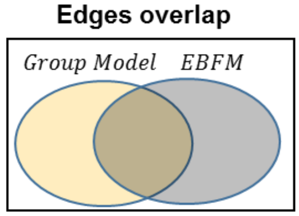
We create virtual agents (with randomly generated cognitive maps) from each group and aggregate their maps. For virtual communities with only one agent type we use *equation S1* and for communities with multiple agent types we use *equation S2* to aggregate individual maps.



3. Baseline comparisons against EBFM

To calculate each group's performance we take into account two measures: Edges overlap (*Edges Similarity*) and the Eigenvalues distance (*ED*) (see Box S1) between the group aggregated model and the EBFM model as the reference point.

Edges overlap



Eigenvalues distance to EBFM

$$ED = \sum_{i=1}^{k^*}\left(\lambda_i^g - \lambda_i^{EBFM}\right)^2$$

$$Edges\ Similarity = \frac{Edg^g \cap Edg^{EBFM}}{Edg^g \cup Edg^{EBFM}}$$

4. Performance index

Performance index is the multiplication of inverted eigenvalues distance and the edge similarity times 100.

$$Performance = \left(\frac{100}{ED}\right) \times (Edges\ Similarity)$$

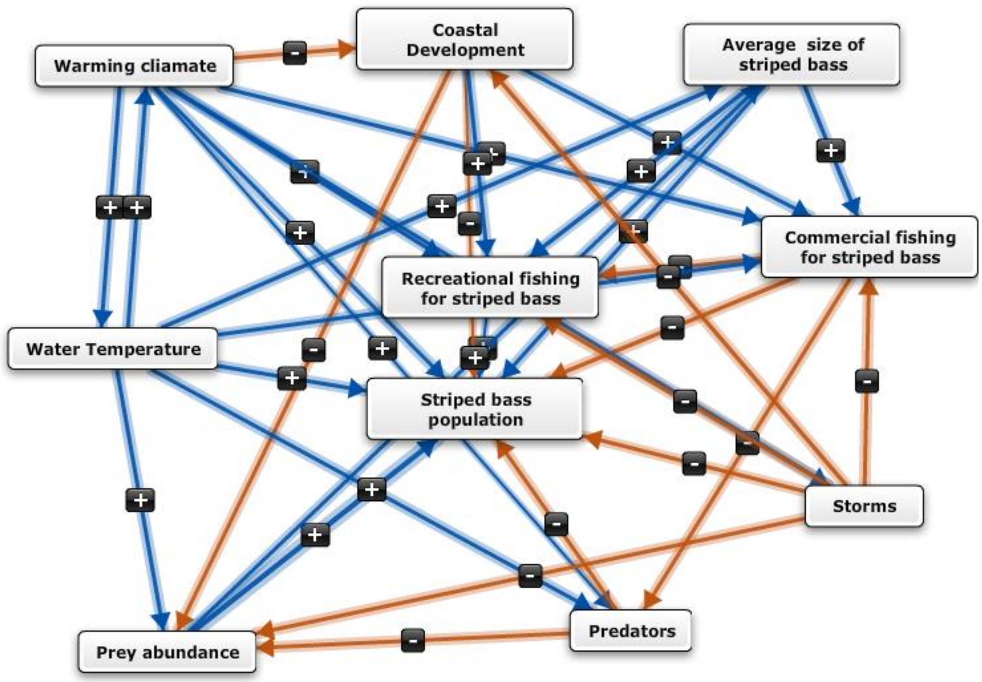


Fig. S1. An example of a fuzzy cognitive map (FCM) representing a mental model about the striped bass fishery. The FCM was created using the *Mental Modeler* online platform at www.mentalmodeler.org. Boxes demonstrate system concepts defined by the individual modeler and arrows indicate causal relationships between concepts.

Fig. S2. All individual fuzzy cognitive maps (FCM) representing the mental models of 32 participants about striped bass fishery in Massachusetts.

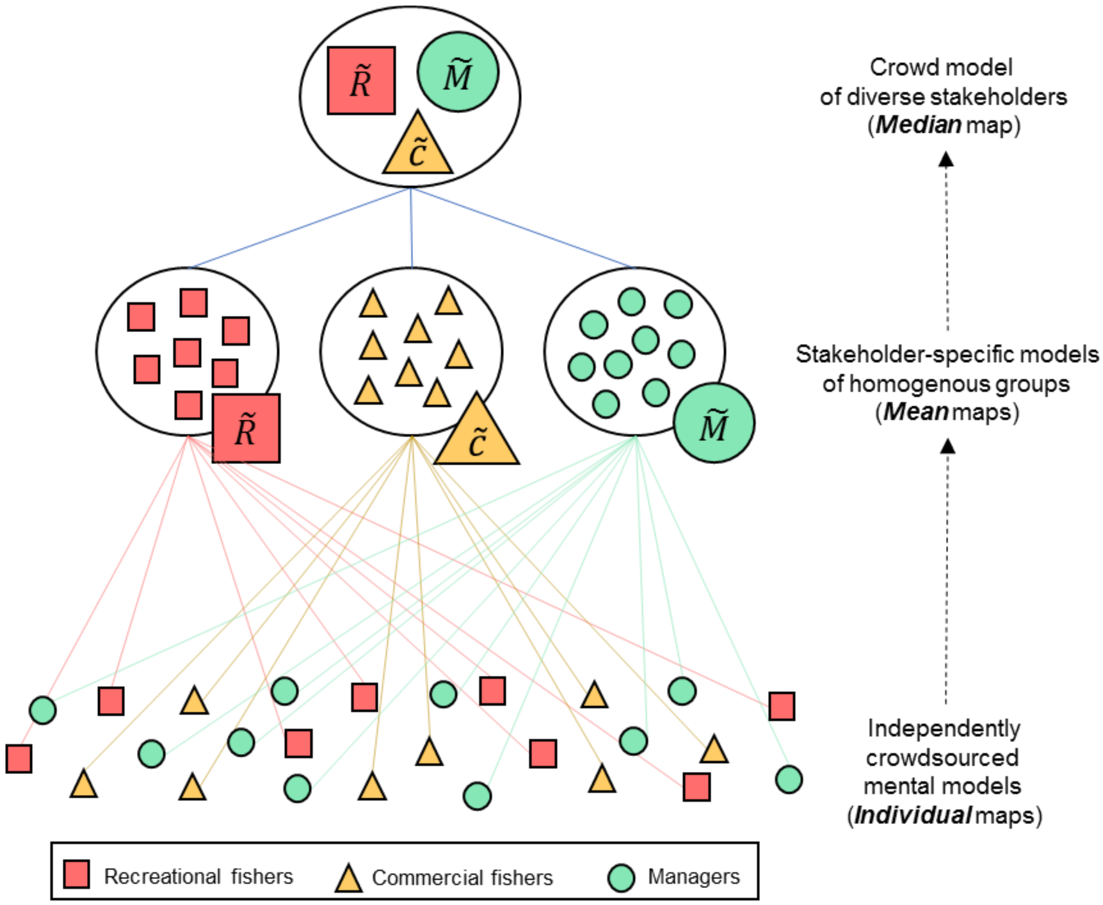


Fig. S3. Multi-level aggregation method. At the first level, individual maps are aggregated by stakeholder groups using the arithmetic mean of their fuzzy cognitive maps' edge weights. In the second level, the resulting group means are aggregated using the median of their edge weights to produce the crowd model.

