Chapter 1: Estimating fishing effort across the landscape: a spatially extensive approach using models to integrate multiple data sources

Authors: Ashley Trudeau[1], Colin J. Dassow[2], Carolyn M. Iwicki[1], Stuart E. Jones[2], Greg G. Sass[3], Christopher T. Solomon[4], Brett T. van Poorten[5], Olaf P. Jensen[6]

1. Graduate Program in Ecology and Evolution, Rutgers University, Department of Marine and Coastal Sciences, 71 Dudley Rd. New Brunswick, NJ 08901
2. Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556
3. Escanaba Lake Research Station, Office of Applied Science, Wisconsin Department of Natural Resources, 3110 Trout Lake Station Drive, Boulder Junction, WI 54512
4. Cary Institute of Ecosystem Studies, Millbrook, NY 12545
5. School of Resource and Environmental Management, Simon Fraser University, Burnaby, BC, Canada, VAS 1S6
6. Department of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ 08901

Abstract

Measuring fishing effort is one important element for effective management of recreational fisheries. Traditional intensive angler intercept survey methods collect many observations on a few water bodies per year to produce highly accurate estimates of fishing effort. However, scaling up this approach to understand landscapes with many systems, such as lake districts, is problematic. In these situations, spatially extensive sampling might be preferable to the traditional intensive sampling method. Here we validate a model-based approach that uses a smaller number of observations collected using multiple methods from many fishing sites to estimate total fishing effort across a fisheries landscape. We distributed on-site and aerial observations of fishing effort across 44 lakes in Vilas County, Wisconsin and then used generalized linear mixed models (GLMMs) to account for seasonal and daily trends as well as lake-specific differences in mean fishing effort. Estimates of total summer fishing effort predicted by GLMMs were on average within 11% of those produced by traditional mean expansion. These estimates required less sampling effort per lake and can be produced for many more lakes per year. In spite of the higher uncertainty associated with model-based estimates

from fewer observations, the improvements associated with the addition of only three aerial observations per lake highlighted the potential for improved precision with relatively few additional observations. Thus, the combination of GLMMs and extensive data collection from multiple sources could be used to estimate fishing effort in regions where intensive data collection for all fishing sites is infeasible, such as lake-rich landscapes. By using these methods of extensive data collection and model-based analysis, managers can produce frequently updated assessments of system states, which are important in developing proactive and dynamic management policies.

Introduction

Recreational fisheries are widespread and socioeconomically important, with about 118 million estimated participants in North America, Europe, and Oceania (Arlinghaus et al., 2015; Tufts et al., 2015). Inland and marine recreational fisheries are responsible for substantial removal of biomass, but in many systems, insufficient data are available to make proactive management decisions with the goal of maintaining sustainable harvest (Cooke and Cowx, 2004; Ihde et al., 2011). In addition, these fisheries are frequently open-access, leaving them particularly vulnerable to overfishing (Cooke and Cowx, 2004; Cox et al., 2002; Post and Parkinson, 2012). Anglers exhibit heterogeneous preferences, which leads them to adjust the location and intensity of their fishing effort in response to changing conditions. This complicates managers' ability to predict fish population dynamics (Carruthers et al., 2018; Wilson et al., 2020). Successful management of recreational fisheries therefore requires understanding fishing effort dynamics across different spatial and temporal scales.

Recreational fisheries are diverse in their spatial extent; their distribution across the landscape; and their availability of catch, effort, and harvest data (FAO, 2012; Kaemingk et al., 2019). Different systems therefore rely on different methods for quantifying fishing effort dynamics, which can include intensive and/or extensive observations of water bodies or access points. The number of water bodies surveyed depends on the abundance of water bodies present in the region as well as the budget limitations of the managing agency (e.g. Cass-Calay and Schmidt, 2009; Chizinski et al., 2014; Malvestuto et al., 1978). Intensive data collection on relatively few locations permits more in-depth sampling of these locations over a wide range of conditions. For example, access point creel surveys assign clerks to select water bodies or access points for stratified-random shifts over much of the year. During these shifts, clerks interview anglers and collect instantaneous counts of angler effort (Newman et al., 1997; Pollock, 1994). For landscapes where water bodies are relatively scarce, intensive data collection satisfactorily balances costs of data collection with accuracy of fishing effort and catch rate estimates. However, intensive data collection regimens can also leave many water bodies with no available data describing fishing effort, catch rates or harvest (Post et al., 2002). Many fisheries landscapes could therefore benefit from extensive data collection, where fewer observations are collected per site, but more water bodies or access points are surveyed (Beard et al., 2011). Fisheries already applying these methods tend to rely on multiple data sources to find the right balance between collecting sufficient observations per site while also surveying as many sites as possible (e.g. Steffe et al., 2008). In contrast, many fisheries that have historically been classified as "small scale" are surveyed through intensive methods in spite of their large spatial extent and/or their high number of access points or fishing sites, such as lake districts (Deroba et al., 2007) and river systems (West and Gordon, 1994). The pool of harvesters within a recreational fisheries landscape is mobile and heterogeneous, and their fishing effort dynamics cannot always be understood by treating small water bodies and fishing sites as independent fisheries (Matsumura et al., 2017; Martin et al., 2017). Many of these fisheries landscapes therefore benefit from a more extensive form of data collection and the integration of multiple data sources (e.g. Smallwood et al., 2012, Askey et al., 2018).

Redistributing data collection to sample all water bodies or access points is not a trivial issue, particularly in lake-rich landscapes or for very large water bodies. For large water bodies with many access points, roving creel survey methods are used to cover more area (Roop et al., 2018; West and Gordon, 1994). Additional extensive survey methods include the use of aerial surveys (Askey et al., 2018; Smucker et al., 2010), cameras (van Poorten et al., 2015), and vehicle counters (Simpson, 2018; van Poorten and Brydle, 2018), often in combination with intensive creel methods (Hartill et al., 2016; van Poorten and MacKenzie 2020). However, when adapting these mixed methods for a particular system, it will not always be possible to produce data compatible with design-based estimates of fishing effort. Traditional methods of estimating fishing effort rely on specific creel designs intended to accommodate variation in fishing effort by temporal strata, such as month or day of the week. Mean effort of a stratum is a mean of means: the mean of daily total effort means within the stratum (Newman et al., 1997). This mean expansion process leverages the central limit theorem to allow Gaussian error propagation to estimate confidence intervals around total fishing effort estimates (Särndal et al., 1978). Disparate systems use different creel designs to achieve this goal (e.g. Chizinski et al., 2014; Lockwood and Rakoczy, 2005; Smallwood et al., 2012), and they are difficult to adapt to non-standard data from supplemental sources.

In contrast, model-based estimation of fishing effort can more easily accommodate multiple data sources and is flexible to system-specific sampling methods. An example of earlier model-based approaches includes a regression method predicting on-site estimates of total fishing effort from instantaneous observations collected by aerial surveys in British Columbia (Tredger, 1992). Askey et al. (2018) demonstrated that the previously employed regression method produced biased estimates and rigorously demonstrated the effectiveness of a generalized linear mixed model-based estimation approach using aerial surveys and on-site data collection from time-lapse cameras. Model-based approaches to estimating fishing effort across multiple fishing sites or water bodies are therefore not new methods, but they have generally been applied to test for differences in fishing effort dynamics among groups (Merten et al., 2018), or to understand ecological and fishery influences on fish growth and productivity (Varkey et al., 2018). Similar models could instead be applied to extensively collected data from multiple sources to estimate waterbody-specific fishing effort over many potential fishing sites.

Despite the availability of multiple data sources for estimating fishing effort, it is not always feasible to survey all fishing sites across a landscape. Models used to estimate total fishing effort could therefore be extended to predict angling effort based on empirical relationships between fishing effort and abiotic and biotic lake variables. Studies of stated and revealed angler preferences have already identified lake characteristics that are particularly attractive to anglers. For example, large lakes that are easily accessible and present high-quality fishing opportunities are more likely to be chosen as angling sites (Hunt, 2005; Reed-Andersen et al., 2000; Hunt and Dyck, 2011). However, anglers have heterogeneous preferences, so it is not immediately clear whether these differences in characteristics among lakes may influence the overall distribution of angling effort (Beardmore et al., 2013; Breffle and Morey, 2000; Curtis and Breen, 2016, Kane et al., 2020). Lake-specific predictors could include some of the many lake morphometric and landscape variables known to influence fishing effort either directly or indirectly through their influence on fish community composition and abundance. In a study estimating total harvest across Wisconsin, Embke et al. (2020) used generalized linear mixed models (GLMMs) with lake characteristics as predictors to estimate harvest on unobserved lakes. If lake characteristics as well as the confounding effects of weather, time of day, and seasonality are also consistent predictors of fishing effort among lakes (i.e. Deroba et al., 2007), at least coarse estimates of fishing effort at unobserved lakes can be produced based on observed lake characteristics.

We tested a model-based approach to estimating fishing effort using extensive data collected in Vilas County, Wisconsin. To accomplish this goal, we examined annual summer fishing effort predictions of GLMMs fit to three datasets. These datasets were collected using different methods that demonstrated tradeoffs between the number of observations per lake and the number of lakes surveyed (Table 1). One dataset was classified as intensive because it included many observations of fewer lakes per year. The second and third datasets were extensive because they contained fewer observations per lake, but many more lakes were surveyed each year. The third dataset additionally included aerial survey observations of the same lakes to test for the value of including a supplemental data source. We completed a series of tests using these datasets to address the following questions: 1) When fit to extensive data, can models detect annual, seasonal, and daily changes in fishing effort? 2) How do fishing effort estimates derived from extensive observations compare to those derived from intensive observations? 3) How well can models fit to extensive data predict total fishing effort on unobserved lakes? 4) How can these model-based methods be applied to predict fishing effort across a fisheries landscape?

Methods

Study area

All observations of angling effort took place in Vilas County, Wisconsin. Vilas County is part of the Northern Highlands Lake District (NHLD), a highly forested, lake-rich region known for its fishing tourism (Peterson et al., 2003). With increasing shoreline residential development and the continued effects of global climate change, the NHLD lake fisheries have shown marked changes in species composition and size structure (Christensen et al., 1996; Sass et al., 2006; G. J. A. Hansen et al., 2015; J. F. Hansen et al., 2015; Embke et al., 2019). The high density of lakes in this region means that intensive creel data are collected infrequently for each surveyed lake. If accurate estimates of fishing effort could instead be derived from extensive data collected over more lakes, managers' understanding of effort dynamics at many lakes of interest could be updated more frequently. Vilas County has 1318 lakes, of which 175 have public access points maintained by the WDNR (Wisconsin Department of Natural Resources, 2009). Since 1995, the Wisconsin Department of Natural Resources (WDNR) has conducted intensive creel surveys on 65 Vilas county lakes (Figure 1, Table 1). Intensive data collection on lakes inhabited by walleye (*Sander vitreus*) in the Ceded Territory (the northern third of Wisconsin) was initiated by the WDNR and the Great Lakes Indian Fish and Wildlife Commission (GLIFWC) in 1987 after the US Seventh Circuit Court of Appeals affirmed the off-reservation hunting, fishing, and gathering rights of Ojibwe tribal members. The WDNR annually selects among all lakes containing walleye using a stratified random design to complete adult walleye population estimates, age-0 walleye relative abundance surveys, and nine-month creel surveys. In addition, each year four "trend" lakes are selected, which are sampled every three years, and most other lakes are surveyed about once every ten years (Cichosz, 2019). The data collected from these surveys are used to manage the joint tribal spearing and recreational angling fishery for walleye in the Ceded Territory of Wisconsin (Hansen et al., 1991).

Data collection

Intensive observations of instantaneous boat counts were collected by the WDNR during 1995-2019 across 65 lakes using a stratified random survey design. On average, five Vilas County lakes were surveyed per year (Tables 1 and A1), and only lakes containing walleye were surveyed (Cichosz, 2019). Survey dates and times were stratified by month, weekend, and mornings and evenings. A creel clerk's 40-hour workweek was randomly assigned to days and times based on these strata. In general, lakes were surveyed for nine months each and visited for about 20 creel shifts per month. November, March, and April were usually omitted from sampling due to perilous ice conditions. Instantaneous counts were completed at two randomly selected times during each shift. Creel clerks circled the lake by boat, counting the number of anglers that were either actively fishing or known to be moving between fishing locations (Gilbert et al., 2013; Rasmussen et al., 1998).

For our extensive experimental creel survey, we completed on-site, instantaneous counts of fishing activity at 38 lakes in Vilas County, WI from mid-May to mid-August of 2018 and 2019 (Figure 1, Supplementary Material A1). Sixty creel shifts in 2018 and 120 shifts in 2019 were stratified by weekends and weekdays as well as by morning (5:30 to 13:30) and evening (13:30 to 21:30) shifts. We randomly assigned at least four of these shifts to each lake, with the restriction that each lake needed to be surveyed at least once on a weekend or holiday. In addition, morning and evening shifts were required to take place at each lake. During each creel shift, we completed three instantaneous boat counts at randomly selected times. If randomly selected count times were less than one hour apart, count times were re-drawn until this criterion was met. If a count was selected to take place before sunrise or after dark, the count was instead completed at sunrise or sunset, respectively, and the new count time was recorded. On average, 13 instantaneous counts were completed per lake during the 6 months total of experimental creel surveys from 2018 and 2019 (Tables A1 and A2). We completed on-site instantaneous counts of fishing effort from a boat, counting the number of fishing boats and shore anglers who were actively fishing at the count time. For each boat or shore angler observed, we recorded whether or not they were angling, the number of passengers, and whether the boats were moving or stationary. Because we counted fishing vessels while the intensive creel survey counted anglers, we converted the intensive raw counts to an approximate number of fishing boats based on the mean number of passengers per boat observed during our extensive on-site counts (μ=2.04, σ=0.95).

