Classifying fishing behavioral diversity using high-frequency movement data

Shay O'Farrell[a,1,2], Iliana Chollett[b,1], James N. Sanchirico[a,c], and Larry Perruso[d]

[a]Department of Environmental Science and Policy, University of California, Davis, CA 95616; [b]Private address, Accony, Louisburgh, Co Mayo F28 A275, Ireland; [c]Resources for the Future, Washington, DC 20036; and [d]Southeast Fisheries Science Center, National Oceanic and Atmospheric Administration Fisheries, Miami, FL 33149

Edited by Bonnie J. McCay, Rutgers University, New Brunswick, NJ, and approved July 12, 2019 (received for review April 19, 2019)

Effective management of social-ecological systems (SESs) requires an understanding of human behavior. In many SESs, there are hundreds of agents or more interacting with governance and regulatory institutions, driving management outcomes through collective behavior. Agents in these systems often display consistent behavioral characteristics over time that can help reduce the dimensionality of SES data by enabling the assignment of types. Typologies of resource-user behavior both enrich our knowledge of user cultures and provide critical information for management. Here, we develop a data-driven framework to identify resource-user typologies in SESs with high-dimensional data. To demonstrate policy applications, we apply the framework to a tightly coupled SES, commercial fishing. We leverage large fisheries-dependent datasets that include mandatory vessel logbooks, observer datasets, and high-resolution geospatial vessel tracking technologies. We first quantify vessel and behavioral characteristics using data that encode fishers' spatial decisions and behaviors. We then use clustering to classify these characteristics into discrete fishing behavioral types (FBTs), determining that 3 types emerge in our case study. Finally, we investigate how a series of disturbances applied selection pressure on these FBTs, causing the disproportionate loss of one group. Our framework not only provides an efficient and unbiased method for identifying FBTs in near real time, but it can also improve management outcomes by enabling ex ante investigation of the consequences of disturbances such as policy actions.

fisheries | human mobility | movement ecology | natural resource management | resilience

Better understanding of human behavior is now an important component of the natural-resource management toolbox (1). The big data revolution that is taking place across multiple social systems and scales holds great promise for incorporating human behavioral data into management. Fisheries in particular have seen enormous growth in the routine gathering and curating of management-relevant data, allowing analysts to measure and understand behaviors at spatial and temporal scales that were previously unthinkable (2–4). A major challenge in the brave new world of big data is how to collapse complex and burgeoning data streams into useable information. In natural resource management, one solution to this dilemma is offered by the behavioral typologies paradigm.

As with many natural resource users, fishers often exhibit consistent behavioral types (5–9), referred variously as personality types, fishing strategies, or fishing styles (Fig. 1). The diversity of types is attributed to personality and character, but also to factors such as vessel characteristics (6). Much research has been devoted to understanding and classifying the behavior of fishers based on sociodemographic and psychological characteristics, and to explaining why fishers make different choices in similar circumstances and how they are likely to respond to disturbances and management interventions (5–9). Typologies are simplifications of complex human behaviors and, as with all summaries, some information loss is inevitable. However, the fishing typology concept allows high-dimensional data to be collapsed into tractable units, which facilitates incorporating fleet-wide behavioral information into management and policy analyses. To date, research on FBTs has been based largely on interviews, which provide valuable data but require considerable resources to obtain and update, and they rely on self-reported information. We offer a complementary approach, whereby fisheries-dependent data that are routinely collected by many agencies can be used in a framework to characterize discrete FBTs using data science tools. Defining FBTs using fisheries' big data offers a generalizable, transferable, and cost-effective approach that can be deployed across entire fleets, arming policymakers, planners, and resource managers with near-real-time behavioral representation of a given fishery to aid in decision making and socioeconomic monitoring.

Many descriptors of fishing behaviors employed in previous research (Fig. 1) can be quantified using extant fisheries-dependent data, which may include trip logbooks and fisheries observer program reports, as well as linked geospatial records from vessel monitoring systems (VMSs). Logbook data have been collected by many fisheries agencies for decades, usually providing trip-level catch and effort data, and often supplemented with economic information such as market prices and fuel costs. Logbooks often record fishing location data too, but these are commonly reported in coarse statistical areas. Recently, the development and deployment of VMS infrastructure has allowed automated tracking of vessel positions. Linking effort, landings, and economic information collated in logbooks to the geospatial information in the VMS datasets not only allows finer-scale analysis of fishing location data for stock assessment (10), but promises new insights when movements are considered as the spatial manifestation of fishers' decisions and behaviors. In the growing field of movement ecology,

Significance

Effective fisheries management is needed to rebuild overfished stocks and prevent future overfishing, and doing so requires an understanding of fishers' behavior. We offer an approach where "big data" routinely collected by many fisheries agencies can be used in a data-driven framework to classify fishers into discrete behavioral types, refining the *métier* concept and facilitating the inclusion of behavioral information into near-real-time fisheries management.

Author contributions: S.O., I.C., and J.N.S. designed research; S.O. and I.C. performed research; S.O., I.C., and L.P. contributed new reagents/analytic tools; S.O. and I.C. analyzed data; and S.O., I.C., J.N.S., and L.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]S.O. and I.C. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: shay.ofarrell.ac@gmail.com.

