

Harvest–release decisions in recreational fisheries

Mark A. Kaemingk, Keith L. Hurley, Christopher J. Chizinski, and Kevin L. Pope

Abstract: Most fishery regulations aim to control angler harvest. Yet, we lack a basic understanding of what actually determines the angler's decision to harvest or release fish caught. We used XGBoost, a machine learning algorithm, to develop a predictive angler harvest–release model by taking advantage of an extensive recreational fishery data set (24 water bodies, 9 years, and 193 523 fish). We were able to successfully predict the harvest–release outcome for 99% of fish caught in the training data set and 96% of fish caught in the test data set. Unsuccessful predictions were mostly attributed to predicting harvest of fish that were released. Fish length was the most essential feature examined for predicting angler harvest. Other important predictive harvest–release features included the number of individuals of the same species caught, geographic location of an angler's residence, distance traveled, and time spent fishing. The XGBoost algorithm was able to effectively predict the harvest–release decision and revealed hidden and intricate relationships that are often unaccounted for with classical analysis techniques. Exposing and accounting for these angler–fish intricacies is critical for fisheries conservation and management.

Résumé : La plupart des règlements relatifs à la pêche visent à contrôler les prises des pêcheurs à la ligne. Une compréhension de base de ce qui détermine réellement la décision d'un pêcheur de conserver ou de relâcher un poisson pêché manque toutefois. Nous avons utilisé XGBoost, un logarithme d'apprentissage automatique, pour élaborer un modèle prédictif de décisions de pêcheurs de conserver ou relâcher un poisson en tirant parti d'un vaste ensemble de données de pêche sportive (24 plans d'eau, 9 années, 193 523 poissons). Nous avons été en mesure de prédire avec succès le résultat (conserver ou relâcher) pour 99 % des poissons pêchés dans l'ensemble de données d'entraînement et 96 % des poissons pêchés dans l'ensemble de données expérimental. Les prédictions inexactes étaient pour la plupart de poissons conservés qui avaient en fait été relâchés. La longueur du poisson est l'aspect examiné le plus important pour la prédiction de la conservation par les pêcheurs. D'autres aspects importants pour prédire la conservation ou le lâcher comprennent le nombre de spécimens de la même espèce pêchés, l'emplacement géographique de la résidence du pêcheur, la distance parcourue et le temps passé à pêcher. L'algorithme XGBoost est arrivé à prédire efficacement les décisions de conserver ou de relâcher et a fait ressortir des relations cachées et complexes dont les méthodes d'analyse classiques ne tiennent souvent pas compte. La reconnaissance et la prise en considération de ces facteurs complexes associés aux pêcheurs et aux poissons sont d'importance clé pour la conservation et la gestion des ressources halieutiques. [Traduit par la Rédaction]

Introduction

Once a fish is caught by an angler, will it be harvested or released? Currently, we lack a basic understanding of this very important decision process. Fish harvest by anglers can alter species abundance and size structure, ultimately affecting biodiversity and trophic dynamics (Cooke and Schramm 2007). An angler's predisposition to harvest a fish caught is arguably one of the most important considerations in fisheries management and conservation (Cooke and Cowx 2004; Birkeland and Dayton 2005). In fact, most management regulations are centered on this harvest–release decision with the intention of controlling harvest (Radomski et al. 2001) typically through the use of designated fishing seasons and limiting the number or size of fish harvested (Hubert and Quist 2010). Given the importance of harvest in recreational fisheries and the long history of regulating harvest, we should be able to predict the harvest–release decision with a high degree of certainty (see Hunt et al. 2002). We have traditionally operated on the basic premise that fish harvest depends primarily on fish size (Allendorf and Hard 2009; Chizinski et al. 2014), but is the angler's decision process really that simple?

Anglers comprise a heterogeneous group that varies in motivations, specializations, and preferences (Johnston et al. 2010; Haab et al. 2012). Thus, we anticipate that the harvest–release decision is complex and depends on both catch and noncatch attributes (Sutton and Ditton 2001; Gwinn et al. 2015). In one of the few studies to directly address the harvest–release decision in recreational fisheries, Hunt et al. (2002) determined that angling effort and catch rates were primarily related to harvest for three species of fish (although fish size was not considered in the assessment). Motivation and social groups were also considered important, but explained less variation (Hunt et al. 2002). In a less direct test of the harvest–release decision, harvest rates did not differ among three distinct segments of German anglers despite differences in their catch orientation (Arlinghaus 2006). The weak explanatory power of general angler motivations appears to underscore the complexity of the harvest decision (Hunt et al. 2002; Arlinghaus 2006). Recreational fisheries reside at the nexus of food and fun (Cooke et al. 2018), and thus our inability to predict angler harvest with a simple explanatory harvest–release model is not particularly surprising.

Received 2 April 2019. Accepted 23 July 2019.

M.A. Kaemingk. Nebraska Cooperative Fish and Wildlife Research Unit, and School of Natural Resources, University of Nebraska, Lincoln, NE 68583, USA.

K.L. Hurley. Nebraska Game and Parks Commission, Lincoln, NE 68503, USA.

C.J. Chizinski. School of Natural Resources, University of Nebraska, Lincoln, NE 68583, USA.

K.L. Pope. US Geological Survey—Nebraska Cooperative Fish and Wildlife Research Unit, and School of Natural Resources, University of Nebraska, Lincoln, NE 68583, USA.

Corresponding author: Mark A. Kaemingk (email: mkaemingk2@unl.edu).

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from [RightsLink](https://www.copyright.com).

Given fish size alone is unlikely to predict an angler's decision to harvest, why have we not developed a more comprehensive angler harvest model? First, management regulations, and particularly those relating to harvest, can be very complex (Radomski et al. 2001). The diversity of management regulations could confound our ability to separate the effects of compliance and an angler's desire to harvest fish, especially in different contexts (e.g., conservative or liberal bag or size limits). Second, trip context (daily or multiday trips) likely influences the harvest–release decision; potential factors include when a fish is caught during the trip (i.e., early or late), weather conditions, distance traveled, and social aspects (i.e., solo or group). Therefore, individual angler heterogeneity or plasticity could complicate harvest predictions. Third, not all fish species are viewed equally among anglers and exist along a harvest–release continuum. Black bass (*Micropterus* spp.) are more likely to be released, whereas crappie (*Pomoxis* spp.) are more likely to be harvested (Colvin 1991; Siepker et al. 2007), providing an example where fish size would not explain the harvest–release decision given different social harvest norms (Arlinghaus 2007; Stensland and Aas 2014). Fourth, we presume that large recreational fishery (as opposed to commercial fishery) data sets are lacking, and the data sets that do exist are used primarily for monitoring and have not been leveraged for research. In any event, robust recreational fishery data sets (i.e., multiyear and water body) are imperative for the development of predictive harvest–release models.