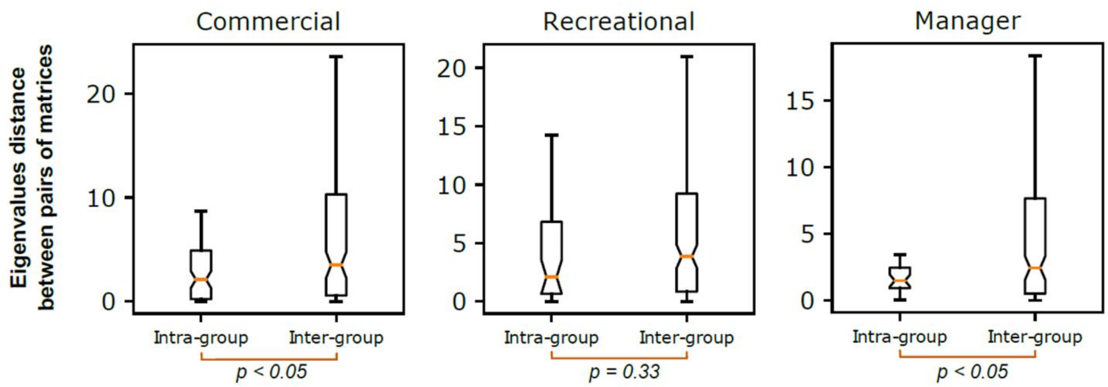


Fig. S4. Intra-group versus Inter-group distances between pairs of cognitive maps. Y-axis in *A* shows the eigenvalues distance between pairs of matrices (see Box S1), and values closer to zero indicate more similar models. For all three groups, individuals within groups (i.e., *intra-group*) have more similar models than individuals across groups (i.e., *inter-group*).

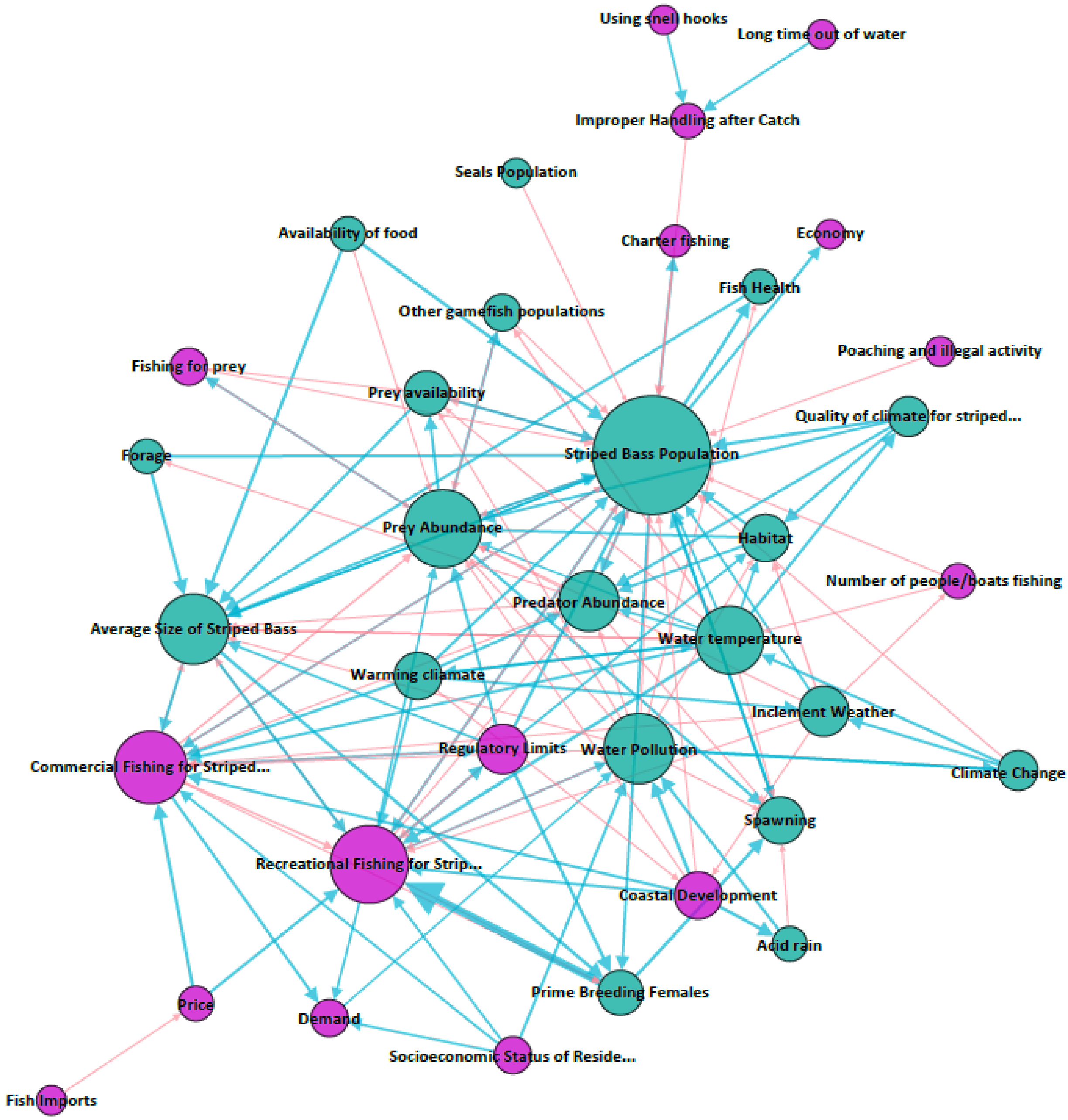


Fig. S5. Aggregated mental model of recreational fishers. Circles demonstrate unique system concepts mentioned by the individuals of type recreational fisher. Ecological-dimension concepts are green and human-dimension components are purple. Weighted blue/red arrows indicate positive/negative causal relationships between concepts. Arrow thickness represents the strength of the causal relationships ranged from -1 to +1. The weights of the arrows are computed using Eq. S1.