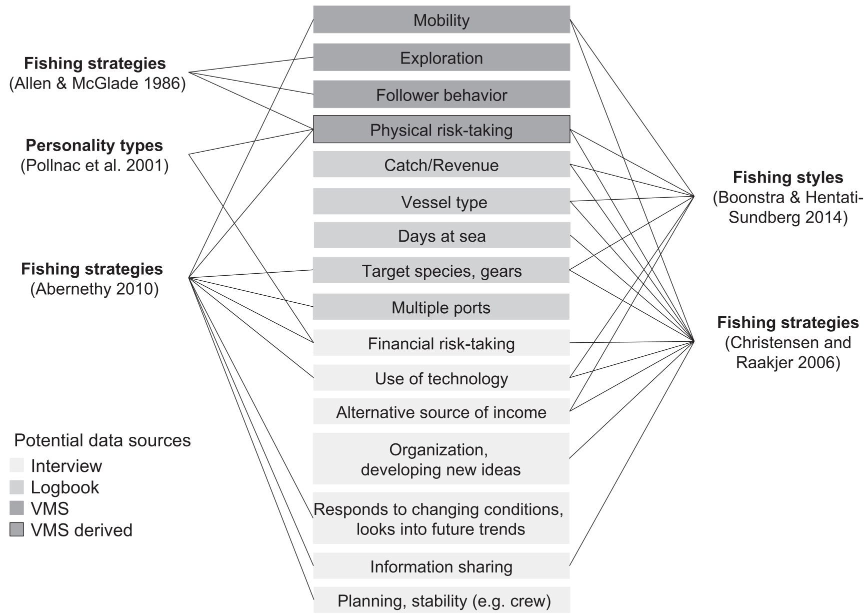

Fig. 1. Characteristics used in the literature to define fishing behavioral types (also known as fishing strategies, personality types, or styles) based on interview data. Darker shading indicates additional data sources that could be used to quantify and refine these characteristics using near-real-time data-driven approaches.

high-resolution tracking data are now successfully being used to characterize animal behaviors and even personalities (11). Although vessel movement paths have to date been used mostly to map activities (e.g., fishing, searching, and cruising, ref. 12) or characterize fishing practices (13), these data also provide insights to spatial behaviors, such as the tendency to explore new grounds (14) or to fish in rough seas, indicating tolerance to risk (15, 16).

Leveraging existing data for the bottom longline fleet participating in the Gulf of Mexico Grouper-Tilefish fishery, we develop an approach to quantifying behavioral and vessel characteristics in fisheries-dependent data and then use clustering to classify these into discrete FBTs. We find that clear and intuitive types emerge from the data and that these agree well with typologies derived from interview data in other studies (5, 7, 17). We go on to demonstrate a management application of our framework by investigating how the FBTs in our case study responded to a series of disturbances that took place in rapid succession. These include the BP *Deepwater Horizon* oil spill, an emergency closure triggered by the Endangered Species Act, the introduction of individual fishing quotas, and a performance-related reduction of the fleet. Preemptive application of our framework would allow managers to explore consequences of many policy interventions prior to implementation.

Results

We quantified 9 characteristics to describe each vessel: mobility, exploration, physical risk taking, revenue (mean and SD), vessel type, days at sea (mean and SD), and use of multiple ports. Our analyses incorporated all of the behavioral characteristics that could be quantified using the fisheries-dependent data listed in Fig. 1, with the exception of: 1) target species and gear, which is less relevant within a species-and-gear-defined métier such as the Grouper-Tilefish longline fleet, and 2) follower behavior, which at this point in time was not possible to quantify (see *Materials and Methods*). Use of multiple ports was found to be highly correlated with mobility and was dropped to avoid skewing the clusters.

Once quantified across the fleet, the vessel characteristics underwent clustering (or unsupervised classification, ref. 18), resulting in 3 well-defined behavioral types (Fig. 2). The characteristics most strongly defining the clusters were exploration (20%), mobility (18%), and variability both in days at sea (16%) and revenue (14%, Fig. 3), with the 3 emergent FBTs being characterized by consistent combinations of these characteristics. FBT1 are fishers with low mobility and less explorative behavior who are risk averse and carry out short trips. Conversely, FBT2 and FBT3 have high mobility and more explorative behavior, are more risk tolerant, and conduct longer trips. These latter groups differ in that FBT3 have higher variability in trip duration and revenue than FBT2. FBT1 coincides well with what Allen and McGlade (5) call "cartesians," with FBT2/FBT3 being typical "stochasts." While cartesians fish in the same locations and are risk averse, stochasts are more risk tolerant and explore with the intention of improving catches (5).

There was a clear change in behavioral types after disturbance. FBT3 vessels mostly exited the fishery (50%), and most of those that remained evolved to become FBT1 or FBT2 (47%). A large proportion of FBT1 vessels exited the fishery (58%), but most vessels that remained active did so by remaining in their own group (32%). FBT2 was the most resilient group to disturbance, with vessels largely remaining in the group after the disturbance (48%: Fig. 4).