In addition, we completed three aerial surveys of the same 38 lakes (plus 6 others) on June 6, July 10, and July 27, 2019. Flights were scheduled based on pilot availability and weather conditions. Volunteer pilots flew a pre-planned flight path in low-wing, single-engine aircraft. The pilot circled each of the target lakes at an altitude of 760 m while the counts took place. Two passengers were present for data collection: one identifying lakes and recording counts and the second locating and counting boats. When conditions allowed, we used binoculars to identify boats containing anglers. We could not always visually identify fishing boats, so unassigned stationary or slow-moving boats were therefore probabilistically classified as fishing or non-fishing based on the proportion of fishing boats among all stationary and slow-moving boats observed during on-site counts. We observed 62% of stationary boats and 80% of slow-moving boats to be fishing during our on-site counts, so each unassigned stationary and slow-moving boat was randomly assigned a classification with a 0.62 or 0.80 probability, respectively, of being classified as a fishing boat.

Traditional mean expansion estimates of fishing effort

Mean expansion estimates of total fishing effort from intensive data compute the sum of mean fishing effort over several strata. Every month of observations makes up one level, and then each month is subdivided into weekday and weekend/holiday strata. Two counts of fishing effort were collected every shift, and these were averaged to estimate each day's mean effort. Daily mean effort was multiplied by the number of daylight hours to estimate that day's total boat hours. The mean of this daily mean total effort was then calculated separately by month and weekday strata, and the sum of these grand means estimated the lake year's total fishing effort. The standard deviation (SD) of angler counts within a stratum was completed according to Rasmussen et al., (1998), and summer fishing effort SD for each lake was calculated as the square root of the summed variance of all strata. This protocol of mean expansion has been demonstrated to accurately estimate total annual fishing effort relative to a census count (Newman et al., 1997). We calculated fishing effort from intensive data only for summer months between May and August. Seven lakes were surveyed intensively and extensively on different years. This overlap allowed us to compare the accuracy and precision of mean-expansion total summer fishing effort estimates with our model-based estimates from extensive data.

When fit to extensive data, can models detect annual, seasonal, and daily changes in fishing effort?

We modeled instantaneous boat counts as a response to the effects of lake, year, day of year, and time of day using GLMMs. We tested the fit of different distributions to our count data using the R package “fitdistrplus”(Delignette-Muller and Dutang, 2015) in R version 3.6.1 (R Core Team, 2019). Because the count data were overdispersed, we fit negative binomial regressions with a log link function. We used autocorrelation function (ACF) plots of standardized residuals to detect significant temporal autocorrelation. Random intercepts incorporated variation due to lake identity that was not accounted for in the explanatory variables (Zuur et al., 2009). By including random intercepts to accommodate lake-specific variation in fishing effort, we allowed the model to pool information across lakes in order to detect general patterns in seasonal and daily fishing effort dynamics. This model was then used to predict hourly instantaneous counts across a summer for each lake. The area under the curve of these predictions then provide estimates of annual summer fishing effort that can be compared to estimates obtained by mean expansion of intensive data.

We used two datasets, the intensive WDNR observations and the extensive experimental data, and compared the ability of GLMMs to detect changes in fishing effort on three subsets of this data: (1) the intensive observations, (2) the extensive on-site observations, and (3) our combined extensive on-site and aerial survey observations. We completed forward model selection of a pre-specified set of increasingly specific predictors by comparing Akaike Information Criterion (AIC) of candidate models. We used a ΔAIC cutoff of -2 for selecting the best-fitting model. The simplest model consisted of only a random intercept by lake. We sequentially added in effects for year, day of year, and hour of day. Seasonality and time of day are already well known predictors of fishing effort (e.g. Mann and Mann-Lang, 2020; Powers and Anson, 2016). By completing forward-selection of nested models, we were able to compare the ability of different datasets to detect increasingly granular dynamics of fishing effort. For the models fit to intensive observations, the year effect was a second random intercept. For the two extensive datasets conducted only over two years, we included a year fixed effect using a dummy variable. To aid convergence, all continuous predictor variables were centered and scaled. We fit these models using the lme4 package version 1.21 (Bates et al., 2015, p. 4). Validity of the models was assessed using the DHARMa package v.0.2.6 (Hartig, 2019), and marginal and conditional r^2 were estimated using the trigamma method with the MuMIn package v.1.43.15 (Barton, 2019).

How do fishing effort estimates derived from extensive observations compare to those derived from intensive observations?

Before comparing model-based to mean expansion predictions, we first validated that generalized linear models fit separately to each lake year of intensive data produced total fishing effort estimates comparable with those produced through mean expansion (Appendix A2, Figures A1 and A2, Tables A3 and A4). After this validation, we then tested the accuracy and precision of total summer fishing effort estimates derived from each of the candidate GLMMs fit in section 2.4. We compared predictions generated by each GLMM with the estimates calculated by mean expansion for the seven lakes surveyed in both datasets. Hourly predictions of instantaneous boat counts from May 1 to August 31 for these lakes were obtained by predicting boat counts at each daylight hour of each day. Continuous prediction variables were centered and scaled according to the mean and standard deviation of the original fit data. Predictions for all models and datasets were produced for all daylight hours of summer, conditional on a mean year effect using the merTools v.0.5.0 R package (Knowles and Frederick, 2019). The area under the curve of each lake's summer predictions was then calculated using the trapezoidal rule, which produced an estimate of total summer fishing effort for each lake. By bootstrapping the model predictions for 5,000 iterations, we obtained a mean estimate of total fishing effort as well as upper and lower 95% prediction intervals. This process was repeated for each of the candidate models. These prediction intervals of model-based estimates of fishing effort were then compared to fishing effort estimates calculated through mean expansion of intensive data. To summarize correspondence between predicted and observed fishing effort for each dataset and model, we compared indices of relative accuracy and precision (I_{RA} and I_{RP}, defined below) of each model's predicted total summer fishing effort versus expanded mean estimates as in Steffe et al. (2008). Some lakes were intensively surveyed over several years. For these lakes, we compared model-based total effort estimates to the mean of all years' mean expansion estimates. The I_{RA} specifies the similarity of two estimates relative to the magnitude of the estimate of interest. A positive I_{RA} indicates that the model-based estimate is higher than that of the mean expansion by some proportion of its overall value, while a negative value indicates a lower estimate.

$$I_{RA} = \frac{GLMM\ estimate - Mean\ expansion\ estimate}{Mean\ expansion\ estimate} \times 100$$

The I_{RP} describes the similarity of each estimates' relative standard error (RSE) as a percentage of the RSE of the estimate of interest. A positive I_{RP} value indicates that the model-based estimate is more precise than that of the mean expansion, or in other words, its standard error is a smaller proportion of its estimate.

$$RSE = \frac{SE_{Estimate}}{Estimate} \times 100$$

$$I_{RP} = \frac{RSE_{Mean\ expansion} - RSE_{GLMM\ estimate}}{RSE_{Mean\ expansion}} \times 100$$

Mean I_{RA} and I_{RP} were then calculated for all lakes surveyed intensively and extensively.

How well can models fit to extensive data predict total fishing effort on unobserved lakes?

We chose the most accurate predictive model from section 2.5 and added covariates describing lake characteristics. We chose variables representing landscape predictors of boating density as described by Hunt et al. (2019). Hunt et al. (2019) modeled the distribution of boating activity in Ontario, Canada as a function of lake surface area, accessibility, human development, and fishing quality. We restricted ourselves to data that were easily obtained for all lakes in a fisheries landscape. Lake surface area is a well-established predictor of fishing effort (e.g. Hunt, 2005), and it is available for all Wisconsin lakes. We also had access to lake-specific availability of public boat ramps and presence of walleye, a popular target species. Each of these variables were obtained from the WDNR lake database. Distance from a resident pool of anglers, either from a nearby urban center or from lake residents, has also been demonstrated to predict fishing effort (Hunt et al., 2011; Wilson et al., 2020). However, given the low and relatively homogeneous population density of Vilas County (Peterson et al., 2003; U.S. Census Bureau, 2010), we judged housing density of the lakeshore to be a more influential source of nearby anglers. We calculated building density (buildings per km shoreline) within 200 m of each lake's shoreline using GIS data obtained from the WDNR and Vilas County. As an additional measure of accessibility, distance to the nearest secondary road was calculated as Euclidean distance from the centroid of a lake to the closest point of the road. Latitude and longitude of each lake was obtained from the WDNR 24K Hydro Geodatabase ("24K Hydro Full Geodatabase for Download," 2017), and road data came from the United States Geological Survey National

Transportation Dataset for Wisconsin (“USGS National Transportation Dataset Downloadable Data Collection,” 2017). Continuous variables were scaled and centered. These models were fit as described in section 2.4, and p-values were estimated based on Wald tests with the null hypothesis that the predictors have no effect on fishing effort and an alpha=0.05.

Models’ ability to predict total effort on unobserved lakes was tested using leave-one-group-out (LOGO) cross validation for models fit to intensive and extensive datasets. All observations from each lake were iteratively removed from the dataset, the models were refit, and the missing values predicted. These predictions were bootstrapped for 5000 iterations to obtain upper and lower 95% prediction intervals for the effort estimates. The I_{RA} and I_{RP} of these estimates were then estimated relative to those produced by mean expansion of intensive data.

How can these methods be applied to predict fishing effort across a fisheries landscape?

The best-performing predictive GLMM was used to estimate total summer fishing effort across all lakes and years surveyed either intensively or extensively in Vilas County. We fit the model to the combined intensive and extensive datasets, including random lake and year effects and fixed effects of weekend, day of year, and a dummy variable indicating the survey method. A full summer of fishing effort was then predicted for each lake over each year represented in the full combined dataset. We obtained 95% prediction intervals by bootstrapping the model predictions for 5000 iterations. Predictions were completed for 100 lakes over 25 years.

Results

When fit to extensive data, models detect presence and shape of annual, seasonal, and daily changes in fishing effort, but underestimate their magnitude.

The best-fit models included a year effect and quadratic effects of day of year and hour of day, which suggests that seasonal and daily patterns of fishing effort were detected by models fit even with few observations per lake (Table 2). The quadratic effect of time of day was the best fitting of all of the functional forms tested for this variable (Tables A5-A7). Weekends and holidays had a consistently positive effect on fishing effort for all datasets. However, the models fit to the intensive dataset were the only models to detect significant quadratic effects of day of year and hour of day on fishing effort (Tables A8-A10). Therefore, while including annual, daily, and hourly effects improved model fit for all of the data sets, it was only the annual and weekend effects that were detectable in the models fit to extensive data. Fixed effects such as day of year, weekend/weekday, and hour of day, explained very little variance in fishing effort (Table 3). Although lake and year random effects consistently explained around 40% of the variance in fishing effort, marginal r^2 values for hourly and daily fixed effects were very low, indicating that they explained < 5% of the variance in instantaneous fishing effort.

Models fit to extensive data produce similar estimates to mean expansion of intensive data, with some reduction in accuracy and precision.

With the exception of Irving Lake (IV), all models fit to the extensive data produced fishing effort estimates with prediction intervals that overlapped with those produced by mean expansion of intensive data (Figure 2). These models all produced mean estimates of fishing effort within 20% of the value of those produced by mean expansion of intensive data (Table 4). The best performing model for the extensive dataset, which included day of year and weekend fixed effects, produced estimates that were, on average, within 11% of the mean expansion estimate. As expected, when the models were fit to intensive data, they produced estimates of fishing effort that were nearly identical to those produced by mean expansion (Table 4, Figure 2).

On an individual lake basis, the effects on accuracy of increasing model complexity were relatively subtle and depended on lake identity. Fishing effort on Irving Lake (IV), for example, was continuously underestimated by all models fit to extensive data. Estimates for Little Arbor Vitae Lake (LV), however, were quite accurate for simple models but became more negatively biased as more parameters were added. Note the differences in total fishing effort predictions for this lake between Figures 2A and 2D. The addition of aerial survey data tended to marginally improve the mean accuracy of predictions for all lakes. More notably, aerial survey data on average improved the precision of fishing effort estimates as measured by I_{RP} (Table 4). Prediction intervals of model estimates based only on on-site extensive observations tended to be, on average, 7 to 10 times wider than the confidence intervals associated with mean expansion. Adding only 3 aerial observations per lake reduced the average width of estimate prediction intervals by nearly half. This improvement in precision suggests that a moderate number of additional samples could result in a substantial reduction in uncertainty associated with these estimates of fishing effort. An exaggerated version of this change can be seen in the predictions for Oxbow Lake (OB), on which fewer on-site observations were recorded. When three aerial observations were added for this lake, the span of the estimate's prediction interval decreased from a width of 16,147 boat hours to 7,724 boat hours, or over 50% (Figure 2C).