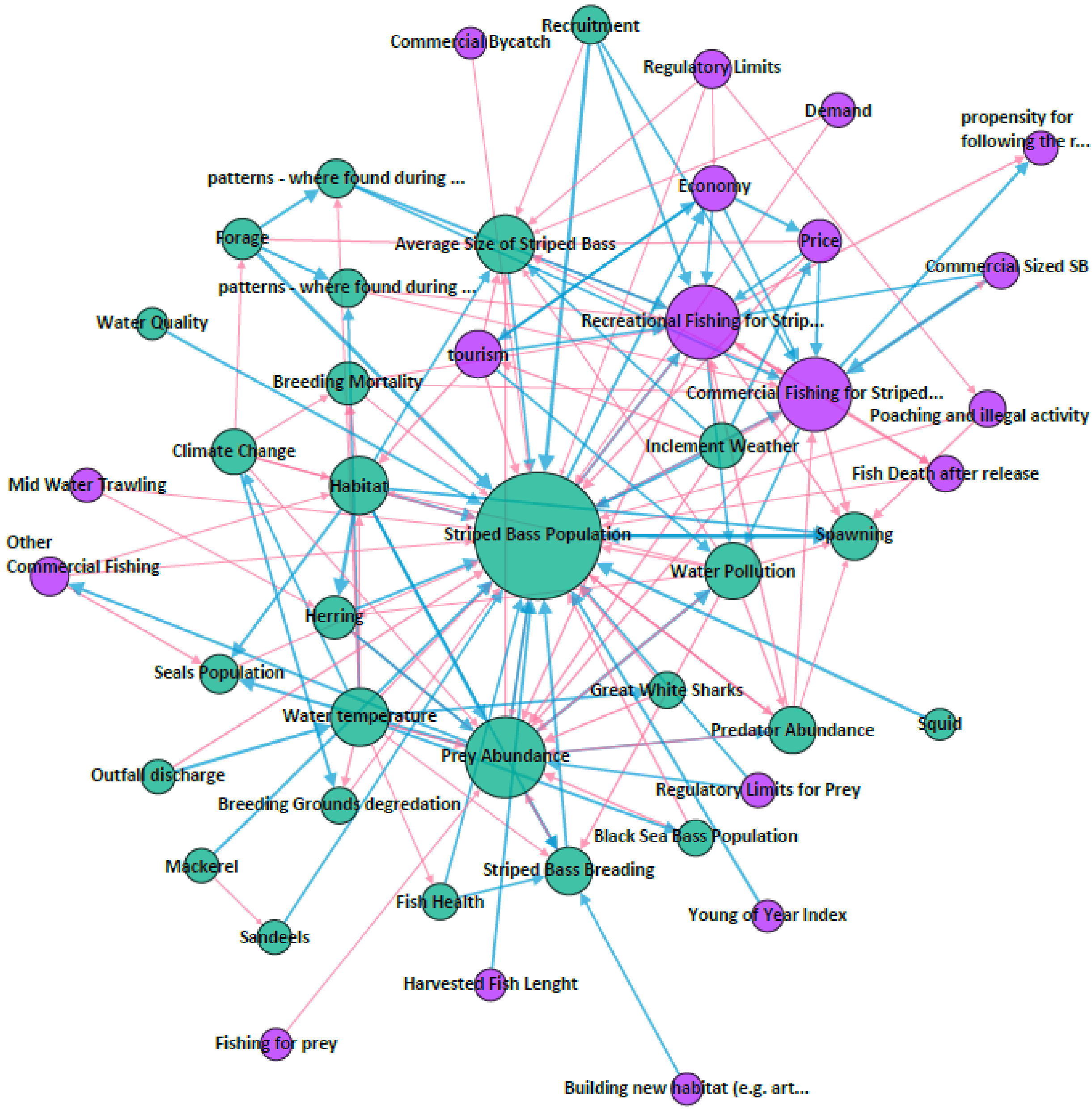


Fig. S6. Aggregated mental model of commercial fishers. Circles demonstrate unique system concepts mentioned by the individuals of type commercial fisher. Ecological-dimension concepts are green and human-dimension components are purple. Weighted blue/red arrows indicate positive/negative causal relationships between concepts. Arrows thickness represents the strength of the causal relationships ranged from -1 to +1. The weights of the arrows are computed using Eq. S1.

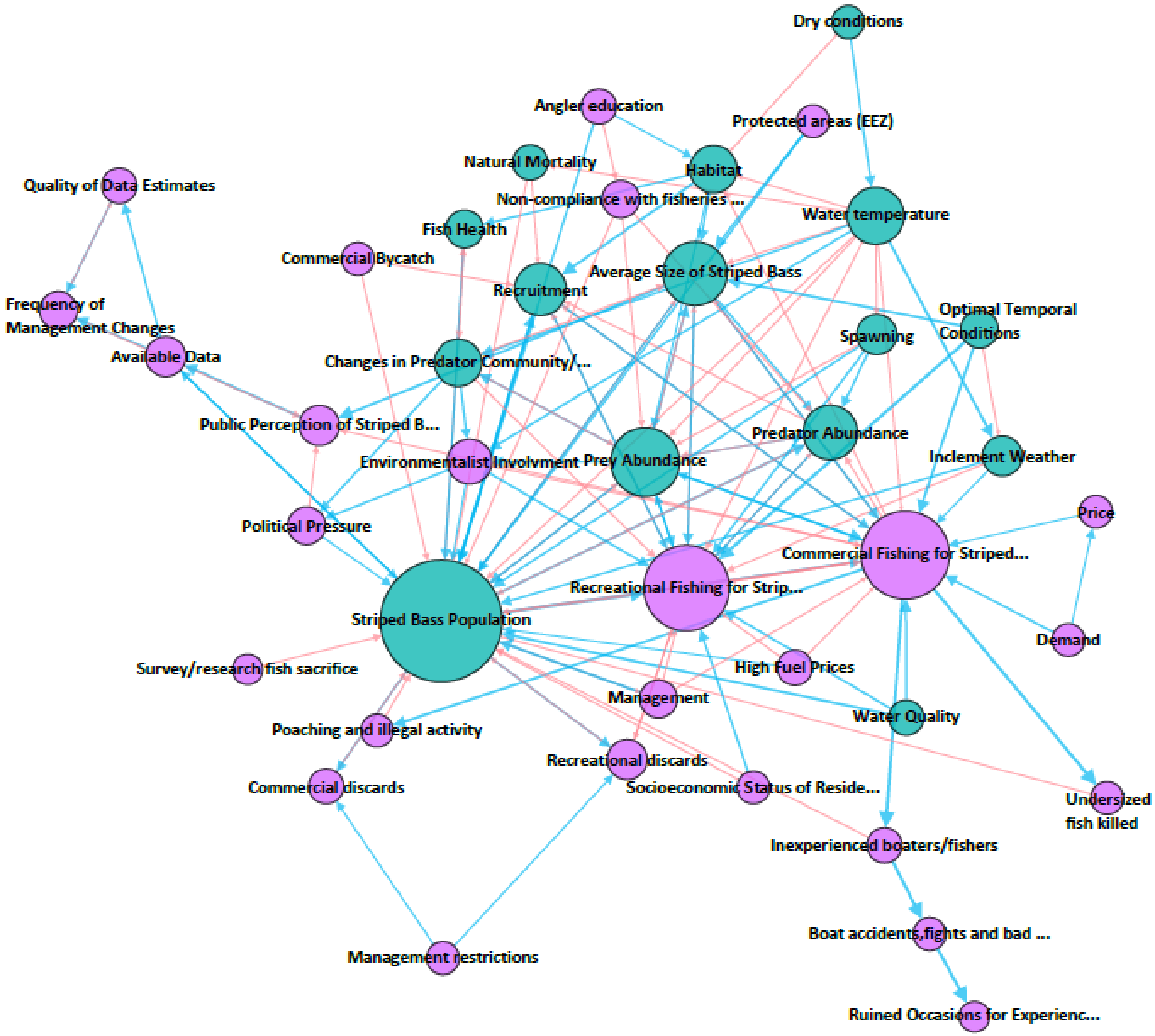


Fig. S7. Aggregated mental model of fisheries managers. Circles demonstrate unique system concepts mentioned by the individuals of type manager. Ecological-dimension concepts are green and human-dimension components are purple. Weighted blue/red arrows indicate positive/negative causal relationships between concepts. The arrows thickness represents the strength of the causal relationships ranged from -1 to +1. The weight of the arrows are computed using Eq. S1.