Herein, we explore an extensive angler survey data set with the intention of (i) developing a predictive harvest–release model and (ii) identifying essential factors that constitute an angler's predisposition to harvest fish. Our approach includes onsite angler surveys (24 water bodies, 9 years) to explore how social and ecological factors influence an angler's decision to harvest fish. We believe an empirical understanding of the harvest–release decision is necessary to avoid undesirable social and ecological consequences. Such insights could lead to more management options and greater effectiveness in controlling harvest, both for overexploited and underexploited (e.g., invasive species) populations. We may also be able to predict how certain harvest regulations may influence subpopulations of anglers based on their choice of species targeted, willingness to travel, trip context, angling method, and other social and ecological factors. This information will ultimately allow for more creative methods and techniques to manage the angler–fish interaction in recreational fisheries.

Materials and methods

Angler surveys

Angler harvest–release information was collected at 24 water bodies in Nebraska, USA, during 2009–2017 from April through October (see online Supplementary material, Table S1¹). Water bodies ranged in size from 8 to 12 141 ha and were located in both urban and rural settings; these water bodies were primarily used for hydropower, irrigation storage, or flood control, though all were actively managed for recreational fishing. Anglers surveyed at Nebraska water bodies targeted a diverse range of fish species (Pope et al. 2016). We collected angler behavior information via in-person interviews at each reservoir according to previously described methods (Malvestuto et al. 1996; Kaemingk et al. 2018). A stratified multistage probability-sampling regime was used to determine days of interviews (Malvestuto et al. 1996). The number of days surveyed per month varied across water bodies, depending on surface area and logistics. Within a month, survey days were stratified into either a week or a weekend day to account for variation in day type. Days were further stratified into a morning or afternoon survey period.

Angler interviews were conducted at the group or party (i.e., individuals travelling together for fishing) level whereby one angler (i.e., party representative) completed the survey for the entire party. From these interviews, we collected catch information (harvested or released, species, length). In addition, we recorded the species sought, fishing start and end times, time fished, party size, angler's residence zip code, distance traveled, and angler type (i.e., bank or boat). We also documented angler and nonangler effort by counting the number of bank and boat anglers, as well as nonfishing boats, during each survey period (Malvestuto et al. 1996; Kaemingk et al. 2018). Finally, we included fish harvest regulations that were specific to each fish capture event, such as whether the fish was legal to harvest (based on length only) and whether there were bag or size limits (i.e., water-body-specific, statewide, or no restrictions).

Data analyses

Harvest–release model development and training

We used a machine learning method for our recreational fishery data set to develop a predictive harvest–release model. Machine learning is a broad field and has the advantage of handling large multidimensional and complex data sets to understand relationships that cannot be revealed by classical methods (Bzdok et al. 2018). Whereas classical techniques are often limited by large amounts of variance, machine learning capitalizes on this variance to reveal intricate relationships. Harvest–release decisions in recreational fisheries are likely complex; thus, we used machine learning on a large data set containing multiple social and ecological variables that spanned several spatial and temporal scales.

We specifically used the Extreme Gradient Boost (XGBoost) algorithm to develop a predictive harvest–release model given the ability and versatility of XGBoost to address complex problems. The XGBoost algorithm is a scalable tree-boosting algorithm that has proven to be a superior method and, as such, has been adopted across many disciplines to tackle complex data mining and machine learning problems (Chen and Guestrin 2016). The most attractive reasons for employing XGBoost lies in the ability to handle sparse data, scalability to a wide variety of scenarios, and excellent computational speeds (Chen and Guestrin 2016). Briefly, XGBoost strongly considers and accounts for model complexity and avoids underfitting or overfitting the data. This bias–variance trade-off is handled by increasing neighborhood (e.g., locality or position of trees) size to avoid increasing variance, unless a complex structure is apparent and then neighborhood size is decreased (Nielsen 2016). The adaptive adjustment of these neighborhoods overcomes traditional methods that struggle to incorporate multidimensional data. The XGBoost algorithm leverages Newton boosting, which is extremely effective for determining tree structure and consequently the neighborhoods (Nielsen 2016).

We selected 21 explanatory or independent variables (Table 1) based on their putative ability to contribute to the harvest–release decision (Hunt et al. 2002). Our data set contained 18 555 interviews that recorded the harvest–release outcome of 193 523 fish caught. Of the fish caught, 133 958 (69%) fish were released and 59 565 (31%) were harvested. Each fish (i.e., experimental unit) was used for either the development or testing of the harvest–release decision model, exclusively. Explanatory variables ranged from more social-oriented (e.g., party size, angler type, size limits) to more social–ecological-oriented (e.g., catch rate, species caught, fish length). Fish length, number caught, and catch rate (i.e., number of fish caught per hour) were normalized from 0 to 1 for each species according to the range in the data set (i.e., 0 length refers to the smallest fish caught). Angler count variables (bank anglers, boat anglers, total anglers (bank and boat), and total boats (an-

¹Supplementary data are available with the article through the journal Web site at <http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2019-0119>.

Table 1. Social and ecological variables ($n = 21$) used to predict angler harvest–release decisions.