Intensive and extensive datasets were collected on different years, potentially limiting our ability to compare estimates of fishing effort. To investigate the influence of year effects on our estimates, we calculated estimates of total fishing effort for each year surveyed using our best-performing model. Fishing effort estimates varied substantially between years, especially for Little Arbor Vitae and Oxbow lakes (Figure 3). These two lakes had produced the least accurate model-based predictions conditional on a mean year effect, but for each of these lakes, the total effort prediction produced for one year was substantially closer to the mean expansion estimates. Much of the difference between mean expansion and model-based fishing effort estimates could therefore be a result of the mismatch in years between intensive and extensive sampling.

Model-based predictions of fishing effort on out-of-sample lakes showed mixed performance.

Predicting fishing effort for specific unobserved lakes required adding covariates describing lake characteristics that may influence fishing effort. Adding these lake variables caused marked changes to the model's conditional and marginal r^2 values (Table 5). Although the fixed effects in GLMMs predicting fishing effort from year, seasonal, and daily effects explained only around 5% of the variance in fishing effort, fixed effects in models containing lake variables explained between 20 and 30%. Because these lake variables took over some of the explanatory ability previously held by the random effects, these models could predict at least a portion of the variation in out-of-sample lakes, i.e. lakes without their own random intercept.

The effect size and significance of these lake variables depended on the dataset to which the model was fit (Table 5). Lake area had a significant positive effect on instantaneous fishing effort in models fit to all three datasets. Distance from lake to the nearest secondary road had no significant effect in any models. In the model fit to intensive data, all lake variables with the exception of distance to road and walleye presence have a significant effect on fishing effort. In the model fit to extensive data, however, lake area and walleye presence were the only significant predictors.

The accuracy of the total fishing effort predictions produced during LOGO cross validation were mixed (Figure 4). On average, the model fit to the extensive dataset containing aerial survey data produced estimates of fishing effort within 11% of those produced by mean expansion (Table 6). However, this small I_{RA} value was largely due to the very high predictions for Black Oak Lake (BK) and the very low predictions for Little Arbor Vitae (LV) offsetting each other. Model-based predictions of fishing effort were similar to the mean expansion estimates for Irving (IV), Birch (BH), Oxbow (OB), and White Birch (WB) lakes. However, this model produced much less accurate predictions for Allequash (AQ), Black Oak, and Little Arbor Vitae lakes. These results could have stemmed from two problems: 1) no lake-specific random intercept was available for the out-of-sample lakes, or 2) the selected lake variables were inconsistent predictors of fishing effort.

To evaluate these two options, the LOGO cross validation process was repeated while retaining the aerial survey observations for the "out-of-sample" lake. This process simulated the scenario of predicting fishing effort based on limited observations as well as lake variable predictors. Retaining these observations, however, did not substantially improve the predictions of total fishing effort (Figure S4). The models fit to the intensive dataset had to be simplified due to an upper limit on computation time. Rather than including both year and daily covariates, the model included only a year random effect, in addition to the lake random effect and lake characteristics that were included in the other models. Out-of-sample predictions of models fit to intensive data tended to reflect those produced by extensive data, with the exception of Irving Lake (IV), where these predictions were much closer to the mean expansion value.

Model-based methods can integrate multiple data sources to predict fishing effort across a fisheries landscape.

By fitting a GLMM to the combined intensive and extensive datasets, we could fit a random intercept to each lake and year surveyed and then predict total summer fishing effort across all lakes for each of the years represented in the datasets. Average hourly fishing effort is highly heterogeneous across the county (Figure 5A, Table S11). Several lakes stood out as having exceptionally high mean hourly fishing effort. For example, Lac Vieux Desert and Little Saint Germain Lake had 603% and 518% higher effort, respectively, than the mean. In addition, while fishing effort varied by year, no trend in overall fishing effort was evident (Figure 5B). Fishing effort in 1995, however, was very high compared to other years.

Discussion

Extensive data collection from multiple data sources is an effective tool for managers to understand fishing effort dynamics across a fisheries landscape. A model-based approach to analyzing this data allows managers to leverage multiple sources of extensive fishing effort data available within their system. By relying on extensively collected data, managers can estimate total fishing effort for many more fishing sites or water bodies than would be possible under an intensive sampling regimen. Further coverage of fisheries landscapes by spatially extensive approaches could be achieved through supplemental data sources such as aerial surveys, camera traps, and drones. With further understanding of predictors of lake use, out-of-sample estimates of fishing effort can further improve landscape coverage.

Evaluating the success of extensive data collection for model-based estimates

On average within the seven lakes evaluated, a model incorporating the effects of lake identity, year, day of year, and weekends predicted total summer fishing effort estimate values within 11% of the value of those obtained by mean expansion. Because the extensive dataset contained fewer observations per lake, some reduction in accuracy was expected. Further, the intensive and extensive observations took place on different years. We therefore remain encouraged that estimation methods using much less data produced similar results to data-rich mean expansion. Mean differences in accuracy among the seven lakes surveyed intensively and extensively were primarily driven by a tendency to underestimate fishing effort on Irving and Little Arbor Vitae lakes and to overestimate fishing effort on Oxbow Lake. The underestimation of fishing effort for Irving Lake highlighted an important consideration for the use of extensively collected data. By chance, two out of four of our experimental creel survey shifts at this lake took place during inclement weather. As a result, the mean instantaneous boat counts collected for this site were not representative of typical fishing effort, and these predictions showed no overlap of prediction intervals with those of mean expansion. When fishing effort estimates were based only on aerial survey data, which by necessity took place during fair weather, predictions of a simple GLMM were very similar to those of mean expansion of intensive data (Figure S3). The effects of poor weather could be accounted for in future applications by including a covariate for severe weather effects in the GLMM. Weather conditions did not obviously influence observations on Little Arbor Vitae, but a higher variation in total annual effort for this large, busy lake may have contributed to the reduced accuracy and precision of its model-based total fishing effort estimates.

Oxbow Lake produced fishing effort estimates with extremely wide prediction intervals. Only 6 instantaneous counts of fishing effort (3 on-site, 3 aerial) took place on this lake, less than half the number of observations collected for other lakes, which likely explains the discrepancy in total effort estimates. Although it was only possible to evaluate predictions for a small number of lakes, these examples demonstrate some of the strengths and limitations of our spatially extensive, model-based method. An extensive data collection scheme can produce reasonably accurate estimates of total fishing effort, but lake specific fishery characteristics and chance conditions during the survey will influence the optimal distribution of observations.

Our results highlight the tradeoffs that managers face in designing surveys to estimate lake-specific fishing effort. For landscapes where potential fishing sites are numerous, conducting extensive rather than intensive surveys may allow improved understanding of fishery dynamics across a broader scale. If, for example, an agency is limited to 500 observations for one summer, there are tradeoffs to consider when deciding how many lakes over which to spread those observations. These data could be used to obtain a highly accurate estimate for three lakes by following the traditional mean expansion protocol. In this case, each of the three lakes would be surveyed on 80 days of the summer with 2 instantaneous boat counts on each day (i.e., 3 lakes x 80 days x 2 observations per day = 480 observations). Alternatively, the agency could survey 31 lakes, spending 8 days surveying each one and completing two instantaneous fishing effort counts per day (i.e. 31 lakes x 8 days x 2 observations per day = 496 observations). Based on our results, transitioning from an intensive sampling regime to extensive sampling should result in, on average, a 3x increase in the width of the prediction intervals, but, in this example, a more than order of magnitude increase in the total number of lakes for which effort estimates are available. The acceptability of these tradeoffs in accuracy and precision associated with greater water body coverage will depend on the management priorities for the region in question.

Some limitations exist in our ability to compare our estimates of fishing effort from extensive data collection to traditional mean expansion of intensive data. When evaluating the accuracy of model-based total fishing effort predictions, we compared prediction intervals for an average survey year with the confidence intervals of the expanded mean total effort calculations. There was no way to account for the effect of the year of the intensive survey when calculating indices of relative abundance and precision, and year effects appear to be the reason for much of the difference in total fishing effort estimates. An additional design-related limitation is the relatively small number of lakes available for comparison of model-based with mean-expansion total effort estimates. Our summary statistics of I_{RA} and I_{RP} generalize the accuracy and precision of estimates within the seven lakes surveyed intensively and extensively, but we have no way of knowing the accuracy and precision of total fishing effort estimates for the other 31 lakes that were extensively surveyed. We can, however, compare our methods and results with those of Askey et al. (2018). Askey et al. (2018) rigorously validated the use of GLMM-based estimates of fishing effort with different sample sizes selected from a large dataset collected by aerial surveys and time-lapse cameras. The smallest sample sizes tested in their article were 10 and 20 observations. Within our limited selection of lakes with extensive and intensive data available, we found similar mean percent inaccuracies for our total effort estimates.

Opportunities for further landscape coverage

Total fishing effort estimates can be improved by integrating supplemental data sources, such as aerial surveys. By including only three additional aerial observations per lake, we substantially improved the accuracy and precision of our estimates. Even without including on-site observations, a small number of aerial observations per lake produced reasonably accurate, if coarse, estimates of total fishing effort (Figure S3). Aerial surveys are ideal for measuring the distribution of fishing effort across many lakes. This method is particularly useful for surveying fisheries with a large spatial extent, such as lake districts (Askey et al., 2018; Hunt et al., 2019; Tredger, 1992), major river systems, (Sindt, 2012) and marine and Great Lakes fisheries (Lockwood and Rakoczy, 2005; Zellmer et al., 2018). Despite its strengths, this method may be too expensive to implement consistently in many fisheries systems and can be limited by severe weather conditions.

Traffic counters and boat launch cameras have also been used to quantify fishing effort and boat traffic (Hunt and Dyck, 2011; Simpson, 2018; van Poorten et al., 2015; van Poorten and Brydle, 2018). These methods can passively collect effort data without the need for creel clerks, but cameras and counters are still expensive and prone to vandalism (van Poorten et al., 2015). The use of drones in fisheries science has been advocated (Kopaska, 2014), and they have been successfully used for identifying derelict or illegal fishing gear (Bloom et al., 2019), counting fish in shallow rivers (Tyler et al., 2018), and monitoring marine protected areas (Miller et al., 2013). Privacy concerns and aviation laws, however, complicate their use in monitoring angling activity for inland fisheries (Duncan, 2016; Lally et al., 2019). Although each of these methods has costs and benefits, they are all potentially fruitful supplemental data sources for model-based estimates of angler effort for different fishery systems.

As we demonstrated, fishing effort data collected through an extensive sampling scheme from multiple sources can be used to understand differences in fishing effort across a broad spatial and temporal scale. Through two years of extensive data collection using on-site and aerial observations, we added coverage of 44 lakes to the combined intensive and extensive fishing effort dataset describing Vilas County. Based on the year effects estimated from 25 years of intensive data, we were able to predict total fishing effort for all lake-year combinations. Although the empirical data does not exist to validate these estimates, this analysis remains a useful demonstration for the potential of extensive data collection and GLMM-based analysis for estimating fishing effort across a lake-rich landscape. Further annual extensive data collection would quickly expand this coverage, as well as allow for the direct comparison of fishing effort between years on a broader scale. These data also have promise for detecting seasonal and daily patterns in fishing effort, which can assist fisheries managers in choosing optimal times for management interventions.

As we found, however, a granular understanding of shifts in angler effort dynamics requires more data than we collected in our extensive sampling scheme. By allowing partial pooling of observations between lakes using lake random intercepts, some generalizable patterns were observed, but more observations per year may be needed to estimate the magnitude of seasonal and daily effects. Alternatively, different lakes may have different diel and seasonal fishing effort patterns. Although the extensive creel survey included fewer lakes than the intensive survey, a wider variety of lakes were surveyed, including lakes with no walleye population, no boat ramp, and lakes with smaller surface areas. Because of this greater variation in lake characteristics, concurrent differences in diel and seasonal fishing effort patterns may have been washed out to non-significance when the GLMMs were fit. In this case, more intensive data collection with more observations per lake may be required to understand lake-specific seasonal and daily patterns. A hypothetical fisheries manager is therefore left to decide whether their goals are best served by investing their limited resources in extensive data collection over a wider spatial extent or intensive data collection within a limited number of systems.