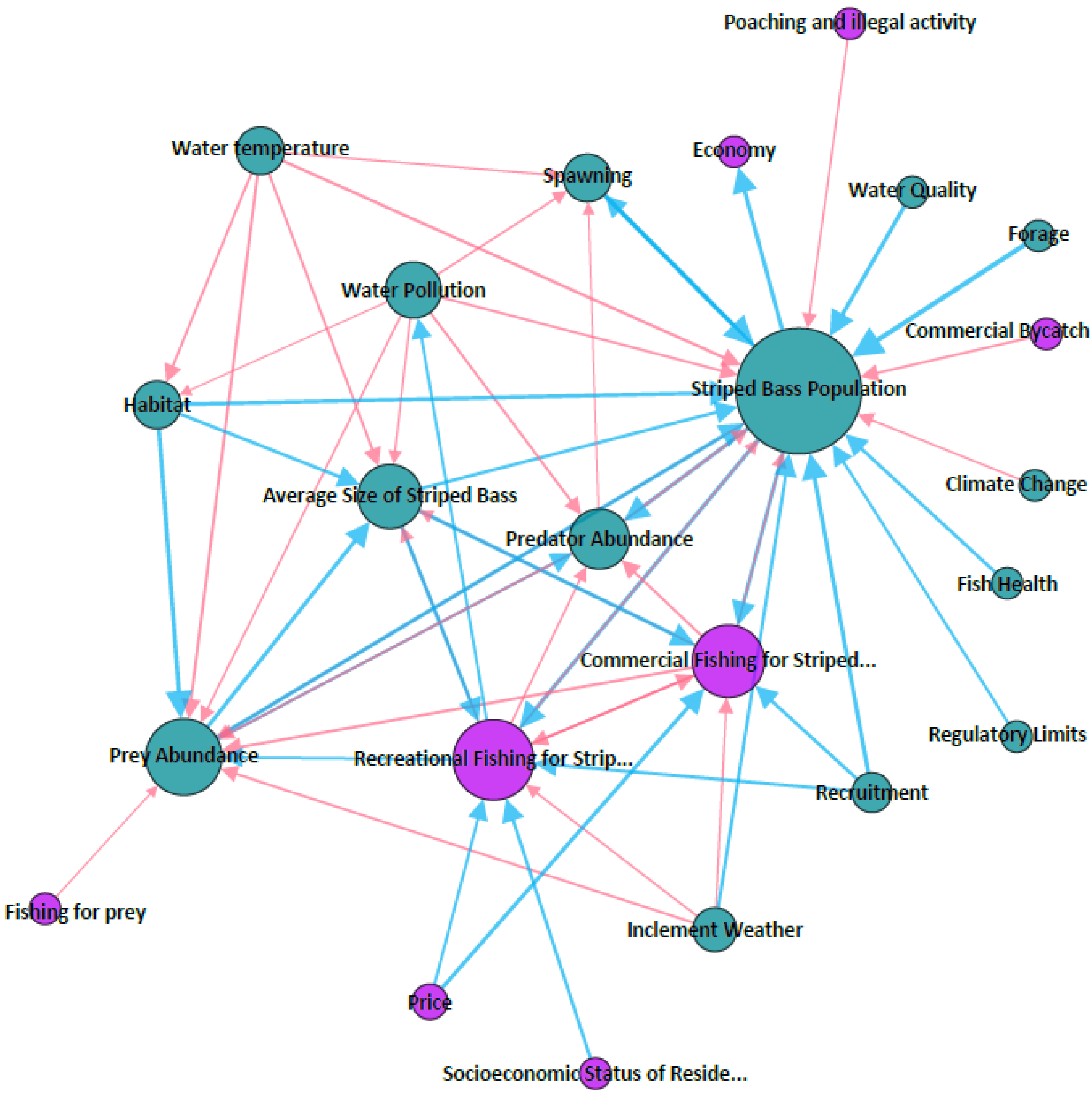


Fig. S8. Aggregated mental model of the diverse crowd. Circles demonstrate a parsimonious list of system concepts mentioned by all individuals of all stakeholder types. This parsimonious list of system concepts is obtained by a multi-level aggregation method. Ecological-dimension concepts are green and human-dimension components are purple. Weighted blue/red arrows indicate positive/negative causal relationships between concepts. Arrows thickness represents the strength of the causal relationships ranged from -1 to +1. The weights of the arrows are computed using Eq. S2.

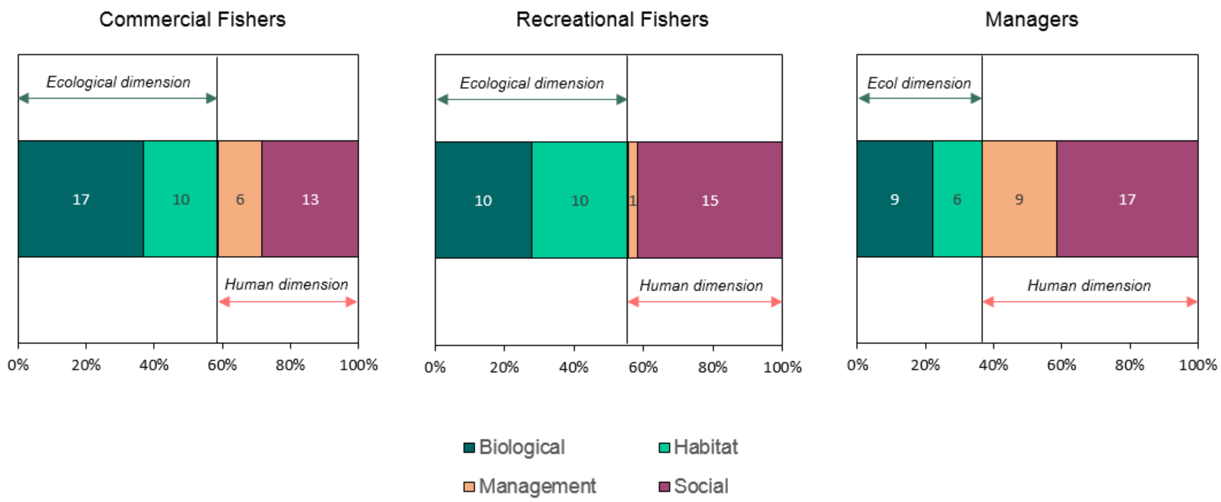


Fig. S9. The frequency and the relative percentage of each category of system concepts across three stakeholder groups. Concepts used in participants' mental models are categorized into two main categories: *Ecological-dimension* and *Human-dimension*. The ecological-dimension is divided into two sub-categories of biological concepts and habitat-related concepts. The human-dimension is divided into two sub-categories of social concepts and management related concepts. We measured the frequency and relative percentage of each sub-category across stakeholder types to determine stakeholder-specific biases. The numbers on bar-graphs indicate the frequency of concepts under each specific category. The x-axis shows the relative percentage.