Discussion

The empirical study of fisher behavior is an underdeveloped area of research (19). This is partially due to the difficulty and expense of obtaining behavioral information and collapsing rich but complex data into information that is tractable for fleet-level analyses (20). By offering a data-driven, near-real-time and cost-effective approach to complement analyses based on interviews as carried out widely within this research area, our study adds to the body of knowledge typifying fisher behavior (5–9).

The FBTs that emerge from our clustering agree well with the cartesian and stochast typologies that have been found by other authors (5, 7, 17) and our results indicate that a series of disturbances drove the selection of these typologies in the bottom longline Grouper-Tilefish fleet in the Gulf of Mexico. Fishers

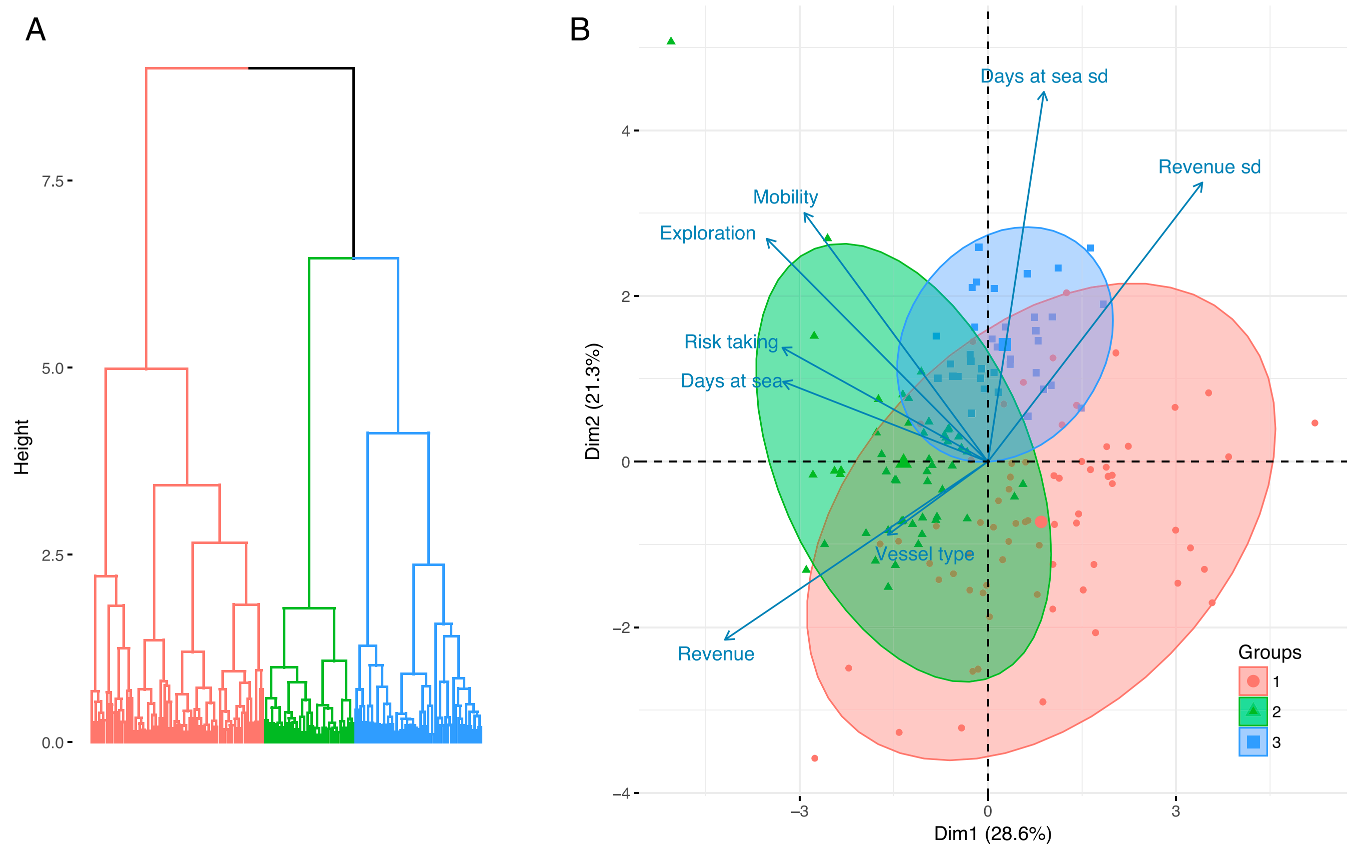

Fig. 2. Fishing behavioral types within the longline Grouper-Tilefish fishery of the Gulf of Mexico. (*A*) Dendrogram showing 3 FBTs. Ensemble analysis of clustering indices suggests the dataset is naturally defined by 3 groups. (*B*) Results of principal component analysis showing 102 vessels before and 54 vessels after the disturbance. Labeled arrows indicate the direction of influence of each characteristic on the 2D ordination plot.