Variable	Description	Type
Species	Species of fish	Nominal
Fish length	Total length of fish	Numeric
Species sought	Primary target species	Nominal
Target species	Species sought and caught	Boolean
Number caught	Number of individuals of the same species caught	Numeric
Catch rate	Catch rate of same species	Numeric
Party size	Number of people in party	Numeric
Zip code	Zip code of angler's residence	Nominal
Distance traveled	Kilometres from centroid of angler's zip code to centroid of water body	Numeric
Month	Month the interview occurred	Nominal
Start time	Hour of day fishing began	Ordinal
Trip length	Minutes spent fishing	Numeric
Angler type	Bank or boat angler	Nominal
Number species caught	Number of unique species caught	Numeric
Bag limits present	Bag limits for species at water body	Ordinal
Size limits present	Size limits for species at water body	Ordinal
Legal for harvest	Fish legal (yes, no) for harvest (only relevant for minimum size limits)	Boolean
Bank anglers	Mean count of bank anglers by day	Numeric
Boat anglers	Mean count of boat anglers by day	Numeric
Total anglers	Mean count of all anglers (boat and bank) by day	Numeric
Total boats	Mean count of all boats (angling and nonangling) by day	Numeric

Note: Fish length, number caught, and catch rate were normalized (0 to 1) by species across the entire data range. Counts of bank anglers, boat anglers, total anglers, and total boats were standardized by water body size (i.e., number per hectare).

glers and nonanglers)) were standardized by water body size (i.e., number per hectare). For some explanatory variables, we further examined species-specific harvest patterns for the most caught species (bluegill (*Lepomis macrochirus*), crappie (*Pomoxis* spp.), wall-eye (*Sander vitreus*), and white bass (*Morone chrysops*)).

The harvest–release decision model was trained using the extreme gradient boosted decision tree algorithm implemented by the XGBoost library in the XGBoost R package (Chen et al. 2015) as initially described by Friedman et al. (2000) and Friedman (2001). We used the 21 explanatory variables (Table 1) for model training to predict the binary harvest–release decision (1 = harvested, 0 = released). Of the fish caught, 75% ($n = 145\ 142$) were randomly used for the training data set, while the remaining 25% ($n = 48\ 381$) were reserved for the test data set. The percentage of fish harvested was similar between the training (31%) and test (31%) data sets. Additionally, the two data sets had similar distributions of species and water body combinations (Fig. S1¹). Hyperparameter tuning (e.g., computationally minimizing classification error) was conducted using tenfold cross-validation and a grid search methodology to control for over-fitting of the model and a final tenfold cross-validation model was trained to identify the optimal number of trees in the final model.

Harvest–release model testing

We used the resultant model to create harvest–release predictions on our test data set that contained known harvest release outcomes. A confusion matrix was created to identify any directional bias in the model's incorrect predictions. Model feature gain, coverage, and frequency metrics were assessed to identify each variable's importance to the prediction process. Feature gain expresses the magnitude of impact or relative contribution of a given variable when it is used in the prediction process. The feature gain metric is calculated as the sum of the given variable's contribution for each tree in the model (expressed as a percentage of all gain metrics). Feature coverage measures the relative number of observations related to a particular variable. Thus, the feature coverage metric is the total count of all harvest–release decisions for all trees in the model that were influenced by the given variable (expressed as a percentage of all cover metrics). Feature frequency accounts for the possibility that a given feature may be used more than once for a given observation and is the relative number of times a variable is used compared with all the

Table 2. Confusion matrix of training and test data set predictions for the harvest–release model.

Data set	Predicted angler decision	Angler decision	
		Released	Harvested
Training	Released	99 813 (69%)	1 104 (1%)
	Harvested	668 (1%)	43 557 (30%)
Test	Released	32 743 (68%)	1 061 (2%)
	Harvested	734 (2%)	13 843 (29%)

other variables used in the trees of the model (expressed as a percent weight of all weights). Essentially, variables with high gain indicate large impact on the final prediction, variables with high coverage are used by a large percentage of the predictions, and variables with high frequency are frequently used in the model decision process. Variables with high gain, coverage, and frequency importance were considered essential model variables because of their contribution and were further evaluated to explain the harvest–release decision. A log-odds impact on the probability of the harvest decision was calculated for each essential model variable using the xgboostExplainer R package (Foster 2017) and used for visual assessment of relationship patterns.

Results

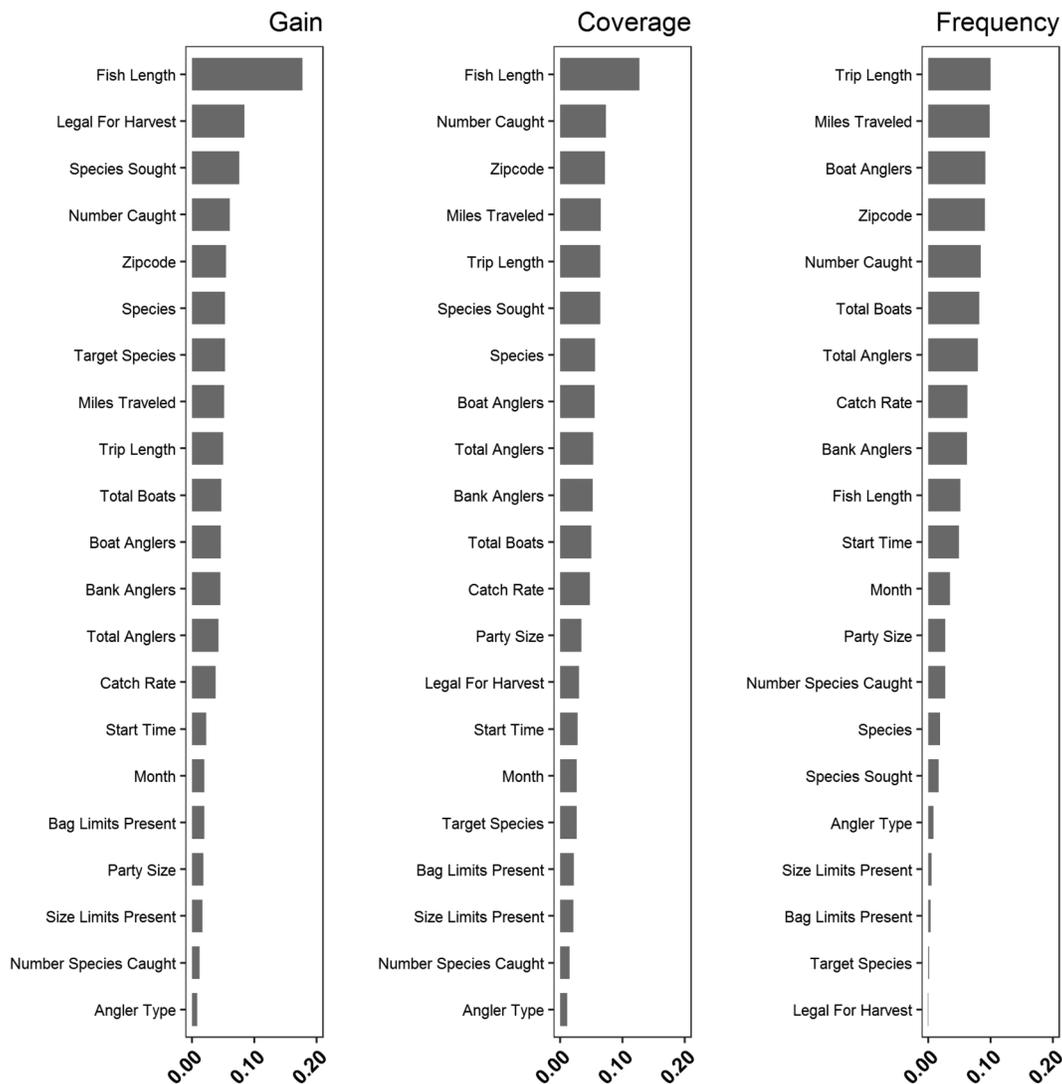
Harvest–release model development and training