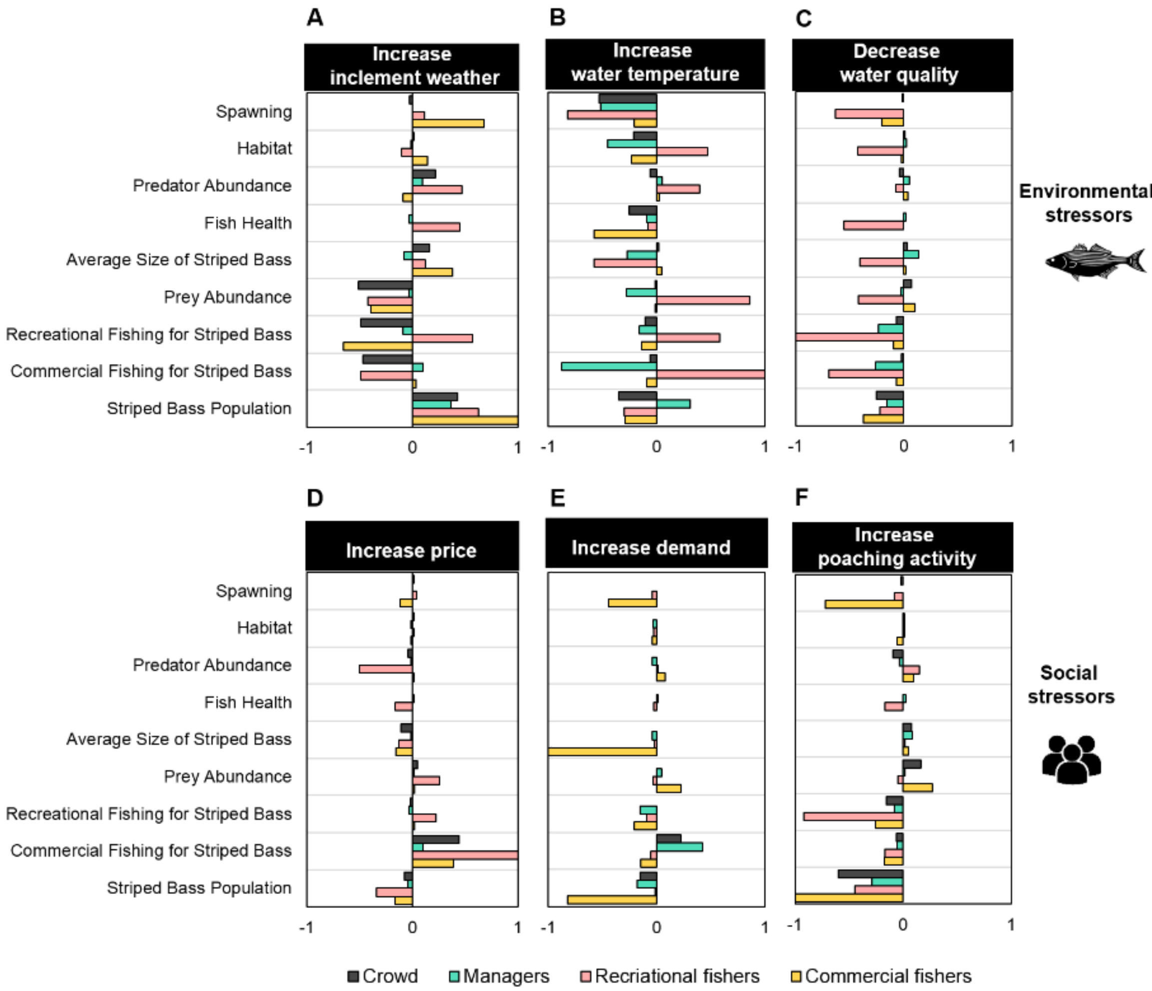


Fig. S10. Relative normalized changes in the value of concepts, simulating the striped bass fishery response to six different scenarios: (*A*) increased inclement weather for fishing, (*B*) increased water temperature, (*C*) decreased water quality, (*D*) increased price of fish, (*E*) increased demand, and (*F*) increased poaching and illegal catches. These scenarios simulate the impacts of social and environmental stressors and were used in experts' subjective evaluations to judge the models' performance in terms of dynamic behavior.

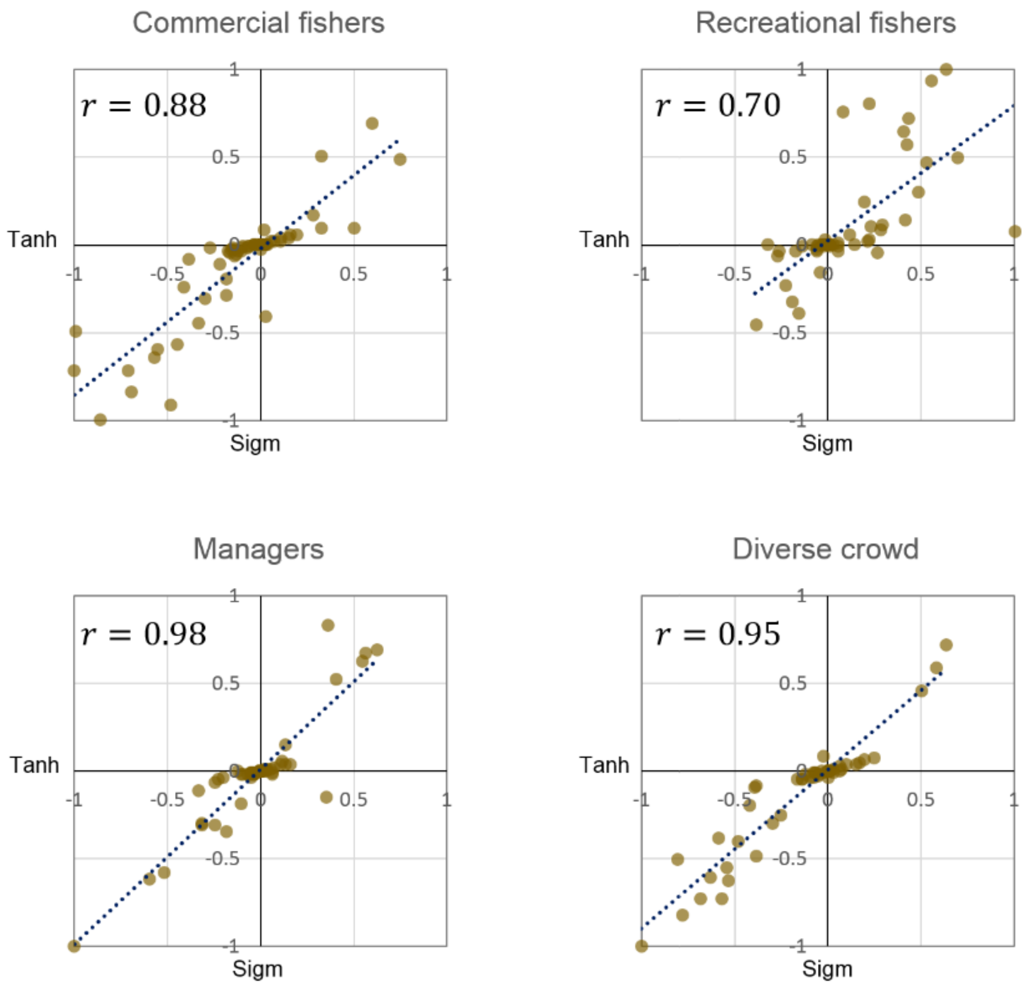


Fig. S11. Pearson Correlation between scenario outcomes produced by sigmoid (x-axis) and hyperbolic tangent (y-axis) as the activation functions. For all four models, outcomes of 6 scenarios (see Fig. S10) produced by sigmoid and hyperbolic tangent functions are significantly correlated at 0.01 level.

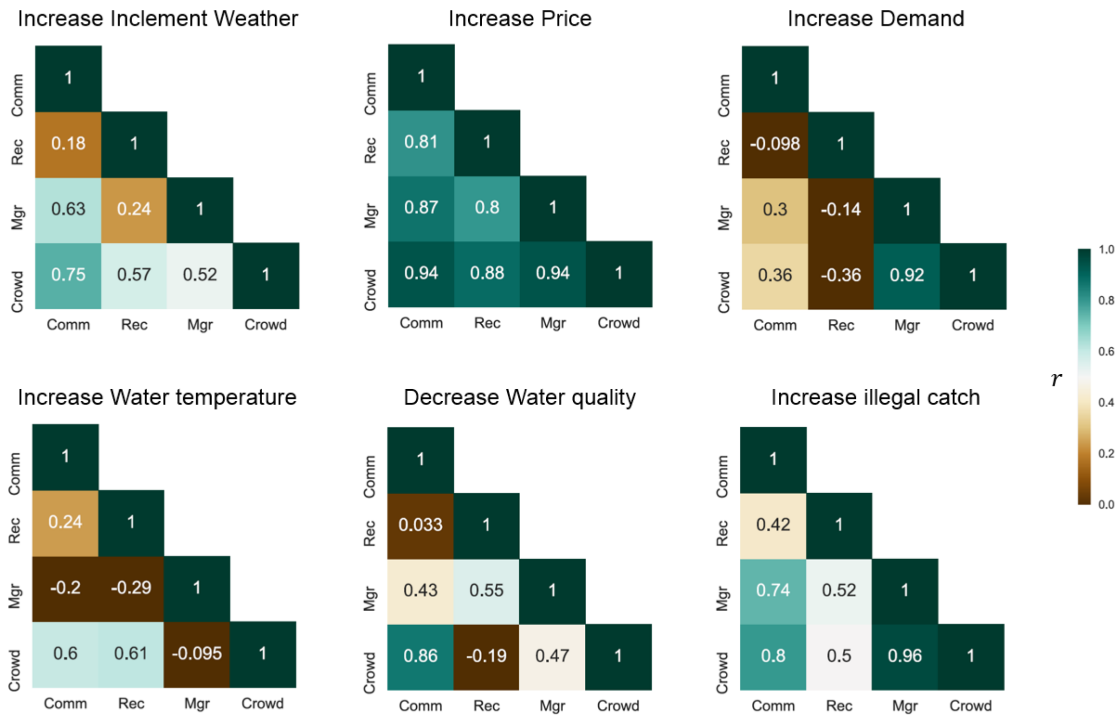


Fig. S12. Correlation matrices of scenario outcomes between four models for 6 scenarios. Each cell represents Pearson correlation coefficient ($\boldsymbol{r}$) between a pair of stakeholder groups. Scenario outcomes for "*increase price*", "*increase illegal catch*", and "*increase inclement weather*" are more positively correlated across groups while the outcomes of scenarios "*increase water temperature*", "*increase demand*", and "*decrease water quality*" are less correlated with some pairs demonstrating negative correlations. Among 6 scenarios, there are no cases where the crowd model strongly contrasts with all three groups.