with high variability in trip duration and revenue were more vulnerable to the disturbances and were more likely to exit the fishery. The vulnerability of inconsistent fishers to disturbances is in line with other work showing that increased income variability can make it difficult for individuals to remain in a fishery (21). The resilience of stochasts to disturbances in our dataset is also in line with work showing that investment in exploration may buffer against shocks (14). The agreement of these findings with the literature indicates that the simplified typologies in our FBT framework succeed in characterizing at least part of the behavioral dynamics of a complex social-ecological system experiencing a series of shocks. Approaches based on observed variables such as the one used here serve as tools to assess behavioral phenomena by describing complex information with quantitative variables and condensing it into categories that can be easily implemented and are informative to natural resource managers (22). Human behavior is, however, complex, and is influenced by several factors. Any given set of observed behaviors may be underpinned by numerous drivers, including the values, perceptions, emotions, and motivations of fishers. A deeper understanding of the underlying drivers of behaviors could be useful for management, although this information is difficult to obtain (6). For example, a fisher's likelihood to comply with regulations is motivated by a complex mix of social, economic, environmental, and political factors. Even though many fishers who are observed to be more compliant may exhibit similar scores along a given axis, such as attitude to risk, this does not mean that all fishers with similar attitudes to risk are necessarily more compliant. As a result, management actions that act on either a single driver or behavior are at risk of producing unintended outcomes (23, 24).

There is a large number of clustering methods available, and the selection of the method can influence the result if the clusters are not well separated, as is often the case with "real world" data (18). To deal with this issue, the results of cluster analyses can often be validated internally or externally (25). For example, the consistency of the results can be validated internally using the original data by iteratively resampling the dataset at random, reclustering, and verifying that the same clusters are found for the different subsets (25). Clusters can also be validated externally by analyzing the same dataset with multiple clustering methods and comparing the results, either informally using visual examination or formally calculating consensus indices (26). Researchers in the social sciences have also dealt with this issue by validating the clusters against independent datasets, selecting examples from each cluster and performing interviews: Agreement between independent sources ensures validity and reliability of the clustering procedure (22, 27, 28).

One of the disturbances affecting the Gulf of Mexico (GoM) Grouper-Tilefish fishery was a 2010 endorsement that removed lower-performing vessels from a fishery that was considered to be overcapitalized, whereby vessels that had a history of landing less than 40,000 lbs of grouper-tilefish per annum were excluded from the bottom longline sector. In our results, it is interesting that lower-performing vessels clustered into 2 discrete types, FBT1 and FBT3. Average performance (mean revenue) was not significantly different between these FBTs (Fig. 3) and the longline endorsement likely removed vessels from both clusters. Indeed, mean revenue ranked only 6th out of 8 variables in terms of its importance to the clustering (Fig. 3). Although performance (catch or revenue) is commonly used to define fishing typologies (6, 7), it is plausible that performance actually arises