Model training hyperparameter values were chosen from three grid searches by minimizing mean validation error and, secondarily, minimizing the number of boosted trees used in the model (Table S2¹). The harvest–release decision model was then trained using the hyperparameter values (Table S3¹). Model accuracy for predicting an angler's predisposition to harvest a given fish was 99% on the training data set (Table 2). The confusion matrix and corresponding 1% error rate indicated that incorrect predictions were 1.7 times more likely to predict a release outcome for a harvested fish (i.e., false negative).

Harvest–release model testing

We were able to successfully predict the harvest–release outcome for 96% of fish on the test data set (Table 2). Of the 4% of

Fig. 1. Level of importance (e.g., higher scores equal greater contribution) across three importance metrics (gain, coverage, frequency) for each independent variable used to predict the angler harvest–release decision (see Materials and methods section for more details). (Note: for miles travelled, 1 mile = 1.609 km.)



predictions that were incorrect, we were 1.4 times more likely to predict a release outcome when the fish was harvested (i.e., false negative).

Five features were consistently important in explaining the harvest–release decision, based on gain, coverage, and frequency importance scores (Fig. 1). These features were fish size, number of individuals of the same species caught, angler’s residence zip code, distance traveled, and time fished (Fig. 1). Although these features were consistently ranked high according to their level of importance across the metrics, features contributed in different manners when predicting the harvest–release decision. For example, fish length was the most important variable for gain and coverage metrics, whereas trip length and miles traveled (1 mile = 1.609 km) were the most important variables for the frequency metric. Therefore, fish length had a large impact on the final prediction and was used by a large percentage of the predictions; trip length and miles traveled, however, were frequently used in the model decision process with less impact on the prediction when they were used.

We revealed unique relationships among the five features with respect to their influence on the harvest–release decision. The importance and influence of fish size on the harvest–release deci-

sion was similar across species and followed a polynomial relationship (Fig. 2). As the number of individuals of the same species caught increased, fish were more likely to be released (Fig. 3). Knowing an angler’s residence was also useful for predicting the harvest–release decision, especially for certain regions of Nebraska where anglers appear to be more release-oriented (e.g., southeast; Fig. 4). Predicting harvest–release decisions for anglers that traveled shorter distances was easier compared with anglers that traveled longer distances (Fig. 5). Finally, the propensity to harvest generally decreased as a function of time spent fishing, although the strength of this pattern appeared to be species-specific (Fig. 6).

Discussion

We were able to accurately predict the harvest–release decision in our recreational fishery data set. This approach required an extensive amount of information and a machine learning technique (XGBoost) to reveal intricate dynamics that are typically obscured or unavailable in more classical assessments. Not surprisingly, an angler’s decision to harvest is complex and is related to multiple social and ecological features. Fish size was an essen-

Fig. 2. The impact of fish length (normalized from 0 (smallest fish caught) to 1 (largest fish caught) for each species) on the angler harvest–release decision across all fish (top panel) and specifically for captured bluegill, crappie, walleye, and white bass (bottom panels); positive log-odds indicate a higher likelihood of harvest, whereas negative log-odds indicate a lower likelihood of harvest.

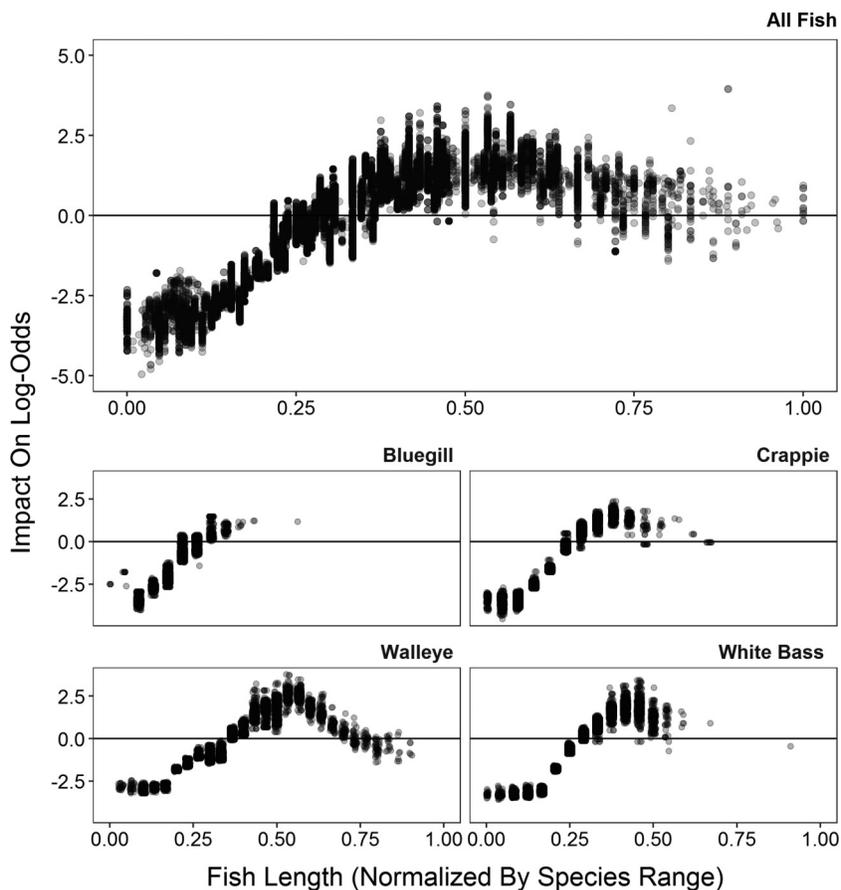
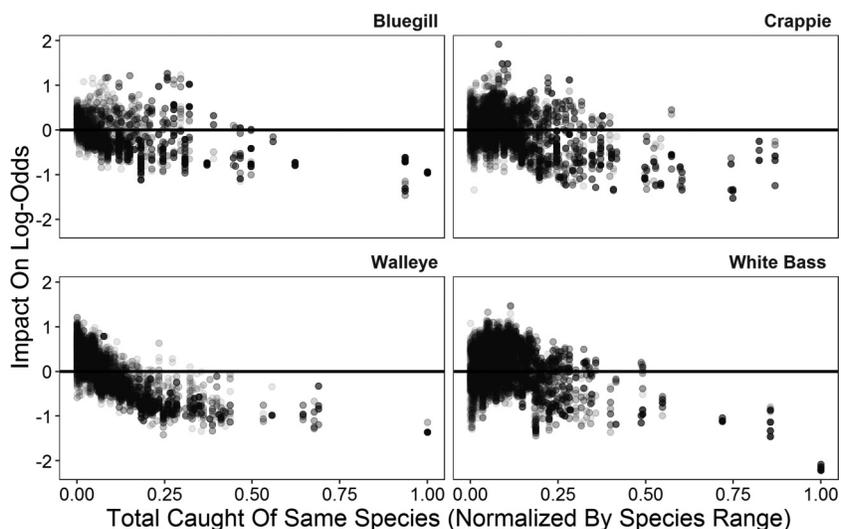