Structure of the Models

	Question	Very Inaccurate	Inaccurate	Somewhat Inaccurate	Neutral	Somewhat Accurate	Accurate	Very Accurate
		1	2	3	4	5	6	7
1	Relationships between **Prey-SB Population-Predator**							
2	Relationships between **Com Fishing-SB Population-Rec Fishing**							
3	Relationships between **Prey-SB Population-Predator** & **Com Fishing-SB Population-Rec Fishing**							
4	Relationships between **Prey-SB Population-Predator** & **Com Fishing-SB Population-Rec Fishing** & **Average Size of SB**							
5	Relationships between **Spawning** and SB Population Dynamics							
6	Relationships between **Water Temperature** and other concepts in the model							
7	Relationships between **Inclement Weather** and other concepts in the model							
8	Relationships between **Habitat** and other concepts in the model							
9	Relationships between **Poaching and illegal activity** and other concepts in the model							
10	Relationships between **Fish Health** and other concepts in the model							
11	Relationships between **Price, Demand** and other concepts							

Fig. S13. Evaluation sheet used at interviews with experts to evaluate the structure of the models. Each expert had to fill out the sheet for 4 blinded models.

Dynamics of the Models

		Very Inaccurate	Inaccurate	Somewhat Inaccurate	Neutral	Somewhat Accurate	Accurate	Very Accurate
	Question	1	2	3	4	5	6	7
1	Scenario outcome of **Increasing water temperature**							
2	Scenario outcome of **Decreasing water quality**							
3	Scenario outcome of **Increasing inclement weather for fishing**							
4	Scenario outcome of **Increasing Price**							
5	Scenario outcome of **Increasing Demand**							
6	Scenario outcome of **Increasing poaching and illegal activities**							
7	**Add Your Scenario**							
8	**Add Your Scenario**							
9	**Add Your Scenario**							

Fig. S14. Evaluation sheet used at interviews with experts to evaluate the dynamic behavior of the models. Each expert had to fill out the sheet for 4 blinded models.

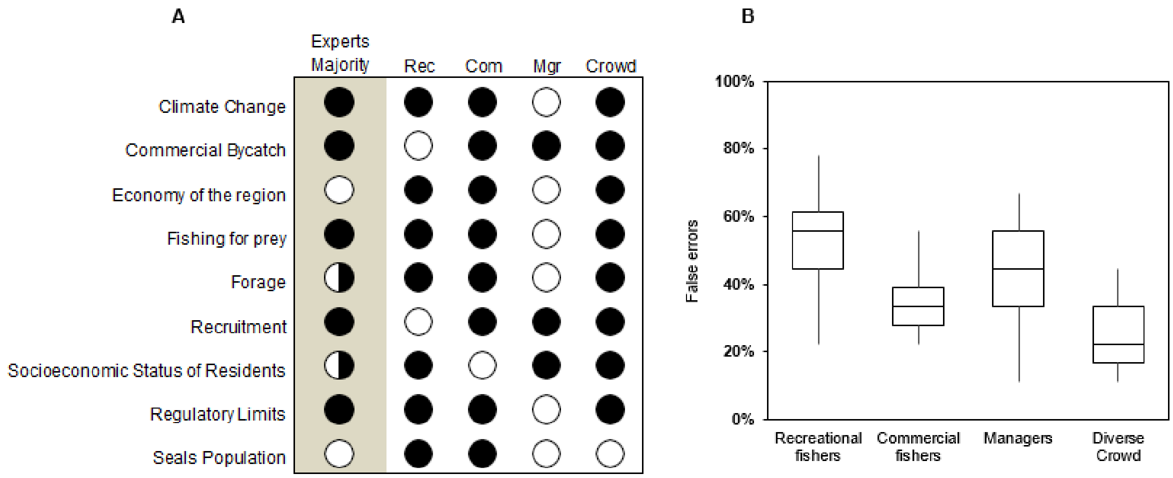


Fig. S15. Expert evaluation of models' components (i.e. nodes). Evaluation was conducted based on experts' scientific knowledge of the system's fundamental components versus trivial or redundant components which were considered superfluous in modeling the striped bass population. One example of an opinion table in (*A*) shows experts' majority opinion about whether a component is necessary (black), superfluous (white) or there is no consensus among experts (half-black, half-white). The percent of false errors according to the experts' majority opinion is shown in (*B*).

List of concepts used in Models

Concepts	**Model #1**	**Model #2**	**Model #3**	**Model #4**	Important must Included	Not Important must Excluded	I don't Know
Climate Change							
Commercial Bycatch							
Economy of the region							
Fishing for prey							
Forage (Available food)							
Recruitment of Striped Bass							
Socioeconomic Status of Residents							
Regulatory Limits							
Seals Population							

Fig. S16. Evaluation sheet used at interviews with experts to evaluate false errors regarding system components that were used by more than one but not all stakeholder groups. Each expert had to fill out the sheet for 4 blinded models.