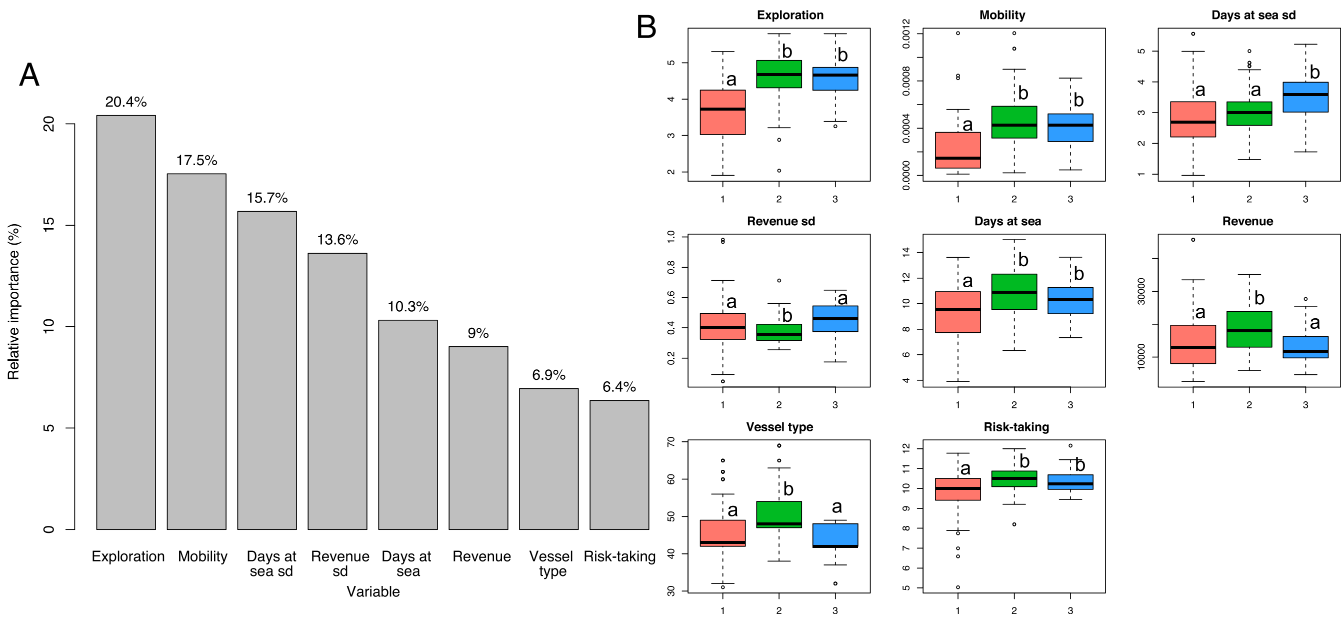

Fig. 3. Characteristics defining each fishing behavioral type. (*A*) Relative importance of characteristics. (*B*) Characteristics of the fleet grouped by behavioral type. In boxplots, lines represent medians, boxes 25% and 75% quantiles, whiskers 1.5 interquartile ranges, and dots outliers. Letters indicate significantly different groups ($P \leq 0.05$) according to Tukey's honest significant difference test, where groups labeled with the same letter, *a* or *b*, are not significantly different to each other.

from other characteristics rather than being a characteristic in itself. Although our performance variable (mean revenue) was not strongly correlated with any other individual clustering variable, future work will test whether performance can be predicted from more complex relationships involving multiple characteristics.

The study of fishing behavioral types is also useful for targeting outreach and stakeholder engagement to management interventions such as spatial closures, temporal closures, quotas, time limits, or gear restrictions. For example, in our case study, the introduction of individual quotas that eliminate the race to fish and spread fishing over a longer season may have greater impacts on risk-tolerant fishers (6), such as our FB2 group. On the other hand, area-based management might have a different impact on each of our 3 groups, such as the more mobile and exploratory vessels in our FB1 and FB2 groups (15, 29, 30).

In addition to enabling the generation of near-real-time information on behavioral dynamics within and across fisheries, our framework offers a number of further uses. As a complement to traditional questionnaire-based data gathering that is common in fisheries science, our approach will allow survey designers to stratify future sampling so that the range of participants included in a survey is representative of the range of FBTs observed in that fishery (22, 28). These FBTs themselves may subsequently be enriched by merging with the questionnaire-derived data. FBTs will also aid in the interpretation of existing and future surveys. For example, a 2015 survey of participants in the GoM Grouper-Tilefish IFQ program (31) found that a majority of participants "disagreed strongly" with statements to the effect that the introduction of individual fishing quotas (IFQs) improved the profitability of the grouper-tilefish (G-T) component of their business by increasing either ex-vessel prices or reducing operating expenses. Assigning FBTs to the participants would allow researchers to ask whether fishers belonging to high-performing typologies such as FBT2 may be less supportive of IFQs as they had been performing well in derby fishing conditions.

Currently 3 levels of information are generally used in fisheries science to classify fishing activity: fleet, fishery, and métier (32), and changes in this information are then used to assess changes

A

		POST-disturbance			
		EXIT	1	2	3
PRE disturbance	1	58.5	31.7	9.8	0
	2	25.9	22.2	48.1	3.7
	3	50.0	26.5	20.6	2.9

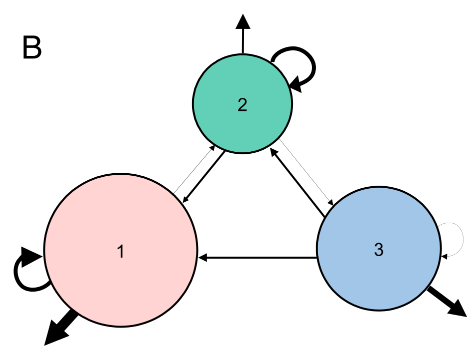

Fig. 4. Transition between behavioral types after disturbance. (*A*) Transition matrix showing the proportion of vessels changing from one behavioral type to another. (*B*) Transition graph, where the size of the circles indicates membership of each behavioral type before disturbance, and the arrows indicate whether vessels remain within the same group (feedback arrows), change group (arrows connecting different FBTs), or exit the fishery (outgoing arrows). Thicker arrows indicate greater proportions transitioning.

in fleets over time (e.g., refs. 33 and 34). Adding simplified behavioral information, such as FBTs, as an additional layer of information would underpin a more complete characterization of the fishery (6, 33) and would aid in policy analyses such as scenario planning and management tasks such as socioeconomic monitoring. Our approach is data-driven, cost-effective, does not require collecting additional data, is replicable in time and transferable to other locations, and can produce easily interpreted outputs for swift inclusion into management and policy (9). Including fisher behavior into routine fisheries analyses can help produce tailored management strategies, or even predict which groups are likely to respond negatively to a management intervention. Doing so would likely improve management success, increase compliance, and reduce enforcement costs (9).

Materials and Methods

Datasets. The majority of longline vessels engaged in the GoM Grouper-Tilefish fishery are converted shrimp trawlers, ranging in length from 10 to 22 m within our dataset. The longline gear consists of a line of baited hooks deployed along the seabed in reef areas of the Gulf of Mexico, which remains in position for a number of hours before being recovered, at which point the fish are removed, the hooks are rebaited, and the line is redeployed (35). This fishery submits logbooks to the Southeast Fisheries Science Center (SEFSC) in Florida, with VMS data being collected by the National Oceanic and Atmospheric Administration (NOAA) Office of Law Enforcement and observer data being collected by the SEFSC office in Galveston, Texas.