Fig. 3. The impact of the number of caught individuals of the same species (normalized from 0 (fewest fish caught) to 1 (most fish caught) for each species) on the angler harvest–release decision; positive log-odds indicate a higher likelihood of harvest, whereas negative log-odds indicate a lower likelihood of harvest.



tial component to the harvest–release decision, and the relationship was similar across species. Past assessments have assumed that fish harvest follows a linear or logistic relationship with fish size (Chizinski et al. 2014). We revealed that harvest is more likely to follow a polynomial relationship with fish size and that predict-

ing the harvest–release outcome for larger fish is difficult. Trip context, such as the number of same species caught and time spent fishing, was also important and suggests that an angler's decision to harvest is not fixed. Finally, it may be useful to consider both the regional differences in angler composition (e.g., zip

Fig. 4. The impact of an angler’s residence (i.e., zip code) on the angler harvest–release decision; positive log-odds indicate a higher likelihood of harvest, whereas negative log-odds indicate a lower likelihood of harvest.

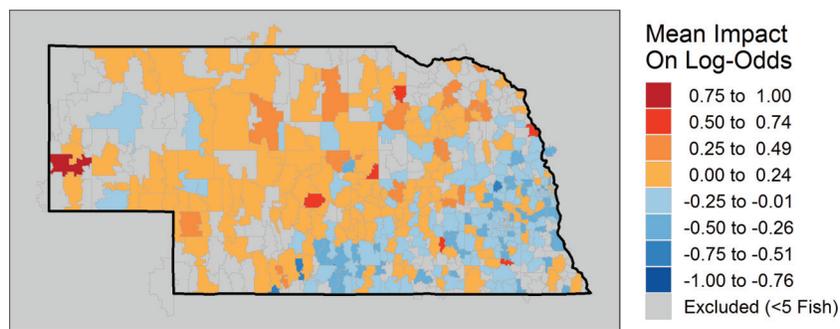


Fig. 5. The impact of distance traveled on the angler harvest–release decision of captured bluegill, crappie, walleye, and white bass; positive log-odds indicate a higher likelihood of harvest, whereas negative log-odds indicate a lower likelihood of harvest.

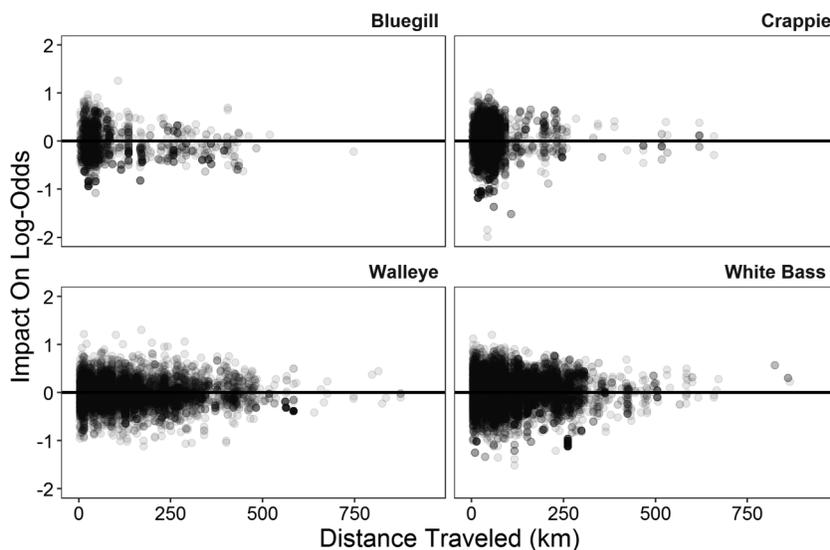
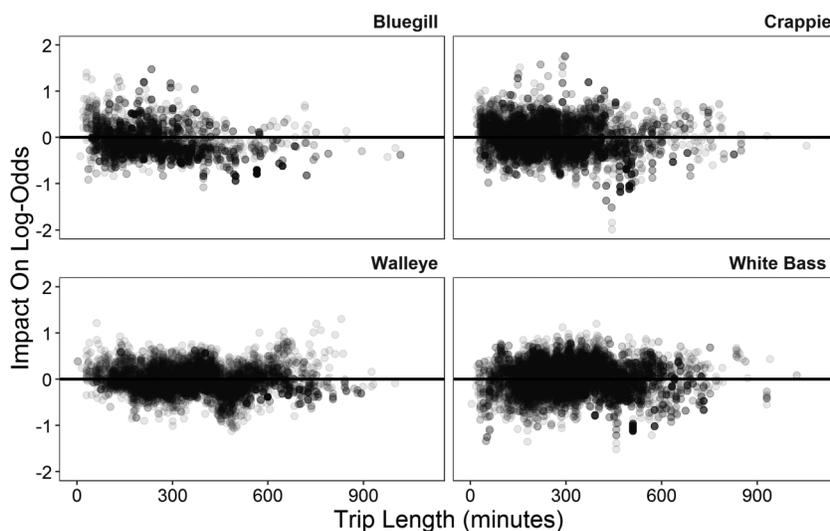