This question of appropriate tradeoffs could be sidestepped if managers could effectively predict fishing effort for unobserved lakes based on lake characteristics. We attempted to predict unobserved fishing effort using easily obtained data, with mixed results. Model predictions overlapped with mean expansion estimates for five out of the seven lakes tested, but total fishing effort for the other two were substantially over- or underestimated. Lakes associated with inaccurate predictions did not have any obvious characteristics in common that could explain this discrepancy. These results could be explained by our use of only easily obtained predictor variables, or they could be an indication that lake characteristics are not consistent, linear predictors of lake-specific fishing effort. We chose lake variables that aligned with characteristics found to predict recreational boating density by Hunt et al. (2019), including lake surface area, walleye presence, and indices of human development and accessibility. Differences in sampling frame between our intensive and extensive data collection resulted in differences in parameter values between models fit to different datasets. For example, intensive data collection in Wisconsin takes place only on lakes containing walleye. Because no contrast was available for this parameter, no walleye effect could be tested. In summer, walleye are also almost exclusively available to boat anglers, potentially explaining the presence of a boat ramp effect in the intensive but not the extensive dataset. Distance to secondary road had no effect on instantaneous fishing effort in any dataset. Most likely, this result stems from measuring distance to road from the centroid of each lake. This metric does not account for the location of boat launches, so the nearest secondary road as measured here may still be inconveniently far away from any access points. Potential explanations for the absence of a building density effect in the extensive data are less clear. The lakes surveyed for both datasets had a similar range in building density values (0-70 buildings per km in the intensive data and 0-80 buildings in the extensive data). It is possible that, similar to diel and seasonal patterns, housing density has a different effect on fishing effort for different lakes. Not all lake residents are interested in fishing, and the presence of some building types such as resorts may be a better predictor of resident fishing effort than the presence of family homes.

Indicators of fishing quality such as angler catch rates or fish population estimates, rather than indirect measurements of accessibility, may improve the predictive ability of these models, but these data are labor-intensive to produce and therefore did not exist for every lake in our extensive dataset. By applying model-based fishing effort predictions over every lake- year combination in the combined intensive and extensive datasets, we identified a handful of extremely high fishing effort lakes, which allowed us to explore potential commonalities between them. The primary characteristic these lakes had in common was their surface area; the lakes with highest mean fishing effort ranged from 350 to over 1600 ha in surface area (Table S11). In contrast, no obvious correlation was found between fishing effort and population abundance or catch rates of popular target species. However, very high fishing effort lakes all tended to have moderate, rather than high or low, catch rates for panfish and muskellunge (Figures A5-A8). Most likely, predicting fishing effort based on lake characteristics would require accounting for nonlinear responses and interactions of lake characteristics, potentially using nonparametric methods such as random forests (e.g. van Poorten et al, 2013). Although out-of-sample predictions of fishing effort were not consistently accurate, we argue that extensive data collection for GLMM-based estimates of total fishing effort is a promising approach for understanding effort dynamics in highly distributed and/or data poor fisheries.

Applications to fisheries management

Our modeling approach proved effective for predicting angler effort across a fisheries landscape; however, other metrics derived from traditional angler intercept surveys, such as angler catch rates and estimates of total catch, are also important for fisheries management. That said, our approach could compliment existing efforts to address these important, additional aspects of fisheries. For example, recent research by Embke et al., (2020) used GLMMs to produce recreational harvest estimates for 267 lakes that were surveyed intensively as well as all unobserved inland lakes across Wisconsin based on abiotic variables and an angler access metric.

Coarse estimates of fishing effort based on spatially extensive observations could further refine harvest estimates on these otherwise unobserved lakes. Additional catch and harvest data can also be collected during spatially extensive sampling of fishing effort through angler intercept interviews (Iwicki et al., in prep). Perhaps most importantly, the different levels of variability associated with fishing effort and harvest estimates based on extensively collected data can identify lakes of greater uncertainty where additional sampling resources should be directed. For example, high-effort and high-variance lakes such as Little Arbor Vitae likely need to be allocated more sampling effort than lakes such as White Birch (Fig. 3).

In addition to its applicability to data-poor fisheries, a model-based approach to generating fishing effort estimates from fewer observations at more fishing sites could be a practical tool for managers who want to implement ecosystem-based management strategies that can respond to fast and slow changes across a fisheries landscape (*sensu* Walker et al., 2012). A transition from a one-size-fits-all management policy to a more diverse set of policies may contribute to a more persistent and resilient fisheries system (Carpenter and Brock, 2004; van Poorten and Camp, 2019). These policies would ideally be dynamic across space and time, which requires faster feedback from data collection describing how interventions have affected fishing effort, catch, and harvest. Although implementing highly dynamic and lake-specific policies is probably an unrealistic goal in lake-rich fisheries, tailored management of different categories of lakes may simultaneously improve system resilience and angler satisfaction by accommodating the preferences of heterogeneous groups of anglers. Strategic collection of fishing effort data over many lakes may therefore be an effective bridge between one-size-fits all policy and model-based implementation of diverse and dynamic policies.

Acknowledgements

This work was supported by the National Science Foundation under grant number 1716066. We would like to thank the University of Notre Dame Environmental Research Center for logistical and financial support. We also thank the Wisconsin Department of Natural Resources for their assistance, including but not limited to data sharing, lake selection, and project design. We thank all team members who assisted with data collection: John Caffarelli, Francie Fink, Amanda Kerkhove, Alex Kisurin, Ana Lopez, Joe Pitti, Alberto Salazar, Mary Solokas, and Danny Szydlowski. Many thanks to Brittni Bertolet, Marco Janssen, Sunny Jardine, Camille Mosley, Randi Notte, Alex Ross, and Dane Whittaker for their helpful suggestions, statistical assistance, and help with aerial surveys. We wish to thank Lighthawk and their volunteer pilots: Bill Menzel and David Weissman for making the aerial surveys possible. We would also like to thank the Office of Advanced Computing at Rutgers for the use of their computing cluster resources and the Manasquan River Marlin and Tuna Club for their financial support. G.G. Sass was also supported by United States Fish and Wildlife Service, Federal Aid in Sportfish Restoration funding, and the Wisconsin Department of Natural Resources.

References

[dataset] 24k Hydro Full Geodatabase for Download [WWW Document], 2017. URL https://data-wi-dnr.opendata.arcgis.com/datasets/0128cce2c06342218725f1069031a4fa (accessed 4.8.20).

Arlinghaus, R., Tillner, R., Bork, M., 2015. Explaining participation rates in recreational fishing across industrialised countries. Fish. Manag. Ecol. 22, 45–55. https://doi.org/10.1111/fme.12075

Askey, P.J., Ward, H., Godin, T., Boucher, M., Northrup, S., 2018. Angler Effort Estimates from Instantaneous Aerial Counts: Use of High-Frequency Time-Lapse Camera Data to Inform Model-Based Estimators. North Am. J. Fish. Manag. 38, 194–209. https://doi.org/10.1002/nafm.10010

Barton, K., 2019. Mu-MIn: Multi-model inference., R package.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67, 1–48. https://doi.org/doi:10.18637/jss.v067.i01

Beard, T.D., Arlinghaus, R., Cooke, S.J., McIntyre, P.B., De Silva, S., Bartley, D., Cowx, I.G., 2011. Ecosystem approach to inland fisheries: research needs and implementation strategies. Biol. Lett. 7, 481–483. https://doi.org/10.1098/rsbl.2011.0046

Beardmore, B., Haider, W., Hunt, L.M., Arlinghaus, R., 2013. Evaluating the Ability of Specialization Indicators to Explain Fishing Preferences. Leis. Sci. 35, 273–292. https://doi.org/10.1080/01490400.2013.780539

Bloom, D., Butcher, P.A., Colefax, A.P., Provost, E.J., Cullis, B.R., Kelaher, B.P., 2019. Drones detect illegal and derelict crab traps in a shallow water estuary. Fish. Manag. Ecol. 26, 311–318. https://doi.org/10.1111/fme.12350

Breffle, W.S., Morey, E.R., 2000. Investigating Preference Heterogeneity in a Repeated Discrete-Choice Recreation Demand Model of Atlantic Salmon Fishing. Mar. Resour. Econ. 15, 1–20. https://doi.org/10.1086/mre.15.1.42629285

Carpenter, S., Brock, W., 2004. Spatial Complexity, Resilience, and Policy Diversity: Fishing on Lake-rich Landscapes. Ecol. Soc. 9. https://doi.org/10.5751/ES-00622-090108

Carruthers, T.R., Dabrowska, K., Haider, W., Parkinson, E.A., Varkey, D.A., Ward, H.G.M., McAllister, M.K., Godin, T., van Poorten, B.T., Askey, P.J., Wilson, K.L., Hunt, L.M., Clarke, A.D., Newton, E., Walters, C., Post, J.R., 2018. Landscape scale social and ecological outcomes of dynamic angler and fish behaviours: processes, data, and patterns. Can. J. Fish. Aquat. Sci. https://doi.org/10.1139/cjfas-2018-0168

Cass-Calay, S.L., Schmidt, T.W., 2009. Monitoring changes in the catch rates and abundance of juvenile goliath grouper using the ENP creel survey, 1973–2006. Endanger. Species Res. 7, 183–193. https://doi.org/10.3354/esr00139

Chizinski, C.J., Martin, D.R., Pope, K.L., Barada, T.J., Schuckman, J.J., 2014. Angler effort and catch within a spatially complex system of small lakes. Fish. Res. 154, 172–178. https://doi.org/10.1016/j.fishres.2014.02.013

Christensen, D.L., Herwig, B.R., Schindler, D.E., Carpenter, S.R., 1996. Impacts of Lakeshore Residential Development on Coarse Woody Debris in North Temperate Lakes. Ecol. Appl. 6, 1143–1149. https://doi.org/10.2307/2269598

Cichosz, T., 2019. Wisconsin Department of Natural Resources 2017-2018 Ceded Territory Fishery Assessment Report (Administrative Report No. 91). Treaty Fisheries Assessment Unit Bureau of Fisheries Management, Madison, Wisconsin.

Cooke, S.J., Cowx, I.G., 2004. The Role of Recreational Fishing in Global Fish Crises. BioScience 54, 857–859. https://doi.org/10.1641/0006-3568(2004)054[0857:TRORFI]2.0.CO;2

Cox, S.P., Beard, T.D., Walters, C., 2002. Harvest Control in Open-Access Sport Fisheries: Hot Rod or Asleep at the Reel? Bull. Mar. Sci. 70, 13.

Curtis, J., Breen, B., 2016. Fisheries management for different angler types (Working Paper No. 529). ESRI Working Paper.

Delignette-Muller, M., Dutang, C., 2015. fitdistrplus: An R Package for Fitting Distributions. J. Stat. Softw. 64, 1–34.

Deroba, J.J., Hansen, M.J., Nate, N.A., Hennessy, J.M., 2007. Temporal Profiles of Walleye Angling Effort, Harvest Rate, and Harvest in Northern Wisconsin Lakes. North Am. J. Fish. Manag. 27, 717–727. https://doi.org/10.1577/M06-125.1

Duncan, J., 2016. Small Unmanned Aircraft Systems (sUAS) AC 107-2 (Advisory Circular). U.S. Department of Transportation Federal Aviation Administration.

Embke, H.S., Beard, T.D., Lynch, A.J., Zanden, M.J.V., 2020. Fishing for Food: Quantifying Recreational Fisheries Harvest in Wisconsin Lakes. Fisheries n/a. https://doi.org/10.1002/fsh.10486

Embke, H.S., Rypel, A.L., Carpenter, S.R., Sass, G.G., Ogle, D., Cichosz, T., Hennessy, J., Essington, T.E., Zanden, M.J.V., 2019. Production dynamics reveal hidden overharvest of inland recreational fisheries. Proc. Natl. Acad. Sci. https://doi.org/10.1073/pnas.1913196116

FAO, 2012. Technical Guidelines for Responsible Fisheries: Recreational fisheries, Technical guidelines for responsible fisheries. Food and Agriculture Organization of the United Nations, Rome.

Gilbert, S., King, R., Kamke, K., Beard, T., 2013. Wisconsin Department of Natural Resources creel clerk manual for lake and reservoir surveys.

Hansen, G.J.A., Carpenter, S.R., Gaeta, J.W., Hennessy, J.M., Vander Zanden, M.J., 2015. Predicting walleye recruitment as a tool for prioritizing management actions. Can. J. Fish. Aquat. Sci. 72, 661–672. https://doi.org/10.1139/cjfas-2014-0513

Hansen, J.F., Sass, G.G., Gaeta, J.W., Hansen, G.A., Isermann, D.A., Lyons, J., Zanden, M.J.V., 2015. Largemouth Bass Management in Wisconsin: Intraspecific and Interspecific Implications of Abundance Increases.

Hansen, M.J., Staggs, M.D., Hoff, M.H., 1991. Derivation of Safety Factors for Setting Harvest Quotas on Adult Walleyes from past Estimates of Abundance. Trans. Am. Fish. Soc. 120, 620–628. https://doi.org/10.1577/1548-8659(1991)120<0620:DOSFFS>2.3.CO;2

Hartig, F., 2019. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models., R package.