All commercial fishing vessels owning a federal reef fish permit are required to complete trip reports on catch and effort information using the Southeast Coastal Fisheries Trip Report logbook. In the present study, a logbook dataset from 103 bottom longline vessels was analyzed.

VMS transponders sending hourly or better reports (pings) have been mandated on commercial reef-fish fishing vessels in the GoM since 2006 and were available for all vessels in the Grouper-Tilefish fishery by early 2007. Each ping consists of the current latitude and longitude of a vessel along with a timestamp, allowing vessel tracks to be mapped with high spatiotemporal resolution.

Preprocessing of the data included the deletion of pings with GPS coordinates from outside the GoM. Ping timestamps were converted to POSIX objects with coordinated universal time (UTC) time zone to match VMS data recording protocol. To derive vessel movement speeds, the interval between each ping's timestamp and the preceding timestamp was obtained, the distance between successive pings was calculated, and then speed was expressed as a linear distance over time (ms^{-1}). Derived vessel speeds above an arbitrary threshold of 20 ms^{-1} were assumed to result from errors and were deleted. Detailed information on the VMS data analysis protocol is available in the literature (2).

The disturbance experienced in the GoM was a combination of spatial closures, effort, and gear restrictions during May–October, 2009 triggered by excessive bycatch of sea turtles (36) and several spatial closures during May 2010–April 2011 triggered by the explosion of the *Deepwater Horizon* offshore drilling rig on April 20, 2010 (37). Here the bottom longline fishery was assessed 2 y before (from May 18, 2007 to May 18, 2008) and 2 y after (from April 19, 2011 to April 19, 2013) the sequence of disturbances. The dataset includes 102 vessels before and 54 vessels after the disturbance.

Input Variables. Each vessel was characterized using 9 behavioral characteristics inferred from logbooks and VMS data: vessel type, use of multiple ports, revenue (mean and SD), days at sea (mean and SD), exploration, mobility, and physical risk taking. Our analysis incorporates all of the behavioral characteristics that can be quantified using fisheries-dependent data listed in Fig. 1, with the exception of follower behavior and target species and gears. Although in principle follower behavior could be described using positional data (11, 38), the methods to do so are challenging and computationally costly, including the classification of multidimensional time series within an unknown lag window.

Vessel type, port use, revenue, and days at sea were extracted from logbook data. Vessel type was described by the length of the vessel in feet, given that longline fishing boats are otherwise relatively similar. The Shannon diversity index (39) was used as a metric to describe port use, by quantifying how many different ports and how evenly they are used by each vessel. We used both mean and SD of trip revenue and trip duration per vessel, as reported in logbooks, to describe revenue and days at sea.

Exploration, mobility, and physical risk taking were extracted from VMS data on a trip basis. Exploration was described using mean cumulative entropy for each trip. Information entropy (40) measures the predictability of a time series and has been used as a metric of exploration (14, 41). Mean home range was used as a metric to describe mobility. A home range is the area in which an individual lives and moves on a regular basis. Different behavioral attitudes toward fishing result in different use of space and home ranges. For each trip, home range was calculated using the method of minimum convex polygon, enclosing all vessels' positions with the exception of the 5% more extreme ones (42). Risk taking was described by the propensity to fish in high-wind conditions (15). To do so, wind speed data were extracted for each VMS position, and then each trip was characterized by its extreme wind conditions (95% percentile of wind speed). The Blended Sea Winds dataset was used as the source of wind data (43). The study data were obtained under a contractual agreement with the US National Marine Fisheries Service (NMFS). The agreement prevents distribution of personally identifiable information, including variables directly included in the analysis. These data are archived at NOAA's Southeast Fisheries Science Center. Researchers under a contractual agreement with NMFS can access the data provided a nondisclosure agreement is signed.

Analyses. A data-driven, objective, and repeatable approach was used to quantify behavioral characteristics, identify distinct behavioral types, and determine what variables define them. Variables found to be highly correlated, with a variance inflation factor value larger than 3 (44), were removed (port diversity was highly related to exploration). Input variables were rescaled between 0 and 1 prior to all analyses to avoid scale-dependent clustering artifacts.

To identify natural groups in the dataset, data were clustered using hierarchical clustering and the Ward method, which minimizes the total within-cluster variance. The optimal number of clusters was identified using an ensemble approach and majority vote from 26 clustering indices. Clustering results were visualized using a dendrogram and principal component analysis.

The relative importance of input variables to defining behavioral types was identified using random forest (45), whereby the vessels were labeled according to the cluster into which they fell, and the labels were then used to perform a supervised classification and determine which variables contributed most to each cluster. Variable importance was quantified as the increase in percent of times a case is misclassified when the variable is permuted (46).

To quantify changes among behavioral types before and after disturbance, a confusion matrix was built which expresses change between types as a percentage of the number of vessels in each type. This information was visualized using a transition graph, where the size of the circles is relative to the number of vessels in each type before the disturbance and the line thickness changes according to the proportion of exchanges.