Fig. 6. The impact of time fished on the angler harvest–release decision of captured bluegill, crappie, walleye, and white bass; positive log-odds indicate a higher likelihood of harvest, whereas negative log-odds indicate a lower likelihood of harvest.



code — urban versus rural) and the spatial arrangement of anglers and water bodies (e.g., distance traveled) on the landscape (Matsumura et al. 2019). Enabling managers and policy makers to predict harvest in recreational fisheries is a powerful manage-

ment tool and could be particularly valuable in situations where overexploitation is a concern.

The prevailing belief is that fish size is one of the most essential aspects of the harvest decision (Fisher 1997). Our findings support

this assertion, with fish size contributing most to gain and coverage importance in the harvest–release model. Previous work predicted that a threshold exists (i.e., self-imposed length limit) as to when a fish will be released or harvested (Chizinski et al. 2014). Fish below this threshold will likely be released and fish above this threshold will likely be harvested, thus balancing the costs and trade-offs of harvest with respect to fish size (Chizinski et al. 2014). Our results support this size trade-off prediction across a range of species, and thus this relationship was not species-specific. This commonly shared size-dependent relationship and outcome is certainly important to understand if harvest regulations are aimed to limit or promote the harvest of certain size groups. For example, setting species-specific harvest restrictions based on size may have little effect if fish will be voluntarily released (especially among more or less harvest-oriented anglers; Colvin 1991; Siepker et al. 2007). Assuming a simple harvest and size relationship across the entire size range could also be misleading for some species, such as walleye, where small and large individuals appear to be released. Predicting angler harvest based on fish size could consider general and specific patterns within the appropriate context.

The context of an angler's trip was influential in predicting the harvest decision. In particular, the number of the same species of fish caught had an effect on harvest–release decisions. The likelihood of a fish being harvested was greater if fewer fish were caught. As the number of fish caught increased, the likelihood of harvest decreased. This outcome could be a function of high-grading or preferentially releasing fish already caught for more preferred fish (e.g., larger fish) caught later in the fishing trip (Coleman et al. 2004). While the presence of restrictive regulations such as bag limits may play a role in this relationship, model features representing such regulations were of low importance in the harvest–release model predictions. The number of fish caught and the decision to harvest may also be a function of time spent fishing. Therefore, the number of fish caught and time fished could interact to shape an angler's decision to harvest. Our study confirms the importance of trip context on the decision to harvest, illustrating that such decisions are dynamic in time, space, and across individual anglers.

Our harvest–release model also highlights that both the geographic location of an angler's residence and distance traveled are important for predicting angler harvest. Spatial heterogeneity of anglers and their proximity to ecological resources on the landscape is pertinent to consider when creating and establishing harvest regulations. It appears this relationship could be related to the rural–urban gradient in Nebraska, with a higher density of people residing in southeastern Nebraska (Fig. S2¹). Anglers living in urban settings are less likely to harvest fish and less likely to travel long distances compared with anglers living in rural settings (Arlinghaus and Mehner 2004; Arlinghaus et al. 2008). Spatial heterogeneity among anglers may lead to emergent properties that create complex cross-scale dynamics (Kaemingk et al. 2018). We are currently unaware if most management agencies are collecting spatial information from anglers and other sportspersons, but our findings highlight the utility of incorporating this type of information for management and conservation purposes. Neglecting to consider spatial properties of angler–fish interactions could lead to unintended consequences (Kaemingk et al. 2018). In our case, setting harvest regulations that consider the catchment area (i.e., spatial draw) of anglers to a particular water body could improve assessing and achieving management goals (Martin et al. 2015, 2017).

Although we successfully predicted the harvest decision and identified the salient factors that constitute an angler's predisposition to harvest, we were surprised that some variables were not more important in the decision process. We expected management regulations (e.g., bag limits, statewide versus water-body-specific) to be more influential in predicting harvest. After all,

these management regulations are primarily intended to control or regulate harvest and are often viewed as the primary vehicle to govern angler behavior (Scrogin et al. 2004). We wonder if our study did not include enough variation in management regulations (e.g., bag and size limits) to be an informative model feature, because many of the water bodies fell under a statewide regulation. Alternatively, management regulations may not be as important as previously thought. There was also an expectation that temporal factors (e.g., month, intraday start time) and weather conditions could alter an angler's decision to harvest (Colvin 1991; Isermann et al. 2005). These findings underscore the complexity of the angler harvest decision and provide justification for testing these relationships with empirical data. It is worth noting that though these factors were of low importance at the angler population level, they may be highly important at an individual level for the harvest–release decision. We encourage future efforts to predict angler harvest in other contexts (marine versus freshwater, domestic versus international), as we anticipate some findings will corroborate our results whereas others will be unexpected and ultimately lead to forming more creative management options and tools.

Angler harvest is a critical component of recreational fisheries that has long been appreciated, but the intricacies of how this decision is made are often overlooked. Recreational harvest can lead to reduced size structure, overexploitation, and a shift in fish community dynamics, trophic cascades, and regime shifts (McPhee et al. 2002; Cooke and Cowx 2004; Arlinghaus and Cooke 2009). Recognizing which social and ecological factors contribute to the harvest decision is therefore important for fisheries management and conservation. In theory, managers could use the angler's decision process to increase or decrease harvest in different contexts. For example, one could predict how effective certain regulations will be by considering both social (e.g., angler residences and willingness to travel) and ecological (e.g., fish abundance and size structure) components of a fishery. A water body residing in an urban environment with high fish size structure may receive less harvest compared with the same water body residing in a rural environment, assuming similar fishing pressure.