Hartill, B.W., Payne, G.W., Rush, N., Bian, R., 2016. Bridging the temporal gap: Continuous and cost-effective monitoring of dynamic recreational fisheries by web cameras and creel surveys. Fish. Res. 183, 488–497. https://doi.org/10.1016/j.fishres.2016.06.002

Hunt, L.M., 2005. Recreational Fishing Site Choice Models: Insights and Future Opportunities. Hum. Dimens. Wildl. 10, 153–172. https://doi.org/10.1080/10871200591003409

Hunt, L.M., Arlinghaus, R., Lester, N., Kushneriuk, R., 2011. The effects of regional angling effort, angler behavior, and harvesting efficiency on landscape patterns of overfishing. Ecol. Appl. 21, 2555–2575. https://doi.org/10.1890/10-1237.1

Hunt, L.M., Dyck, A., 2011. The Effects of Road Quality and Other Factors on Water-Based Recreation Demand in Northern Ontario, Canada. For. Sci. 57, 281–291. https://doi.org/10.1093/forestscience/57.4.281

Hunt, L.M., Morris, D.M., Drake, D.A.R., Buckley, J.D., Johnson, T.B., 2019. Predicting spatial patterns of recreational boating to understand potential impacts to fisheries and aquatic ecosystems. Fish. Res. 211, 111–120. https://doi.org/10.1016/j.fishres.2018.11.007

Ihde, T.F., Wilberg, M.J., Loewensteiner, D.A., Secor, D.H., Miller, T.J., 2011. The increasing importance of marine recreational fishing in the US: Challenges for management. Fish. Res. 108, 268–276. https://doi.org/10.1016/j.fishres.2010.12.016

Kaemingk, M., Chizinski, C., Allen, C., Pope, K., 2019. Ecosystem size predicts social-ecological dynamics. Ecol. Soc. 24. https://doi.org/10.5751/ES-10961-240217

Kane, D.S., Kaemingk, M.A., Chizinski, C.J., Pope, K.L., 2020. Spatial and temporal behavioral differences between angler-access types. Fish. Res. 224, 105463. https://doi.org/10.1016/j.fishres.2019.105463

Knowles, J., Frederick, C., 2019. merTools: tools for analyzing mixed effect regression models, R package.
Kopaska, J., 2014. Drones—A Fisheries Assessment Tool? Fisheries 39, 319–319. https://doi.org/10.1080/03632415.2014.923771
Lally, H.T., O'Connor, I., Jensen, O.P., Graham, C.T., 2019. Can drones be used to conduct water sampling in aquatic environments? A review. Sci. Total Environ. 670, 569–575. https://doi.org/10.1016/j.scitotenv.2019.03.252
Lockwood, R.N., Rakoczy, G.P., 2005. Comparison of Interval and Aerial Count Methods for Estimating Fisher Boating Effort. North Am. J. Fish. Manag. 25, 1331–1340. https://doi.org/10.1577/M04-176.1
Malvestuto, S.P., Davies, W.D., Shelton, W.L., 1978. An Evaluation of the Roving Creel Survey with Nonuniform Probability Sampling. Trans. Am. Fish. Soc. 107, 255–262. https://doi.org/10.1577/1548-8659(1978)107<255:AEOTRC>2.0.CO;2
Mann, B.Q., Mann-Lang, J.B., 2020. Trends in shore-based angling effort determined from aerial surveys: a case study from KwaZulu-Natal, South Africa. Afr. J. Mar. Sci. 0, 1–13. https://doi.org/10.2989/1814232X.2020.1740785
Martin, D.R., Shizuka, D., Chizinski, C.J., 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. https://doi.org/10.1016/j.fishres.2017.05.007
Matsumura, S., Beardmore, B., Haider, W., Dieckmann, U., Arlinghaus, R., 2017. Ecological, angler and spatial heterogeneity drive social and ecological outcomes in an integrated landscape model of freshwater recreational fisheries. bioRxiv 227744. https://doi.org/10.1101/227744
Merten, W., Rivera, R., Appeldoorn, R., Serrano, K., Collazo, O., Jimenez, N., 2018. Use of video monitoring to quantify spatial and temporal patterns in fishing activity across sectors at moored fish aggregating devices off Puerto Rico. Sci. Mar. Barc. 82, 107–117. https://doi.org/10.3989/scimar.04730.09A
Miller, D.G.M., Slicer, N.M., Hanich, Q., 2013. Monitoring, control and surveillance of protected areas and specially managed areas in the marine domain. Mar. Policy 39, 64–71. https://doi.org/10.1016/j.marpol.2012.10.004
Newman, S.P., Rasmussen, P.W., Andrews, L.M., 1997. Comparison of a Stratified, Instantaneous Count Creel Survey with a Complete Mandatory Creel Census on Escanaba Lake, Wisconsin. North Am. J. Fish. Manag. 17, 321–330. https://doi.org/10.1577/1548-8675(1997)017<0321:COASIC>2.3.CO;2
Peterson, G.D., Beard, T.D., Beisner, B.E., Bennett, E.M., Carpenter, S.R., Cumming, G.S., Dent, C.L., Havlicek, T.D., 2003. Assessing Future Ecosystem Services: a Case Study of the Northern Highlands Lake District, Wisconsin. Conserv. Ecol. 7. http://www.consecol.org/vol7/iss3/art1/
Pollock, K.H., 1994. Angler Survey Methods and Their Applications in Fisheries Management, Special Publication. American Fisheries Society.
Post, John.R., Parkinson, E.A., 2012. Temporal and spatial patterns of angler effort across lake districts and policy options to sustain recreational fisheries. Can. J. Fish. Aquat. Sci. 69, 321–329. https://doi.org/10.1139/f2011-163
Post, J.R., Sullivan, M., Cox, S., Lester, N.P., Walters, C.J., Parkinson, E.A., Paul, A.J., Jackson, L., Shuter, B.J., 2002. Canada's Recreational Fisheries: The Invisible Collapse? Fisheries 27, 6–17. https://doi.org/10.1577/1548-8446(2002)027<0006:CRF>2.0.CO;2

Powers, S.P., Anson, K., 2016. Estimating Recreational Effort in the Gulf of Mexico Red Snapper Fishery Using Boat Ramp Cameras: Reduction in Federal Season Length Does Not Proportionally Reduce Catch. North Am. J. Fish. Manag. 36, 1156–1166. https://doi.org/10.1080/02755947.2016.1198284

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rasmussen, P.W., Staggs, M.D., Beard, T.D., Newman, S.P., 1998. Bias and Confidence Interval Coverage of Creel Survey Estimators Evaluated by Simulation. Trans. Am. Fish. Soc. 127, 469–480. https://doi.org/10.1577/1548-8659(1998)127<0469:BACICO>2.0.CO;2

Reed-Andersen, T., Bennett, E.M., Jorgensen, B.S., Lauster, G., Lewis, D.B., Nowacek, D., Riera, J.L., Sanderson, B.L., Stedman, R., 2000. Distribution of recreational boating across lakes: do landscape variables affect recreational use? Freshw. Biol. 43, 439–448. https://doi.org/10.1046/j.1365-2427.2000.00511.x

Roop, H.J., Poudyal, N.C., Jennings, C.A., 2018. Assessing Angler Effort, Catch, and Harvest on a Spatially Complex, Multi-Lake Fishery in Middle Georgia. North Am. J. Fish. Manag. 38, 833–841. https://doi.org/10.1002/nafm.10179

Särndal, C.-E., Thomsen, I., Hoem, J.M., Lindley, D.V., Barndorff-Nielsen, O., Dalenius, T., 1978. Design-Based and Model-Based Inference in Survey Sampling [with Discussion and Reply]. Scand. J. Stat. 5, 27–52.

Sass, G.G., Kitchell, J.F., Carpenter, S.R., Hrabik, T.R., Marburg, A.E., Turner, M.G., 2006. Fish Community and Food Web Responses to a Whole-lake Removal of Coarse Woody Habitat. Fisheries 31, 321–330. https://doi.org/10.1577/1548-8446(2006)31[321:FCAFWR]2.0.CO;2

Simpson, G., 2018. Use of a Public Fishing Area Determined by Vehicle Counters with Verification by Trail Cameras. Nat. Resour. 09, 188–197. https://doi.org/10.4236/nr.2018.95012

Sindt, T., 2012. 2012 Ohio River Angler Survey. Ohio Department of Natural Resources.

Smallwood, C.B., Pollock, K.H., Wise, B.S., Hall, N.G., Gaughan, D.J., 2012. Expanding Aerial–Roving Surveys to Include Counts of Shore-Based Recreational Fishers from Remotely Operated Cameras: Benefits, Limitations, and Cost Effectiveness. North Am. J. Fish. Manag. 32, 1265–1276. https://doi.org/10.1080/02755947.2012.728181

Smucker, B.J., Lorantas, R.M., Rosenberger, J.L., 2010. Correcting Bias Introduced by Aerial Counts in Angler Effort Estimation. North Am. J. Fish. Manag. 30, 1051–1061. https://doi.org/10.1577/M09-193.1

Steffe, A.S., Murphy, J.J., Reid, D.D., 2008. Supplemented Access Point Sampling Designs: A Cost-Effective Way of Improving the Accuracy and Precision of Fishing Effort and Harvest Estimates Derived from Recreational Fishing Surveys. North Am. J. Fish. Manag. 28, 1001–1008. https://doi.org/10.1577/M06-248.1

Tredger, C.D., 1992. Estimation of Angler Effort using Index Boat Counts (Fisheries Technical Circular No. 94). Ministry of Environment, Lands and Parks, Fisheries Branch, Conservation Section, Victoria, B.C.

Tufts, B.L., Holden, J., DeMille, M., 2015. Benefits arising from sustainable use of North America's fishery resources: economic and conservation impacts of recreational angling. Int. J. Environ. Stud. 72, 850–868. https://doi.org/10.1080/00207233.2015.1022987

Tyler, S., Jensen, O.P., Hogan, Z., Chandra, S., Galland, L.M., Simmons, J., 2018. Perspectives on the Application of Unmanned Aircraft for Freshwater Fisheries Census. Fisheries 43, 510–516. https://doi.org/10.1002/fsh.10167

U.S. Census Bureau, 2010. 2010 Census of Population adn Housing: Population and Housing Unit Counts. U.S. Department of Commerce.

[dataset] USGS National Transportation Dataset Downloadable Data Collection [WWW Document], 2017. URL https://catalog.data.gov/dataset/usgs-national-transportation-dataset-ntd-downloadable-data-collectionde7d2 (accessed 4.8.20).

van Poorten, B.T., Brydle, S., 2018. Estimating fishing effort from remote traffic counters: Opportunities and challenges. Fish. Res. 204, 231–238. https://doi.org/10.1016/j.fishres.2018.02.024

van Poorten, B.T., Camp, E., 2019. Addressing Challenges Common to Modern Recreational Fisheries with a Buffet-Style Landscape Management Approach. Rev. Fish. Sci. Aquac. 27, 393–416. https://doi.org/10.1080/23308249.2019.1619071

van Poorten, B.T., Carruthers, T.R., Ward, H.G.M., Varkey, D.A., 2015. Imputing recreational angling effort from time-lapse cameras using an hierarchical Bayesian model. Fish. Res. 172, 265–273. https://doi.org/10.1016/j.fishres.2015.07.032

van Poorten, B.T., Cox, S.P., Cooper, A.B., 2013. Efficacy of harvest and minimum size limit regulations for controlling short-term harvest in recreational fisheries. Fish. Manag. Ecol. 20, 258–267. https://doi.org/10.1111/j.1365-2400.2012.00872.x

van Poorten, B.T., MacKenzie, C.J.A., 2020. Using Decision Analysis to Balance Angler Utility and Conservation in a Recreational Fishery. North Am. J. Fish. Manag. 40, 29–47. https://doi.org/10.1002/nafm.10377

Varkey, D.A., van Poorten, B., McAllister, M.K., Godin, T., Parkinson, E., Kumar, R., 2018. Describing growth based on landscape characteristics and stocking strategies for rainbow trout. Fish. Res. 208, 229–238. https://doi.org/10.1016/j.fishres.2018.06.015

Walker, B.H., Carpenter, S.R., Rockstrom, J., Crépin, A.-S., Peterson, G.D., 2012. Drivers, "Slow" Variables, "Fast" Variables, Shocks, and Resilience. Ecol. Soc. 17. http://dx.doi.org/10.5751/ES-05063-170330

West, R.J., Gordon, G.N.G., 1994. Commercial and recreational harvest of fish from two Australian coastal rivers. Mar. Freshw. Res. 45, 1259–1279. https://doi.org/10.1071/mf9941259

Wilson, K.L., Foos, A., Barker, O.E., Farineau, A., Gisi, J.D., Post, J.R., 2020. Social–ecological feedbacks drive spatial exploitation in a northern freshwater fishery: A halo of depletion. J. Appl. Ecol. n/a. https://doi.org/10.1111/1365-2664.13563

Wisconsin Department of Natural Resources, 2009. Wisconsin Lakes.

Zellmer, A.J., Burdick, H., Medel, I., Pondella, D.J., Ford, T., 2018. Aerial surveys and distribution models enable monitoring of fishing in Marine Protected Areas. Ocean Coast. Manag. 165, 298–306. https://doi.org/10.1016/j.ocecoaman.2018.08.027

Zuur, A., Ieno, E., Walker, N., Saveliev, A., Smith, G., 2009. Mixed Effects Models and Extensions in Ecology with R, 1st ed. Springer.

Tables

Table 1: Characteristics of the three datasets we evaluated when estimating lake-specific total fishing effort.