All data analyses were conducted in R (ver. 3.4.3: 21). The R package rerddap (47) was used to extract wind speed data from NOAA's ERDDAP server. The package geosphere was used to calculate distances between VMS pings (48). The packages vegan (49) and adehabitatHR (42) were used to calculate port diversity and home range. Finally, NbClust (50) was used to perform clustering, factoextra (51) to identify the optimal number of clusters, and randomForest (46) to calculate the relative importance of input variables.

ACKNOWLEDGMENTS. We thank Elizabeth Scott-Denton of the Galveston Laboratory, Southeast Fisheries Science Center, NOAA Fisheries, for supplying Gulf of Mexico observer data, and Kelly Spalding (National Marine Fisheries Service VMS Program Manager, NOAA, 1315 East West Highway, Silver Spring, MD 20910) and Carlos Rivero (Southeast Fisheries Science Center, Beaufort Laboratory) for supplying VMS data. Funding for the work came from an NSF Coastal Science, Engineering and Education for Sustainability (SEES) grant awarded to J.N.S. (grant 2500-1565-00) and a NMFS subaward to J.N.S. (award 201502599-01) through the Spatial Economics Toolbox for Fisheries (FishSET). The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect those of NOAA or the US Department of Commerce.

1. C. M. Anderson *et al*., How commercial fishing effort is managed. *Fish Fish.* **20**, 268–285 (2019).
2. S. O'Farrell *et al*., Improving detection of short-duration fishing behaviour in vessel tracks by feature engineering of training data. *ICES J. Mar. Sci.* **74**, 1428–1436 (2017).
3. D. A. Kroodsma *et al*., Tracking the global footprint of fisheries. *Science* **359**, 904–908 (2018).
4. T. Russo *et al*., A model combining landings and VMS data to estimate landings by fishing ground and harbor. *Fish. Res.* **199**, 218–230 (2018).
5. P. M. Allen, J. M. McGlade, Modelling complex human systems: A fisheries example. *Eur. J. Oper. Res.* **30**, 147–167 (1987).
6. W. J. Boonstra, J. Hentati-Sundberg, Classifying fishers' behaviour. An invitation to fishing styles. *Fish Fish.* **17**, 78–100 (2016).