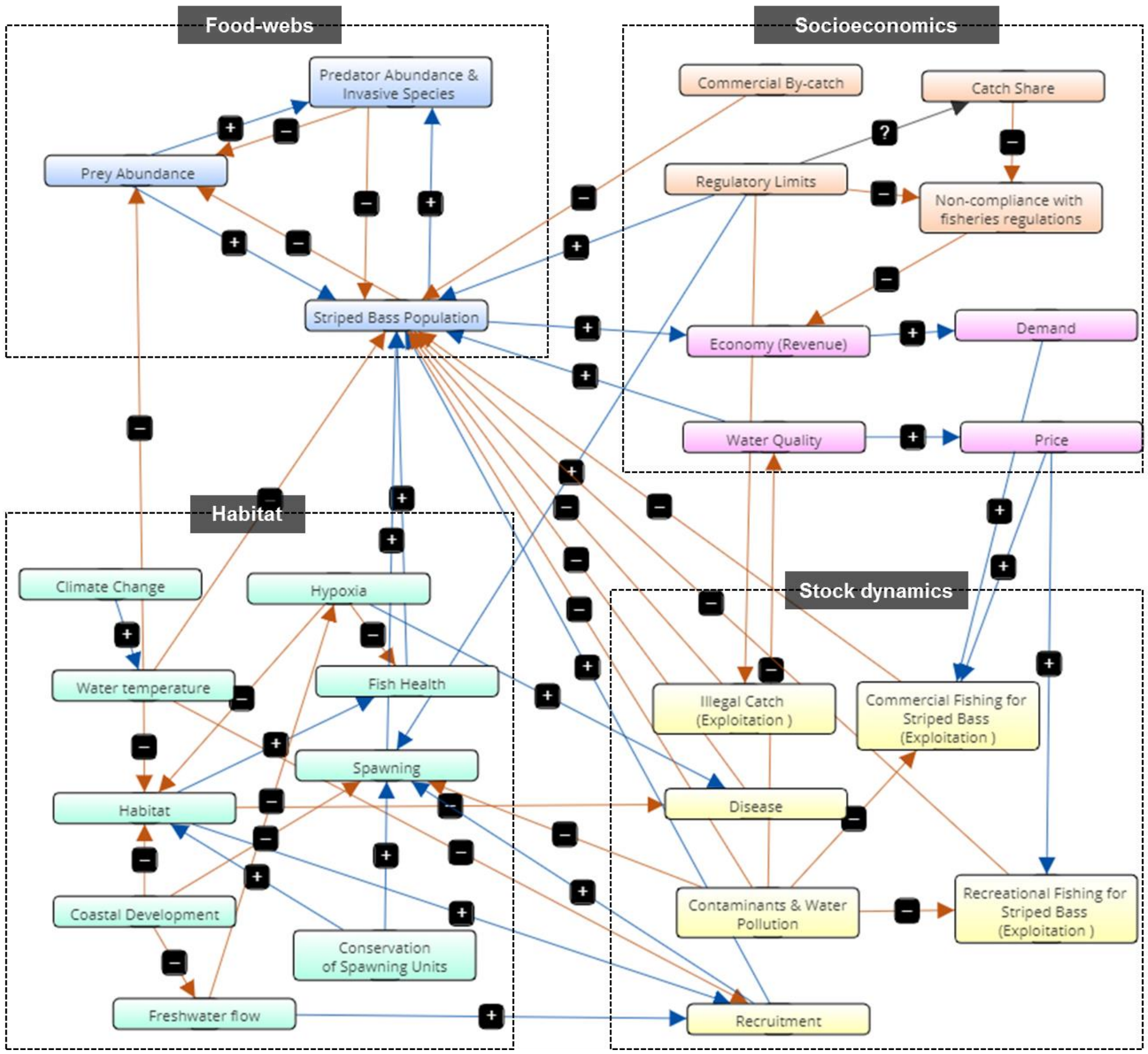


Fig. S17. Ecosystem-Based Fishery Management (EBFM) model. This striped bass EBFM model has four critical ecosystem considerations (i.e., sub-models): *Habitat suitability*, *Stock dynamics*, *Food-webs*, and *Socioeconomics*, each illustrated by a unique color. This model is developed base on a published report by EBFM Striped Bass Species Team at Maryland Sea Grant (see ref. (29)) and we graphically re-produced it using *Mental Modeler* tool (www.mentalmodeler.org).

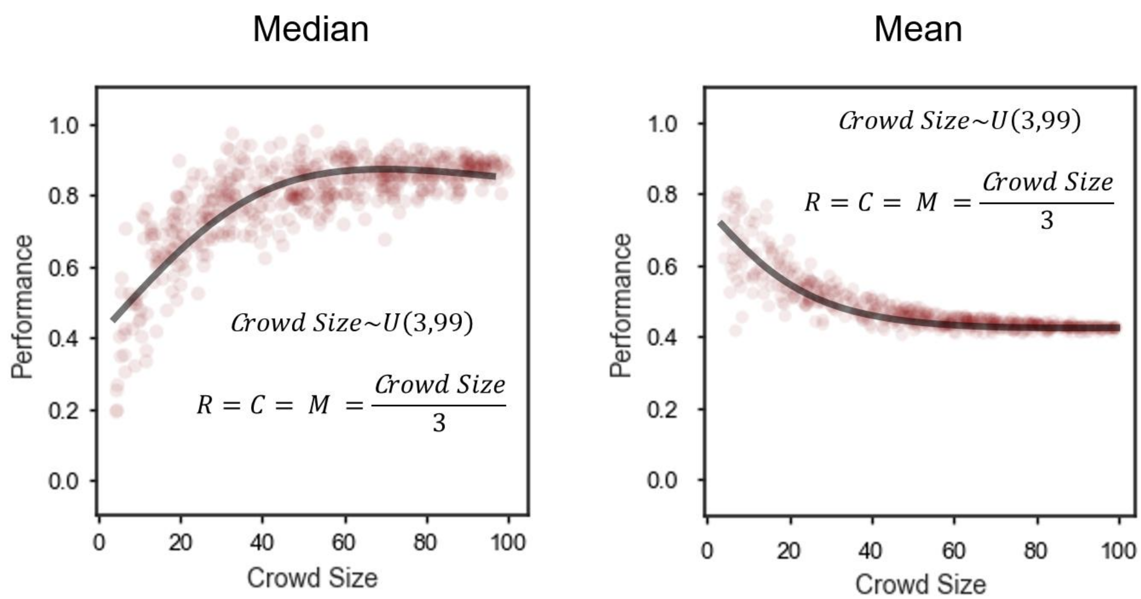
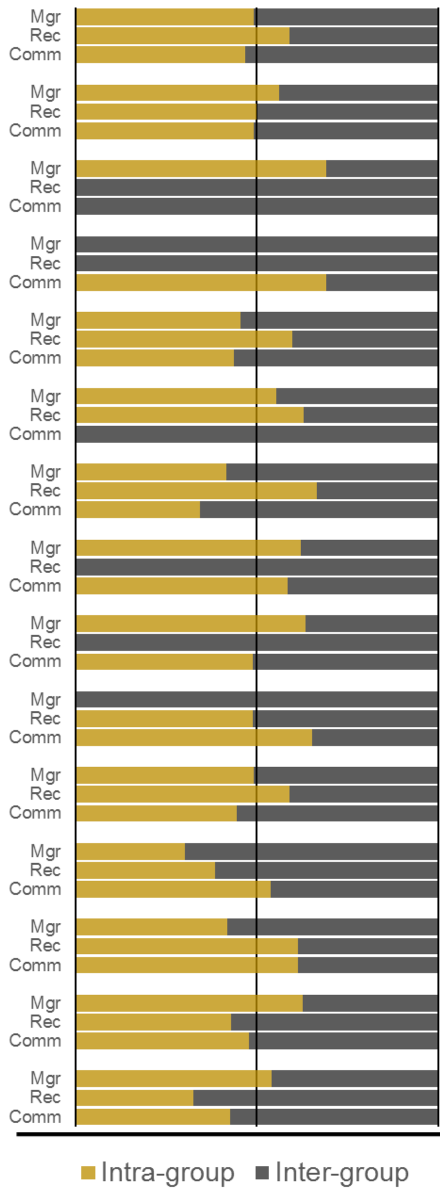


Fig. S18. Comparison of the performance of virtual crowds generated by *median* and *mean* mechanisms while the crowd size increases. Results in this figure illustrate 10,000 replicates for generating virtual crowds with a random size (i.e., the crowd size in each replicate is a random number between 3 and 99 drawn from a uniform distribution), such that the number of agents from types recreational fisher ($\boldsymbol{R}$), commercial fisher ($\boldsymbol{C}$), and manager ($\boldsymbol{M}$) are the same. Unlike the *median* mechanism, crowd models generated by a simple-averaging method (i.e. *mean*) disproportionately cumulate any information from diverse types of stakeholders without any mechanisms for filtering out information for which no majority agreement (i.e. concurrence) exists. In contrast, even though using the *median* method to aggregate models across diverse groups of stakeholders may cause information loss (i.e., filtering out information for which no clear agreement exists), it produces a parsimonious model that more accurately represents system complexities and interdependences.

Table S1. The number of participants from each stakeholder type and the number of nodes and connections used in their mental models. The mean and standard deviation of number of concepts (i.e. nodes) and connections (i.e. edges) are shown by stakeholder types.