	Intensive dataset	Extensive dataset	Extensive dataset with aerial surveys
Sampling methods	On-site observations	On-site observations	On-site observations Aerial surveys
Number of years surveyed	25	2	2
Number of lakes surveyed	65	38	44
Mean number of lakes surveyed per year (SD)	4.9 (2.6)	21 (7.1)	29.5 (19.1)

Table 2: AIC values for each model fit to each dataset. Each model contains its listed predictors as well as all predictors listed for the models above it. Values for ΔAIC are the difference between that model's AIC and that of the model containing only a random lake effect. The best fit model for all datasets is in bold.

Model	Intensive data		On-site extensive data		On-site and aerial survey extensive data	
	AIC	ΔAIC	AIC	ΔAIC	AIC	ΔAIC
$(1 \vert Lake)$	90206		1360.1		1725.8	
$+ Year$	89883	-323	1350.2	-9.9	1713.3	-12.5
$+ Day\ of\ year$ $+ Day\ of\ year^2$ $+ Weekend$	88766	-1440	1346.6	-13.5	1708.5	-17.3
$+ \boldsymbol{Hour\ of\ day}$ $+ \boldsymbol{Hour\ of\ day^2}$	**87948**	**-2258**	**1338.9**	**-21.2**	**1700.0**	**-25.8**

Table 3: Marginal and conditional r^2 values for each model fit to each dataset. Each model contains its listed predictors as well as all predictors listed for the models above it.

Model	Intensive data		On-site extensive data		On-site and aerial extensive data	
	Marginal r^2	Conditional r^2	Marginal r^2	Conditional r^2	Marginal r^2	Conditional r^2
$(1\|Lake)$		0.36		0.38		0.39
$+Year$		0.39	0.035	0.46	0.021	0.43
$+Day\ of\ year$	0.023	0.43	0.047	0.50	0.031	0.45
$+Day\ of\ year^2$						
$+Weekend$						
	0.044	0.46	0.065	0.52	0.044	0.46
$+Hour\ of\ day$						
$+Hour\ of\ day^2$						

Table 4: Mean indices of accuracy and precision for model-based estimates of total summer fishing boat hours relative to mean expansion estimates. (N=7) Each model contains its listed predictors as well as all predictors listed for the models above it.

Model	Intensive data		On-site extensive data		On-site and aerial extensive data	
	Mean I_{RA}(SD)	Mean I_{RP} (SD)	Mean I_{RA}(SD)	Mean I_{RP} (SD)	Mean I_{RA}(SD)	Mean I_{RP} (SD)
$(1\|Lake)$	8.06 (6.48)	73.95 (5.51)	-5.50 (43.68)	-48.82 (61.24)	1.93 (42.57)	-7.67 (25.70)
$+ Year$	4.80 (12.03)	67.27 (7.44)	18.28 (58.46)	-51.57 (63.67)	11.98 (48.60)	-9.15 (25.46)
$+ Day\ of\ year$ $+ Day\ of\ year^2$ $+ Weekend$	-0.91 (11.73)	67.67 (7.24)	-8.13 (51.55)	-72.46 (67.02)	-10.86 (39.35)	-23.25 (24.22)
$+ Hour\ of\ day$ $+ Hour\ of\ day^2$	-4.58 (12.35)	69.79 (8.93)	-11.31 (46.68)	-74.82 (61.75)	-13.71 (36.11)	-26.86 (24.88)

Table 5: Parameters of a GLMM predicting fishing effort from seasonality and lake variables as fit to each dataset. Parameters with significant effects are in bold.

Model parameters	Intensive data		On-site extensive data		On-site and aerial extensive data	
	Coefficient (SE)	P value	Coefficient (SE)	P value	Coefficient (SE)	P value
Intercept	-0.66 (0.55)	0.23	-1.51 (0.39)	0.0001	-1.35 (0.31)	<0.0001
Lake area (ha)	**0.56 (0.12)**	**<0.0001**	**0.47 (0.13)**	**0.0002**	**0.50 (0.10)**	**<0.0001**
Building density	**0.25 (0.10)**	**0.01**	0.09 (0.13)	0.50	0.06 (0.10)	0.55
Boat ramp present	**0.71 (0.22)**	**0.001**	0.14 (0.42)	0.74	0.19 (0.33)	0.56
Walleye present	0.72 (0.54)	0.18	**1.40 (0.34)**	**<0.0001**	**1.23 (0.26)**	**<0.0001**
Distance to road	-0.02 (0.09)	0.78	-0.06 (0.11)	0.61	-0.10 (0.09)	0.25
Year 2018			**-0.25 (0.09)**	**0.006**	**-0.21 (0.06)**	**0.0009**
Day of year	**1.21 (0.09)**	**<0.0001**	1.06 (0.78)	0.17	0.75 (0.58)	0.27
Day of year^2	**-1.22 (0.09)**	**<0.0001**	-1.13 (0.77)	0.14	-0.85 (0.67)	0.21
Weekend	**0.47 (0.01)**	**<0.0001**	**0.21 (0.11)**	**0.05**	**0.18 (0.09)**	**0.04**
Marginal r^2	0.23		0.26		0.28	
Conditional r^2	0.43		0.34		0.35	

Table 6: Mean indices of relative accuracy and precision of out-of-sample model predictions relative to mean expansion estimates of intensive data. (N=7)

Model	Intensive data		On-site extensive data		On-site and aerial extensive data	
	Mean I_{RA}(SD)	Mean I_{RP} (SD)	Mean I_{RA}(SD)	Mean I_{RP} (SD)	Mean I_{RA}(SD)	Mean I_{RP} (SD)
$(1\|Lake) + Year$ $+ Day\ of\ year$ $+ Day\ of\ year^2$ $+ Weekend$ $+ Lake\ area$ $+ Building\ density$ $+ Boat\ ramp$ $+ Walleye\ presence$ $+ Distance\ to\ road$	-26.17 (76.46)	987.31 (400.09)	-16.16 (64.71)	42.28 (63.90)	-10.66 (58.99)	88.11 (76.00)

Figures

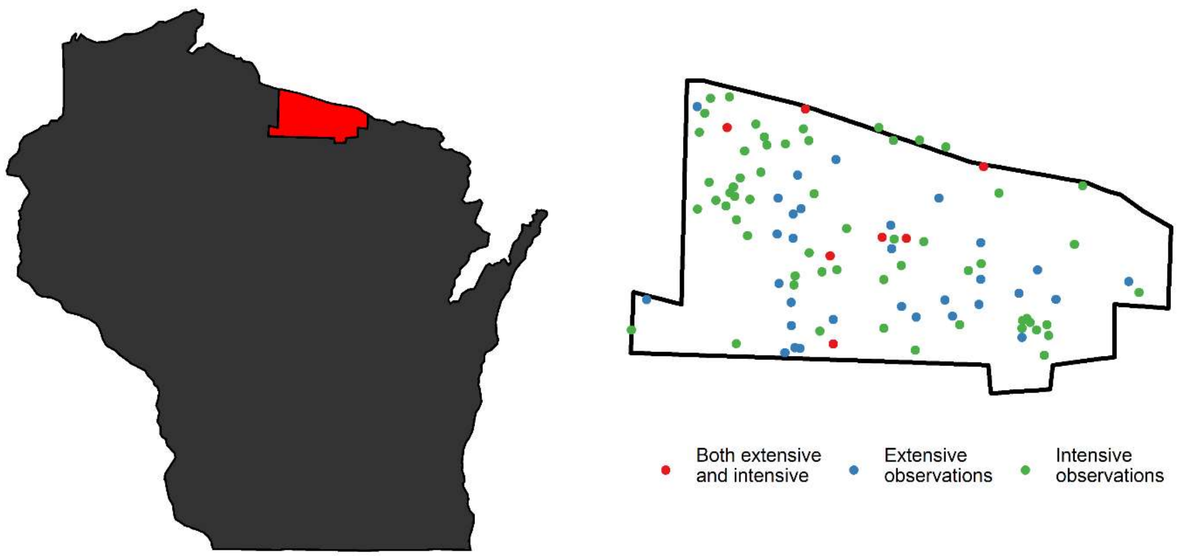

Figure 1: Map of Vilas County, WI showing location of lakes intensively surveyed by WDNR (green), extensively surveyed by our experimental creel survey (blue), and surveyed by both (red).

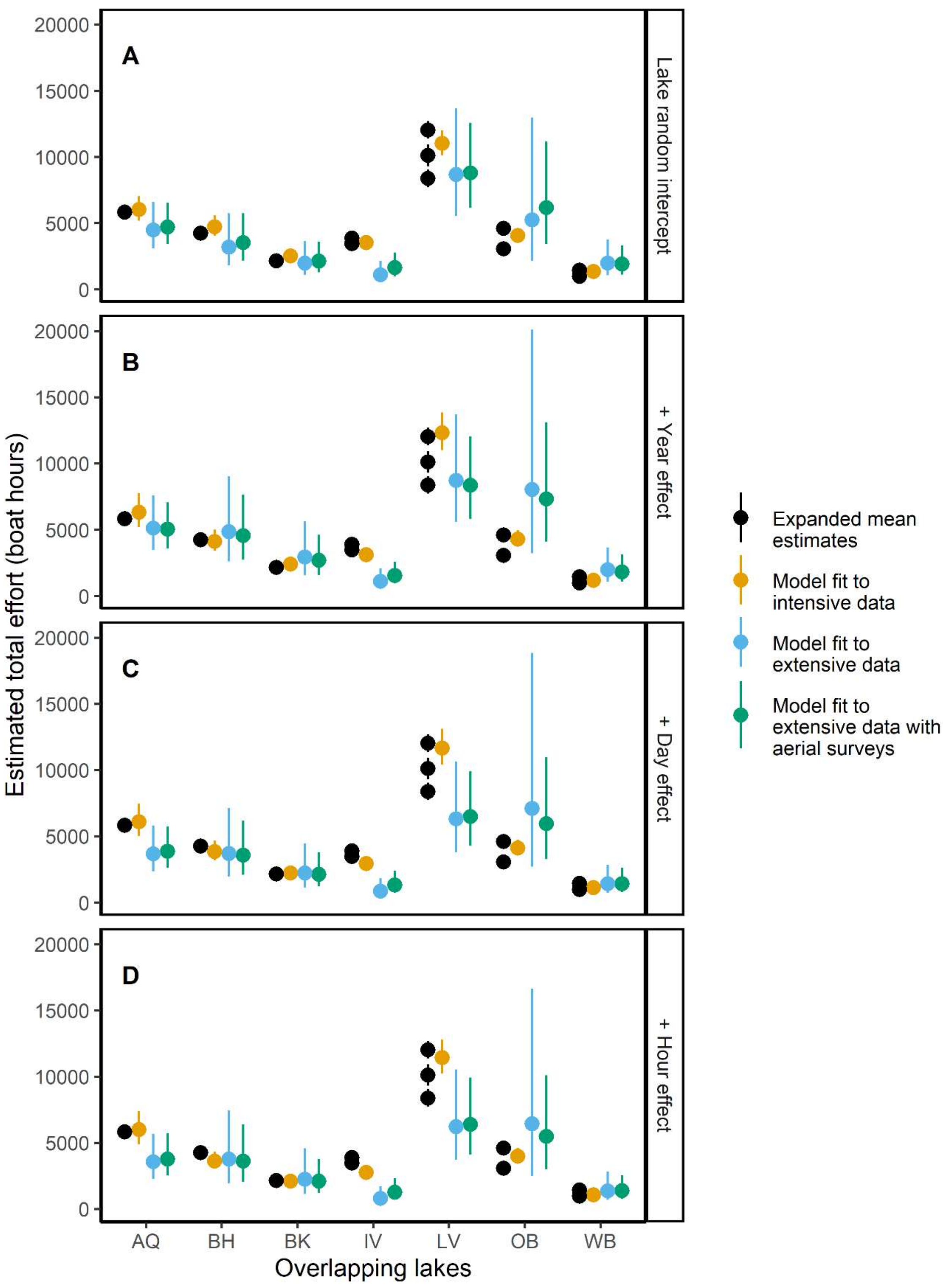

Figure 2: Comparison of total summer fishing effort estimates between mean expansion (black), and area under the curve of GLMM predictions fit to extensive data (colors). Parameters added to each model are indicated by the labels on the right. Points are mean estimates, and bars show 95% prediction intervals. Lakes that were intensively surveyed multiple years by the WDNR have multiple estimates depicted along with their 95% prediction intervals.