7. A.-S. Christensen, J. Raakjær, Fishermen's tactical and strategic decisions: A case study of Danish demersal fisheries. *Fish. Res.* **81**, 258–267 (2006).
8. R. B. Pollnac, R. S. Pomeroy, I. H. T. Harkes, Fishery policy and job satisfaction in three southeast Asian fisheries. *Ocean Coast. Manage.* **44**, 531–544 (2001).
9. K. Abernethy, "Fishing for what?: Understanding fisher decision-making in southwest England," PhD thesis, University of East Anglia, (2010).
10. J. Lee, A. B. South, S. Jennings, Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (VMS) data. *ICES J. Mar. Sci.* **67**, 1260–1271 (2010).
11. O. Spiegel, S. T. Leu, C. M. Bull, A. Sih, What's your move? Movement as a link between personality and spatial dynamics in animal populations. *Ecol. Lett.* **20**, 3–18 (2017).
12. R. Joo, S. Bertrand, J. Tam, R. Fablet, Hidden Markov models: The best models for forager movements? *PLoS One* **8**, e71246 (2013).
13. R. Joo, O. Salcedo, M. Gutierrez, R. Fablet, S. Bertrand, Defining fishing spatial strategies from VMS data: Insights from the world's largest monospecific fishery. *Fish. Res.* **164**, 223–230 (2015).
14. S. O'Farrell *et al*., Disturbance modifies payoffs in the explore-exploit trade-off. *Nat. Commun.* **10**, 3363 (2019).
15. L. Pfeiffer, T. Gratz, The effect of rights-based fisheries management on risk taking and fishing safety. *Proc. Natl. Sci. Acad. U.S.A.* **113**, 2615–2620 (2016).
16. R. B. Pollnac, J. J. Poggie, Happiness, well-being and psychocultural adaptation to the stresses associated with marine Fishing. *Hum. Ecol. Rev.* **15**, 194–200 (2008).
17. D. S. Holland, J. G. Sutinen, Location choice in New England trawl fisheries: Old habits die hard. *Land Econ.* **76**, 133–149 (2000).
18. I. Witten, E. Frank, *Practical Machine Learning Tools and Techniques* (Morgan Kaufmann, San Francisco, ed. 2, 2005).
19. I. E. van Putten *et al*., Theories and behavioural drivers underlying fleet dynamics models. *Fish Fish.* **13**, 216–235 (2012).
20. N. Deporte, C. Ulrich, S. Mahévas, S. Demaneche, F. Bastardie, Regional métier definition: A comparative investigation of statistical methods using a workflow applied to international otter trawl fisheries in the North Sea. *ICES J. Mar. Sci.* **69**, 331–342 (2012).
21. S. C. Anderson *et al*., Benefits and risks of diversification for individual Fishers. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10797–10802 (2017).
22. A. Lavoie *et al*., Ground-truthing social vulnerability indices of Alaska fishing communities. *Coast. Manage.* **46**, 359–387 (2018).
23. E. A. Fulton, A. D. M. Smith, D. C. Smith, I. E. van Putten, Human behaviour: The key source of uncertainty in fisheries management. *Fish Fish.* **12**, 2–17 (2010).
24. W. J. Boonstra, S. Birnbaum, E. Björkvik, The quality of compliance: Investigating Fishers' responses towards regulation and authorities. *Fish Fish.* **18**, 682–697 (2017).
25. P. Legendre, L. F. Legendre, *Numerical Ecology* (Elsevier, 2012).
26. M. Z. Rodriguez *et al*., Clustering algorithms: A comparative approach. *PLoS One* **14**, e0210236 (2019).
27. R. B. Pollnac, T. Seara, L. L. Colburn, M. Jepson, Taxonomy of USA east coast fishing communities in terms of social vulnerability and resilience. *Environ. Impact Assess. Rev.* **55**, 136–143 (2015).
28. S. L. Smith, R. B. Pollnac, L. L. Colburn, J. Olson, Classification of coastal communities reporting commercial fish landings in the US Northeast region: Developing and testing a methodology. *Mar. Fish. Rev.* **73**, 41–61 (2011).
29. T. A. Branch *et al*., Fleet dynamics and fishermen behavior: Lessons for fisheries managers. *Can. J. Fish. Aquat. Sci.* **63**, 1647–1668 (2006).
30. S. J. Bucaram, J. W. White, J. N. Sanchirico, J. E. Wilen, Behavior of the Galapagos fishing fleet and its consequences for the design of spatial management alternatives for the red spiny lobster fishery. *Ocean Coast. Manage.* **78**, 88–100 (2013).
31. NOAA Fisheries, "Final Report on the Survey of Participants in the Gulf of Mexico Grouper-Tilefish (G-T) Individual Fishing Quota (IFQ) Program" (Southeast Fisheries Science Center, Miami, FL, 2015).
32. ICES, "Report of the Study Group on the development of fishery-based forecasts" (ICES Document CM 2003/ACFM: 08 Ref. D., 2003).
33. C. Ulrich *et al*., Challenges and opportunities for fleet- and métier-based approaches for fisheries management under the European Common Fishery Policy. *Ocean Coast. Manage.* **70**, 38–47 (2012).
34. S. Davie, C. Lordan, Examining changes in Irish fishing practices in response to the cod long-term plan. *ICES J. Mar. Sci.* **68**, 1638–1646 (2011).
35. H. F. Prytherch "A descriptive survey of the bottom longline fishery in the Gulf of Mexico" (1983).
36. D. G. Foster, J. R. Pulver, E. Scott-Denton, C. Bergmann, Minimizing bycatch and improving efficiency in the commercial bottom longline fishery in the Eastern Gulf of Mexico. *Fish. Res.* **196**, 117–125 (2017).
37. A. McCrea-Strub *et al*., Potential impact of the Deepwater Horizon oil spill on commercial fisheries in the Gulf of Mexico. *Fisheries (Bethesda, Md.)* **36**, 332–336 (2011).
38. C. Amornbunchornvej *et al*., FLICA: A framework for leader identification in coordinated activity. arXiv:160301570 (2016).
39. M. O. Hill, Diversity and evenness: A unifying notation and its consequences. *Ecology* **54**, 427–432 (1973).
40. C. E. Shannon, A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.* **5**, 3–55 (2001).
41. O. Krishna, T. Yamasaki, A. Helo, R. Pia, K. Aizawa, Developmental changes in ambient and focal visual processing strategies. *Electronic Imaging* **2017**, 224–229 (2017).
42. C. Calenge, The package "adehabitat" for the R software: A tool for the analysis of space and habitat use by animals. *Ecol. Modell.* **197**, 516–519 (2006).
43. H.-M. Zhang, J. J. Bates, R. W. Reynolds, Assessment of composite global sampling: Sea surface wind speed. *Geophys. Res. Lett.* **33**, L17714 (2006).
44. A. F. Zuur, E. N. Ieno, A protocol for conducting and presenting results of regression-type analyses. *Methods Ecol. Evol.* **7**, 636–645 (2016).
45. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).
46. A. Liaw, M. Wiener, Classification and regression by randomForest. *R. News* **2**, 18–22 (2002).
47. S. Chamberlain, rerddap: General Purpose Client for "ERDDAP" Servers, Version 0.4.2. https://CRAN.R-project.org/package=rerddap. Accessed 21 October 2017.
48. R. J. Hijmans, geosphere: Spherical Trigonometry, Version 1.5-5. https://CRAN.R-project.org/package=geosphere. Accessed 17 July 2017.
49. J. Oksanen *et al*., vegan: Community Ecology Package, Version 2.5-3. https://CRAN.R-project.org/package=vegan. Accessed 13 July 2017.
50. M. Charrad, N. Ghazzali, V. Boiteau, A. Niknafs, NbClust: An R package for determining the relevant number of clusters in a data set. *J. Stat. Softw.* **61**, 1–36 (2014).
51. A. Kassambara, F. Mundt, factoextra: Extract and Visualize the Results of Multivariate Data Analyses, Version 1.0.5. https://CRAN.R-project.org/package=factoextra. Accessed 13 July 2017.