Stakeholder group	Number of Participants	Number of Nodes (*N*) Mean (SD)	Number of Connections (*C*) Mean (SD)
Recreational fishers	13	11.54 (4.01)	29.85 (20.53)
Commercial fishers	11	11.45 (2.84)	23.45 (12.41)
Fisheries managers	8	12.00 (3.21)	27.25 (7.87)
Total	32	11.63 (3.35)	27.00 (15.32)

Table S2. Intra-group versus inter-group variations based on the ratio of disagreement to agreement ($\boldsymbol{\gamma}$) for three stakeholder groups and for the major causal relationships. Commercial fishers, recreational fishers, and managers in 67%, 60%, and 53% of reviewed relationships, respectively, demonstrate higher agreement within group compared to across groups. In 80% of cases, there are at least two groups for which the *inter-group* $\boldsymbol{\gamma}$ (dark gray color bar) is higher than *intra-group* $\boldsymbol{\gamma}$ (mustard yellow color bar), suggesting that stakeholders within each group constructs more similar maps compared to stakeholders across groups.



$\gamma = \frac{Disagree}{Agree}$	From	To
Mgr / Rec / Comm	Predator abundance	Striped bass population
Mgr / Rec / Comm	Striped bass population	Predator abundance
Mgr / Rec / Comm	Demand	Commercial fishing
Mgr / Rec / Comm	Water temperature	Fish health
Mgr / Rec / Comm	Water temperature	Habitat
Mgr / Rec / Comm	Water temperature	Spawning
Mgr / Rec / Comm	Water temperature	Prey abundance
Mgr / Rec / Comm	Water quality	Striped bass population
Mgr / Rec / Comm	Fish health	Striped bass population
Mgr / Rec / Comm	Habitat	Striped bass population
Mgr / Rec / Comm	Spawning	Striped bass population
Mgr / Rec / Comm	Prey abundance	Striped bass population
Mgr / Rec / Comm	Water temperature	Striped bass population
Mgr / Rec / Comm	Recreational fishing	Striped bass population
Mgr / Rec / Comm	Commercial fishing	Striped bass population

SI References

1. P. N. Johnson-Laird, *Mental models: Towards a cognitive science of language, inference, and consciousness* (Harvard University Press, 1983).
2. S. Mohammed, B. C. Dumville, Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *J. Organ. Behav. Int. J. Ind. Occup. Organ. Psychol. Behav.* **22**, 89–106 (2001).
3. C. Eden, On the nature of cognitive maps. *J. Manag. Stud.* **29**, 261–265 (1992).
4. B. Kosko, Fuzzy cognitive maps. *Int. J. Man. Mach. Stud.* **24**, 65–75 (1986).
5. S. Gray, J. Hilsberg, A. McFall, R. Arlinghaus, The structure and function of angler mental models about fish population ecology: The influence of specialization and target species. *J. Outdoor Recreat. Tour.* **12**, 1–13 (2015).
6. A. C. Stier, *et al.*, Integrating expert perceptions into food web conservation and management. *Conserv. Lett.* **10**, 67–76 (2017).
7. S. A. Gray, S. Gray, L. J. Cox, S. Henly-Shepard, Mental modeler: a fuzzy-logic cognitive mapping modeling tool for adaptive environmental management in *2013 46th Hawaii International Conference on System Sciences*, (2013), pp. 965–973.
8. D. Koutra, A. Parikh, A. Ramdas, J. Xiang, Algorithms for graph similarity and subgraph matching in *Proc. Ecol. Inference Conf*, (2011).
9. J. Navajas, T. Niella, G. Garbulsky, B. Bahrami, M. Sigman, Aggregated knowledge from a small number of debates outperforms the wisdom of large crowds. *Nat. Hum. Behav.* **2**, 126–132 (2018).
10. S. A. Gray, E. Zanre, S. R. J. Gray, "Fuzzy cognitive maps as representations of mental models and group beliefs" in *Fuzzy Cognitive Maps for Applied Sciences and Engineering*, (Springer, 2014), pp. 29–48.
11. A. B. Kao, I. D. Couzin, Modular structure within groups causes information loss but can improve decision accuracy. *Philos. Trans. R. Soc. B Biol. Sci.* **374**, 20180378 (2019).
12. P. Aminpour, *et al.*, Wisdom of stakeholder crowds in complex social--ecological systems. *Nat. Sustain.*, 1–9 (2020).
13. A. Voinov, *et al.*, Tools and methods in participatory modeling: Selecting the right tool for the job. *Environ. Model. Softw.* **109**, 232–255 (2018).
14. C. D. Stylios, P. P. Groumpos, Modeling complex systems using fuzzy cognitive maps. *IEEE Trans. Syst. Man, Cybern. A Syst. Humans* **34**, 155–162 (2004).
15. A. J. Jetter, R. C. Sperry, Fuzzy cognitive maps for product planning: using stakeholder knowledge to achieve corporate responsibility in *2013 46th Hawaii International Conference on System Sciences*, (2013), pp. 925–934.
16. A. Jetter, W. Schweinfort, Building scenarios with Fuzzy Cognitive Maps: An exploratory study of solar energy. *Futures* **43**, 52–66 (2011).
17. P. J. Giabbanelli, S. A. Gray, P. Aminpour, Combining fuzzy cognitive maps with agent-based modeling: Frameworks and pitfalls of a powerful hybrid modeling approach to understand human-environment interactions. *Environ. Model. Softw.* **95**, 320–325 (2017).
18. E. I. Papageorgiou, J. L. Salmeron, "Methods and algorithms for fuzzy cognitive map-based modeling" in *Fuzzy Cognitive Maps for Applied Sciences and Engineering*, (Springer, 2014), pp. 1–28.
19. P. P. Groumpos, C. D. Stylios, Modelling supervisory control systems using fuzzy cognitive maps. *Chaos, Solitons & Fractals* **11**, 329–336 (2000).
20. I. Á. Harmati, L. T. Kóczy, Notes on the Dynamics of Hyperbolic Tangent Fuzzy Cognitive Maps in *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, (2019), pp. 1–6.
21. E. I. Papageorgiou, J. L. Salmeron, A review of fuzzy cognitive maps research during the last decade. *IEEE Trans. Fuzzy Syst.* **21**, 66–79 (2012).
22. S. Bueno, J. L. Salmeron, Benchmarking main activation functions in fuzzy cognitive maps. *Expert Syst. Appl.* **36**, 5221–5229 (2009).
23. J. Kneebone, W. S. Hoffman, M. J. Dean, D. A. Fox, M. P. Armstrong, Movement patterns and stock composition of adult Striped Bass tagged in Massachusetts coastal waters. *Trans. Am. Fish. Soc.* **143**, 1115–1129 (2014).

24. H. B. Bigelow, W. C. Schroeder, *Fishes of the Gulf of Maine* (US Government Printing Office, 1953).
25. G. A. Nelson, B. C. Chase, J. Stockwell, Food habits of striped bass (Morone saxatilis) in coastal waters of Massachusetts. *J. Northwest Atl. Fish. Sci.* **32** (2003).
26. G. A. Nelson, B. C. Chase, J. D. Stockwell, Population consumption of fish and invertebrate prey by striped bass (Morone saxatilis) from coastal waters of northern Massachusetts, USA. *J. Northw. Atl. Fish. Sci* **36**, 111–126 (2006).
27. R. Murphy Jr, S. Scyphers, S. Gray, J. H. Grabowski, Angler attitudes explain disparate behavioral reactions to fishery regulations. *Fisheries* **44**, 475–487 (2019).
28. Atlantic States Marine Fisheries Commission, "Addendum IV to Amendment 6 to the Atlantic Striped Bass Interstate Fishery Management Plan" (2014).
29. , "Ecosystem-Based Fisheries Management for Chesapeake Bay: Striped Bass Background and Issue Briefs." (2009).
30. C. E. Shannon, A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).