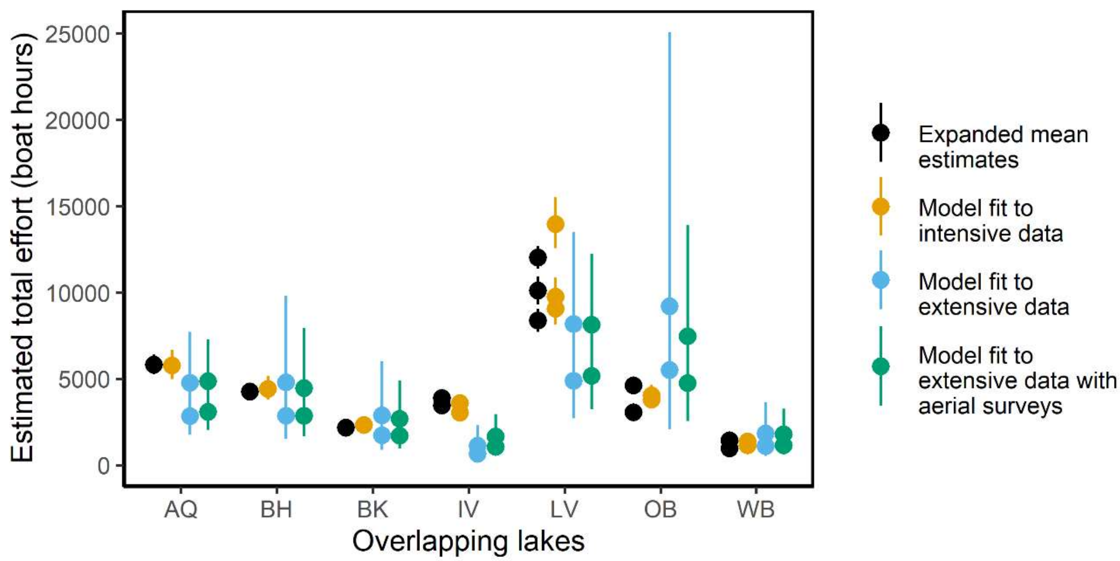

Figure 3: Total summer fishing effort estimates from mean expansion (black) and GLMM predictions incorporating lake, year, day of year, and weekend effects (colors) for every year the lake was surveyed. GLMM predictions from extensive data were always produced for the summers of 2018 and 2019, and mean-expansion estimates and GLMM predictions from intensive data are depicted for the years intensively surveyed. Points are mean estimates for each year observed by the dataset, and bars show 95% prediction intervals.

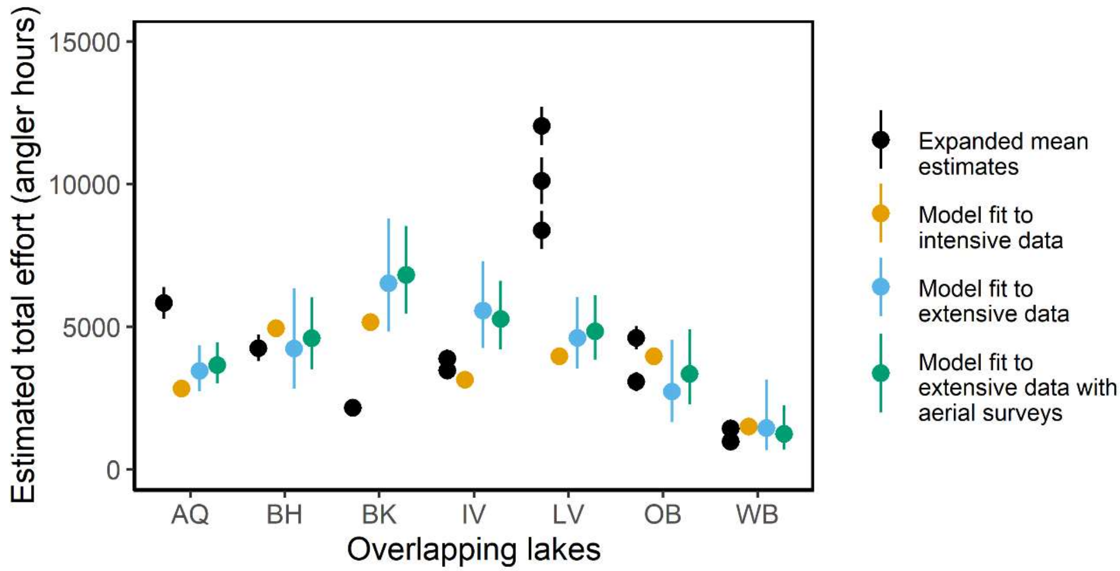

Figure 4: Out-of-sample total summer fishing effort predictions for lakes that were surveyed both extensively and intensively. Lakes that were intensively surveyed multiple years by the WDNR have multiple estimates depicted along with their 95% prediction intervals. Estimates were predicted based on lake characteristics, seasonality, and the grand mean random lake intercept through LOGO cross validation.

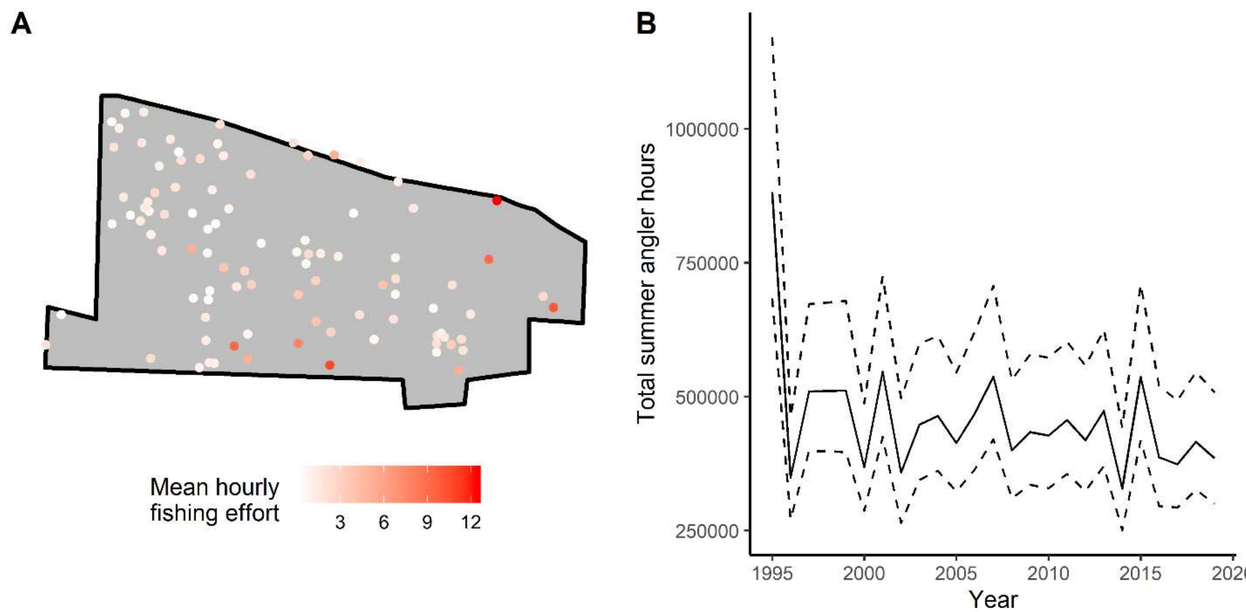

Figure 5: Lake-specific values of the random intercept for each of the 100 lakes surveyed either intensively or extensively in Vilas County, WI (A), and a time series of total annual summer fishing effort across each of these lakes for every year of observations (B).

Supplementary materials

Table S1: Summary of the observations collected intensively and extensively in Vilas county, WI. Standard deviation given in parentheses.

	Intensive dataset	Extensive dataset	Extensive dataset with aerial surveys
Number of lakes surveyed	65	38	44
Number of years surveyed	25	2	2
Mean number of observations per lake	337 (252)	12.4 (4.06)	13.3 (5.8)
Mean number of observations per year	374 (479)	235 (84.9)	294 (168)
Mean number of lakes per year	4.9 (2.6)	21 (7.1)	29.5 (19.1)
Mean number of observations per lake per year	182 (67.5)	11.2 (1.93)	10.5 (4.7)

Table S2: Years surveyed and number of observations for lakes surveyed by both intensively and extensively. Oxbow Lake only received 1 on-site visit (3 instantaneous counts) in 2018 for administrative reasons.

Lake name	Water body identification code (WBIC)	Lake ID	Years surveyed intensively	Number of instantaneous boat counts	Years surveyed extensively	Number of instantaneous boat counts (on-site and/or aerial)
Birch Lake	2311100	BH	1997	170	2018	10
					2019	2
Oxbow Lake	2954800	OB	2008	174	2018	3
			2018	170	2019	3
Allequash Lake	2332400	AQ	2010	176	2018	10
					2019	14
Black Oak Lake	1630100	BK	2011	168	2018	12
					2019	3
Irving Lake	2340900	IV	2001	169	2019	17
			2011	168		
Little Arbor Vitae Lake	1545300	LV	1996	172	2019	14
			2007	170		
			2017	168		
White Birch Lake	2340500	WB	2001	170	2019	14
			2011	168		

A1: Lake selection process

Lakes were selected as part of a larger, multi-objective study of lakes in this region. Initially, fifty lakes were randomly selected from all Vilas County lakes, 35 of which had lake associations, and 15 with no lake associations. All lakes fulfilled the following criteria based on WIDNR data:

- Lakes located completely within Vilas county, not crossing any county or state boundaries
- Have a public launch
- Contain largemouth bass
- Not directly connected to other lakes, so limited connectivity for fish and anglers (No “chained” lakes)
- Surface area less than 500 acres

Distributions of chemical, biological, and morphometric variables across lakes were visually checked using histograms comparing distributions of selected lakes with those of all Vilas county lakes.

Feedback from team members was solicited, and their suggestions were incorporated into an updated lake list based on availability of new data and consideration of logistical constraints. All prior filtering criteria were retained, the maximum lake size was increased to 618 acres (250 hectares) to accommodate the largest sized lakes we could effectively electrofish.

The list of selected lakes was sent to colleagues affliated with the WIDNR for consideration of a Scientific Collectors Permit. They provided feedback on this list indicating areas for revision. Potentially problematic lakes were labeled for the following reasons:

- All smallmouth rather than largemouth bass

- No history of largemouth bass catches in creel data
- Extremely low fish populations
- Difficult to access by shock boat
- Connectivity to other lakes allowing movement of fish and anglers
- Negative encounters with residents

Because of this feedback, 19 lakes were removed from the previous selection. These lakes were replaced with lakes suggested by WIDNR colleagues. The final lake list therefore contained 22 lakes that were randomly selected and 19 lakes suggested as replacements because of their largemouth bass populations,

Note: Low density largemouth bass lakes were retained to achieve a continuum of bass densities and to retain representative low-bass lakes.

A2. Model validation

We needed to establish that, given the same intensive dataset, a generalized linear model-based estimate of total fishing effort is functionally equivalent to an expanded mean estimate. We therefore first compared the total summer fishing effort estimates derived from mean expansion to estimates of total summer fishing effort predicted by negative binomial generalized linear models (GLMs) fit separately to each lake year.

An equivalent model-based estimation approach to mean expansion was developed by fitting a negative binomial generalized linear model (GLM) separately to each lake year of intensive count data with the following parameterization:

$$Count \sim June + July + August + Weekend + June * Weekend + July * Weekend + August * Weekend$$

Instantaneous fishing effort was predicted by dummy variables categorizing the day as belonging to month and weekday strata, as in the mean expansion protocol. By including interaction effects between month and weekend, different weekend effects were estimated for each month. To estimate total summer fishing effort, counts of fishing effort were then predicted for each daylight hour between May 1 and August 31. By calculating the area under the curves of the predictions using the trapezoidal rule, we could then estimate total boat hours for the summer on a particular lake and year as well as upper and lower 95% prediction intervals. Because both sets of estimates were based on the same data and predictors, and because effort on all lakes was estimated separately, total summer fishing effort should be comparable as estimated by both methods.

A more efficient modeling approach may instead fit a quadratic effect of day of year and hour of day to estimate seasonal and daily changes in fishing effort. It would use fewer degrees of freedom than monthly dummy variables and would therefore be a more effective approach to modeling fishing effort using extensive data. Therefore, we additionally fit a negative binomial GLM to each lake year of intensive data with the following form:

$$Count \sim Day\ of\ year + Day\ of\ year^2 + Weekend + Hour\ of\ day + Hour\ of\ day^2$$

Each of these model-based estimates were compared to expanded mean estimates by calculating an index of relative accuracy (I_{RA}) and an index of relative precision (I_{RP}).

When fit to intensively sampled observations, model-based approaches produced very similar results to the standard approach of mean expansion. Negative binomial GLMs were fit to each lake-year of intensive observations, and the area under the curve of the predictions successfully matched the stratified mean total estimates for summer fishing effort (Figure S2). As expected, the model parameterization that more closely matched the stratification of the mean expansion protocol generated nearly identical estimates (Figure S2A). When the monthly dummy variables were replaced by a quadratic effects of day of year and hour of day, some minor deviations from the mean expansion estimates were evident (Figure S2B). For the seven lakes that were surveyed both intensively and extensively, all estimates of total summer fishing effort were effectively the same, with some differences in the width of their confidence intervals (Figure S3).

Model-based estimates of total fishing effort that included month effects of a dummy variable produced estimates that were equally as accurate and precise relative to the expanded mean estimates (Table S3). When models instead included a quadratic effect of day of year, estimates of total effort tended to be lower but more precise than those produced by mean expansion (Table S4).