We hope that our findings will contribute to a much deeper understanding of the harvest–release decision in recreational fisheries. Our intention is that other studies will employ machine learning techniques to understand this complex decision process. We wonder if fish size, the number of same species caught, geographic location of an angler's residence, distance traveled, and time spent fishing will remain important in other social–ecological settings that vary in angler heterogeneity and resource quality. We believe more research in this area is warranted given the importance of understanding angler harvest in recreational fisheries and hope our novel approach will reinvigorate investigations into the harvest decisions of anglers.

Acknowledgements

We thank T. Anderson, R. Barg, B. Bird, D. Bohnenkamp, Z. Brashears, D. Brundrett, K. Carpenter, M. Cavallaro, P. Chvala, N. Cole, M. Coll, C. Dietrich, L. Dietrich, M. Dedinsky, C. Depue, D. Dobesh, D. Eichner, B. Eifert, H. Evans, A. Fandrich, A. Fedele, R. Foley, R. Fusselman, J. Glenn, A. Glidden, R. Grandi, A. Gray, J. Hair, A. Hanson, B. Harmon, C. Huber, S. Huber, H. Hummel, C. Hothan, J. Johnson, C. Knight, L. Kowalewski, R. Lawing, D. Liess, J. Lorensen, N. Luben, A. Maple, G. Maynard, B. McCue, J. Meirgard, J.P. Montes, C. Nelson, B. Newcomb, C. Niehoff, L. Ohlman, A. Park, A. Pella, M. Petsch, R. Pierson, B. Porter, B. Roberg, P. Rossmeier, C. Ruskamp, J. Rydell, J. Ryschon, T. Sanders, A. Schiltz, J. Schuckman, S. Sidel, M. Smith, J. Spicha, M. Staab, P. Stolberg, D. Thompson, J. Walrath, N. Weaver, and J. Yates for assistance in the field. Mel provided inspiration. We thank A. Taylor and two anonymous reviewers for improving this

manuscript. This project was funded by Federal Aid in Sport Fish Restoration project F-182-R, which was administered by the Nebraska Game and Parks Commission, and by the Nebraska Public Power District agreement No. 4200002717. The Institutional Review Board for the Protection of Human Subjects approved the research protocol (IRB Project ID 14051). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government. C. Chizinski was supported by Hatch funds through the Agricultural Research Division at the University of Nebraska–Lincoln and from Federal Aid in Wildlife Restoration projects W-120-T, administered by the Nebraska Game and Parks Commission. The Nebraska Cooperative Fish and Wildlife Research Unit is jointly supported by a cooperative agreement among the US Geological Survey, the Nebraska Game and Parks Commission, the University of Nebraska, the US Fish and Wildlife Service, and the Wildlife Management Institute. The authors declare no conflict of interest.

References

- Allendorf, F.W., and Hard, J.J. 2009. Human-induced evolution caused by natural selection through harvest of wild animals. *Proc. Natl. Acad. Sci.* **106**(Suppl. 1): 9987. doi:10.1073/pnas.0901069106.
- Arlinghaus, R. 2006. On the apparently striking disconnect between motivation and satisfaction in recreational fishing: the case of catch orientation of German anglers. *North Am. J. Fish. Manag.* **26**(3): 592–605. doi:10.1577/M04-220.1.
- Arlinghaus, R. 2007. Voluntary catch-and-release can generate conflict within the recreational angling community: a qualitative case study of specialised carp, *Cyprinus carpio*, angling in Germany. *Fish. Manag. Ecol.* **14**(2): 161–171. doi:10.1111/j.1365-2400.2007.00537.x.
- Arlinghaus, R., and Cooke, S.J. 2009. Recreational fisheries: socioeconomic importance, conservation issues and management challenges. *In* *Recreat. Hunt. Conserv. Rural Livelihoods Sci. Pract.* pp. 39–58.
- Arlinghaus, R., and Mehner, T. 2004. A management-orientated comparative analysis of urban and rural anglers living in a metropolis (Berlin, Germany). *Environ. Manage.* **33**(3): 331–344. doi:10.1007/s00267-004-0025-x. PMID:15037955.
- Arlinghaus, R., Bork, M., and Fladung, E. 2008. Understanding the heterogeneity of recreational anglers across an urban–rural gradient in a metropolitan area (Berlin, Germany), with implications for fisheries management. *Fish. Res.* **92**(1): 53–62. doi:10.1016/j.fishres.2007.12.012.
- Birkeland, C., and Dayton, P.K. 2005. The importance in fishery management of leaving the big ones. *Trends Ecol. Evol.* **20**(7): 356–358. doi:10.1016/j.tree.2005.03.015. PMID:16701393.
- Bzdok, D., Altman, N., and Krzywinski, M. 2018. Statistics versus machine learning. *Nat. Methods* **15**: 233–234. doi:10.1038/nmeth.4642.
- Chen, T., and Guestrin, C. 2016. XGBoost: a scalable tree boosting system. *In* *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 13 August 2016. pp. 785–794.
- Chen, T., He, T., Benesty, M., Khotilovich, V., and Tang, Y. 2015. Xgboost: extreme gradient boosting. R Package Version 04-2: 1–4.
- Chizinski, C.J., Martin, D.R., Hurley, K.L., and Pope, K.L. 2014. Self-imposed length limits in recreational fisheries. *Fish. Res.* **155**: 83–89. doi:10.1016/j.fishres.2014.02.022.
- Coleman, F.C., Figueira, W.F., Ueland, J.S., and Crowder, L.B. 2004. The impact of united states recreational fisheries on marine fish populations. *Science*, **305**(5692): 1958–1960. doi:10.1126/science.1100397. PMID:15331771.
- Colvin, M.A. 1991. Population characteristics and angler harvest of white crappies in four large Missouri reservoirs. *North Am. J. Fish. Manag.* **11**(4): 572–584. doi:10.1577/1548-8675(1991)011<0572:PCAAHO>2.3.CO;2.
- Cooke, S.J., and Cowx, I.G. 2004. The role of recreational fishing in global fish crises. *BioScience*, **54**(9): 857–859. doi:10.1641/0006-3568(2004)054[0857:TRORFI]2.0.CO;2.
- Cooke, S.J., and Schramm, H.L. 2007. Catch-and-release science and its application to conservation and management of recreational fisheries. *Fish. Manag. Ecol.* **14**(2): 73–79. doi:10.1111/j.1365-2400.2007.00527.x.
- Cooke, S.J., Twarddek, W.M., Lennox, R.J., Zolderdo, A.J., Bower, S.D., Gutowsky, L.F.G., et al. 2018. The nexus of fun and nutrition: recreational fishing is also about food. *Fish. Res.* **19**(2): 201–224. doi:10.1111/faf.12246.
- Fisher, M.R. 1997. Segmentation of the Angler population by catch preference, participation, and experience: a management-oriented application of recreation specialization. *North Am. J. Fish. Manag.* **17**(1): 1–10. doi:10.1577/1548-8675(1997)017<0001:SOTAPB>2.3.CO;2.
- Foster, D. 2017. xgboostExplainer. R Package Version 10.
- Friedman, J.H. 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* **29**(5): 1189–1232.
- Friedman, J., Hastie, T., and Tibshirani, R. 2000. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Ann. Stat.* **28**(2): 337–407. doi:10.1214/aos/1016218223.
- Gwinn, D.C., Allen, M.S., Johnston, F.D., Brown, P., Todd, C.R., and Arlinghaus, R. 2015. Rethinking length-based fisheries regulations: the value of protecting old and large fish with harvest slots. *Fish. Res.* **16**(2): 259–281. doi:10.1111/faf.12053.
- Haab, T., Hicks, R., Schnier, K., and Whitehead, J.C. 2012. Angler heterogeneity and the species-specific demand for marine recreational fishing. *Mar. Resour. Econ.* **27**(3): 229–251. doi:10.5950/0738-1360-27.3.229.
- Hubert, W., and Quist, M. (Editors). 2010. *Inland fisheries management in North America*. 3rd ed. Am. Fish. Soc., Bethesda, Md.
- Hunt, L., Haider, W., and Armstrong, K. 2002. Understanding the fish harvesting decisions by anglers. *Hum. Dimens. Wildl.* **7**(2): 75–89. doi:10.1080/10871200290089355.
- Isermann, D.A., Willis, D.W., Lucchesi, D.O., and Blackwell, B.G. 2005. Seasonal harvest, exploitation, size selectivity, and catch preferences associated with winter Yellow Perch Anglers on South Dakota Lakes. *North Am. J. Fish. Manag.* **25**(3): 827–840. doi:10.1577/M04-026.1.
- Johnston, F.D., Arlinghaus, R., and Dieckmann, U. 2010. Diversity and complexity of angler behaviour drive socially optimal input and output regulations in a bioeconomic recreational-fisheries model. *Can. J. Fish. Aquat. Sci.* **67**(9): 1507–1531. doi:10.1139/F10-046.
- Kaemingk, M.A., Chizinski, C.J., Hurley, K.L., and Pope, K.L. 2018. Synchrony — an emergent property of recreational fisheries. *J. Appl. Ecol.* **55**(6): 2986–2996. doi:10.1111/1365-2664.13164.
- Malvestuto, S.P., Murphy, B., and Willis, D. 1996. Sampling the recreational creel. *Fish. Tech.* 2nd ed. Am. Fish. Soc., Bethesda, Md. pp. 591–623.
- Martin, D.R., Chizinski, C.J., and Pope, K.L. 2015. Reservoir Area of Influence and Implications for Fisheries Management. *North Am. J. Fish. Manag.* **35**(2): 185–190. doi:10.1080/02755947.2014.975299.
- Martin, D.R., Shizuka, D., Chizinski, C.J., and Pope, K.L. 2017. Network analysis of a regional fishery: Implications for management of natural resources, and recruitment and retention of anglers. *Fish. Res.* **194**: 31–41. doi:10.1016/j.fishres.2017.05.007.
- Matsumura, S., Beardmore, B., Haider, W., Dieckmann, U., and Arlinghaus, R. 2019. Ecological, Angler, and spatial heterogeneity drive social and ecological outcomes in an integrated landscape model of freshwater recreational fisheries. *Rev. Fish. Sci. Aquac.* **27**(2): 170–197. doi:10.1080/23308249.2018.1540549.
- McPhee, P.D., Leadbitter, D., and Skilleter, A.G. 2002. Swallowing the bait: is recreational fishing in Australia ecologically sustainable? *Pac. Conserv. Biol.* **8**(1): 40–51. doi:10.1071/PC020040.
- Nielsen, D. 2016. Tree boosting with XGBoost—Why does XGBoost win every machine learning competition? M.Sc. thesis. Department of Mathematical Sciences, Norwegian University of Science and Technology.
- Pope, K.L., Chizinski, C.J., Wiley, C.L., and Martin, D.R. 2016. Influence of anglers' specializations on catch, harvest, and bycatch of targeted taxa. *Fish. Res.* **183**: 128–137. doi:10.1016/j.fishres.2016.05.025.
- Radomski, P.J., Grant, G.C., Jacobson, P.C., and Cook, M.F. 2001. Visions for recreational fishing regulations. *Fisheries*, **26**(5): 7–18. doi:10.1577/1548-8446(2001)026<0007:VFRFR>2.0.CO;2.
- Scrogin, D., Boyle, K., Parsons, G., and Plantinga, A.J. 2004. Effects of regulations on expected catch, expected harvest, and site choice of recreational Anglers. *Am. J. Agric. Econ.* **86**(4): 963–974. doi:10.1111/j.0002-9092.2004.00646.x.
- Sieper, M., Ostrand, K., Cooke, S., Philipp, D., and Wahl, D. 2007. A review of the effects of catch-and-release angling on black bass, *Micropterus* spp.: implications for conservation and management of populations. *Fish. Manag. Ecol.* **14**(2): 91–101. doi:10.1111/j.1365-2400.2007.00529.x.
- Stensland, S., and Aas, Ø. 2014. The role of social norms and informal sanctions in catch-and-release angling. *Fish. Manag. Ecol.* **21**(4): 288–298. doi:10.1111/fme.12078.
- Sutton, S.G., and Ditton, R.B. 2001. Understanding Catch-and-Release Behavior Among U.S. Atlantic Bluefin Tuna Anglers. *Human Dimensions of Wildlife*, **6**(1): 49–66. doi:10.1080/10871200152668698.