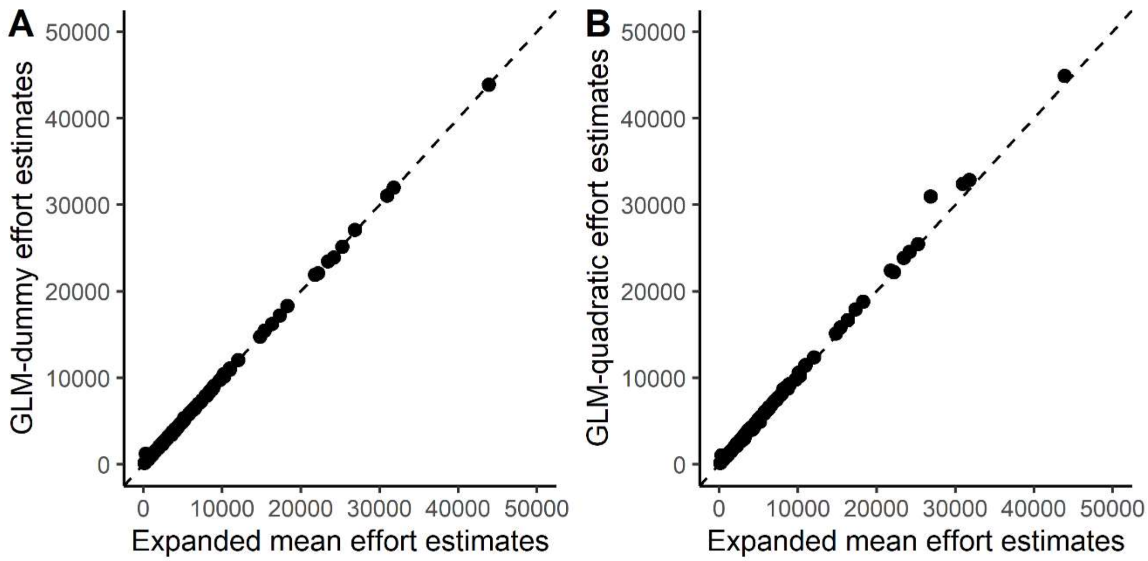

Figure S1: Comparison of model-based estimates of total fishing effort with WIDNR stratified mean estimates for all lake years.

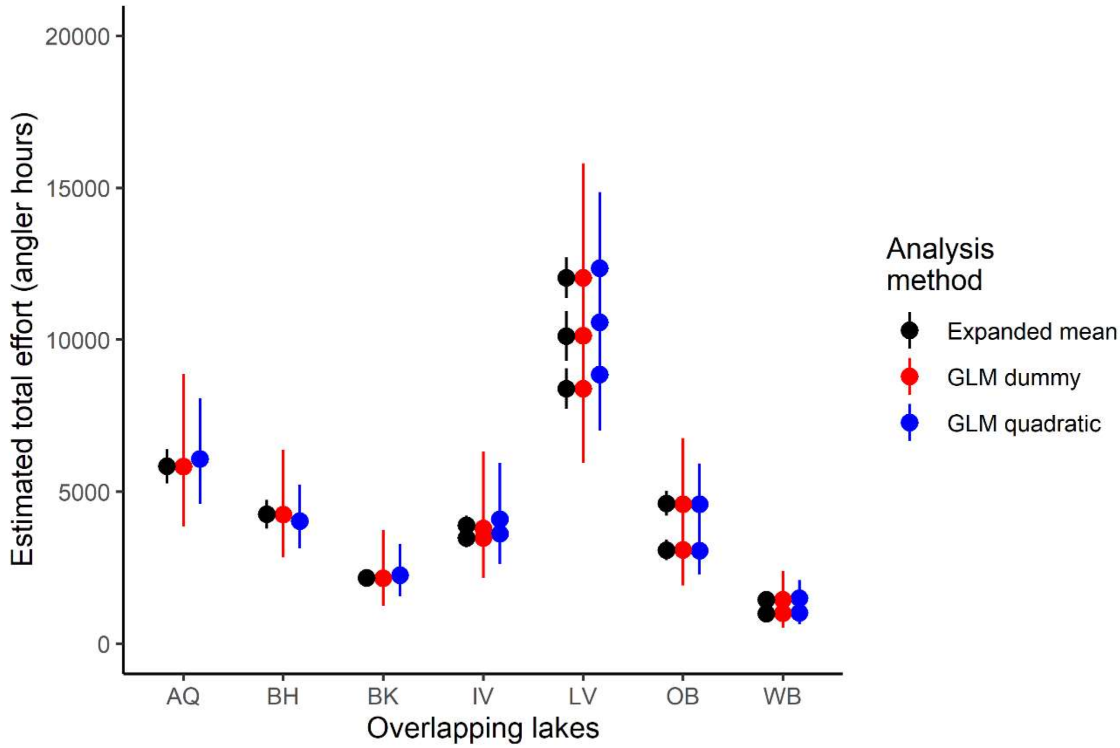

Figure S2: Comparison of total estimated fishing effort and 95% confidence intervals for expanded mean estimates and for two functional forms of a GLM. Lakes that were surveyed multiple years by the WIDNR have multiple estimates depicted along with their 95% confidence intervals. Confidence intervals are wider for GLM estimates, even though they used the same data as the expanded mean estimates. Including the quadratic effect somewhat reduces the width of the GLM confidence intervals.

Table S3: Indices of relative accuracy (I_{RA}) and precision (I_{RP}) for GLM predictions of total summer fishing boat hours day effects relative to expanded mean estimates.

Lake year	Expanded mean total estimate (SD)	GLM prediction (SE)	I_{RA} of GLM prediction	I_{RP} of GLM prediction
AQ 2010	5830.1 (284.5)	5828.8 (1551.5)	-0.02	-81.6667
BH 1997	4253.2 (239.8)	4247.6 (1089.5)	-0.13	-78.0189
BK 2011	2163.5 (111.1)	2142.5 (816.5)	-0.98	-86.5234
IV 2001	3886.6 (163.7)	3783.7 (1294.4)	-2.72	-87.6843
IV 2011	3470.4 (160.2)	3474.3 (1094.4)	0.11	-85.3432
LV 1996	12035.2 (340.0)	12032.5 (1926.1)	-0.02	-82.3508
LV 2007	8383.6 (341.3)	8389.5 (1777.8)	0.07	-80.7875
LV 2017	10118.0 (417.6)	10129.9 (2200.6)	0.12	-80.9993
OB 2008	4610.6 (208.5)	4585.8 (1107.6)	-0.54	-81.2776
OB 2018	3073.9 (169.7)	3077.5 (965.0)	0.12	-82.3964
WB 2001	1439.2 (76.5)	1439.8 (476.0)	0.05	-83.922
WB 2011	981.6 (80.4)	996.6 (513.0)	1.50	-84.0815

Table S4: Indices of relative accuracy (I_{RA}) and precision (I_{RP}) for GLM predictions of total summer fishing boat hours with quadratic day and hour of day effects relative to expanded mean estimates.

Lake year	Expanded mean total estimate (SD)	GLM quadratic prediction (SE)	I_{RA} of GLM quadratic prediction	I_{RP} of GLM quadratic prediction
AQ 2010	5830.1 (284.5)	6078.2 (1013.9)	4.08	-70.75
BH 1997	4253.2 (239.8)	4034.2 (601.2)	-5.43	-62.16
BK 2011	2163.5 (111.1)	2250.7 (523.3)	3.88	-77.91
IV 2001	3886.6 (163.7)	4088.6 (946.6)	4.94	-81.80
IV 2011	3470.4 (160.2)	3611.3 (708.7)	3.90	-76.47
LV 1996	12035.2 (340.0)	12351.6 (1275.7)	2.56	-72.65
LV 2007	8383.6 (341.3)	8852.3 (1208.4)	5.29	-70.18
LV 2017	10118.0 (417.6)	10568.6 (1439.5)	4.26	-69.70
OB 2008	4610.6 (208.5)	4591.0 (679.7)	-0.43	-69.46
OB 2018	3073.9 (169.7)	3054.0 (546.7)	-0.65	-69.16
WB 2001	1439.2 (76.5)	1489.4 (304.0)	3.37	-73.96
WB 2011	981.6 (80.4)	1006.9 (316.7)	2.51	-73.94
		Mean (SD)	**2.36 (3.09)**	**-72.35 (5.04)**
		Mean (SD)	**0.02 (0.004)**	**-82.9 (2.71)**

Table S5: Goodness-of-fit diagnostics for models fit to intensive fishing effort count data as additional parameters are added. Values of ΔAIC for alternate specifications for time of day are the difference between that model's AIC and that of model 3. The best-fit model is in bold.

	Predictors	AIC	ΔAIC	BIC	logLik	Deviance	df	Marginal r^2	Conditional r^2
1	$(1\|Lake)$	90206		90230	-45100	90200	21873		0.36
2	$+ (1\|Year)$	89883	-323	89914	-44937	89875	21872		0.39
3	$+ Day\ of\ year$ $+ Day\ of\ year^2$ $+ Weekend$	88766	-1440	88822	-44376	88752	21869	0.023	0.43
Alternate specifications for time of day:									
4	$+ Hours\ to\ /from\ dark$ $+ Hours\ to\ /from\ dark^2$ $+ Morning$	88393	-1813	88473	-44186	88373	21866	0.032	0.45
5	$+ Hours\ to\ /from\ dark$ $+ Morning$	88399	-1807	88471	-44190	88381	21867	0.032	0.45
6	$\boldsymbol{+ Hour\ of\ day}$ $\boldsymbol{+ Hour\ of\ day^2}$	**87948**	**-2258**	**88020**	**-43965**	**87930**	**21867**	**0.044**	**0.46**

Table S6: Goodness-of-fit diagnostics for models fit to extensive fishing effort count data. Values of ΔAIC for alternate specifications for time of day are the difference between that model's AIC and that of model 3. The best-fit model is in bold.

Model	Predictors	AIC	ΔAIC	BIC	logLik	deviance	df	Marginal r^2	Conditional r^2
1	$(1\|Lake)$	1360.1		1372.5	-677.0	1354.1	467		0.38
2	$+\ year$ 2018	1350.2	-9.9	1366.8	-671.1	1342.2	466	0.035	0.46
3	$+\ Day\ of\ year$ $+\ Day\ of\ year^2$ $+\ Weekend$	1346.6	-13.5	1375.7	-666.3	1332.6	463	0.047	0.50
Alternate specifications for time of day:									
4	$+\ Hours\ to\ /from\ dark$ $+\ Hours\ to\ /from\ dark^2$ $+\ Morning$	1345.5	-14.6	1387.0	-662.8	1325.5	460	0.055	0.50
5	$+\ Hours\ to\ /from\ dark$ $+\ Morning$	1344.3	-15.8	1381.7	-663.2	1326.3	461	0.055	0.50
6	$\boldsymbol{+\ Hour\ of\ day}$ $\boldsymbol{+\ Hour\ of\ day^2}$	**1338.9**	**-21.2**	**1376.3**	**-660.4**	**1320.9**	**461**	**0.065**	**0.52**

Table S7: Goodness-of-fit diagnostics for models fit to extensive creel and aerial survey effort count data. Values for ΔAIC for alternate specifications of time of day are the difference between that model's AIC and that of model 3. The best-fit model is in bold.

Model	Predictors	AIC	ΔAIC	BIC	logLik	deviance	df	Marginal r^2	Conditional r^2
1	$(1\|Lake)$	1725.8		1739.0	-859.9	1719.8	588		0.39
2	$+ year$ 2018	1713.3	-12.5	1730.9	-852.7	1705.3	587	0.021	0.43
3	$+ Day\ of\ year$ $+ Day\ of\ year^2$ $+ Weekend$	1708.5	-17.3	1739.2	-847.3	1694.5	584	0.031	0.45
Alternate specifications for time of day:									
4	$+ Hours\ to\ /from\ dark$ $+ Hours\ to\ /from\ dark^2$ $+ Morning$	1708.9	-16.9	1752.7	-844.4	1688.9	581	0.038	0.45
5	$+ Hours\ to\ /from\ dark$ $+ Morning$	1707.0	-18.8	1746.5	-844.5	1689.0	582	0.037	0.45
6	$+ \boldsymbol{Hour\ of\ day}$ $+ \boldsymbol{Hour\ of\ day^2}$	**1700.0**	**-25.8**	**1739.4**	**-841**	**1682.0**	**582**	**0.044**	**0.46**

Table S8: Parameter estimates for the best-fit model to intensive fishing effort count data.

Model	Random effects	Variance of random intercept (SD)	Fixed effects	Fixed effect coefficients (SE)	Z value	P value
$(1\|Lake) + (1\|year) + Day\ of\ year + Day\ of\ year^2 + Weekend + Hour\ of\ day + Hour\ of\ day^2$	$Lake$	1.06 (1.03)	$Intercept$	0.43 (0.14)	3.11	0.0019*
	$Year$	0.065 (0.25)	$Day\ of\ year$	1.25 (0.086)	14.44	<0.0001*
			$Day\ of\ year^2$	-1.25 (0.086)	-14.48	<0.0001*
			$Hour\ of\ day$	1.42 (0.05)	26.75	<0.0001*
			$Hour\ of\ day^2$	-1.31 (0.052)	-24.94	<0.0001*
			$Weekend\ or\ holiday$	0.48 (0.015)	32.63	<0.0001*

Table S9: Parameter estimates for the best-fit model to extensive fishing effort count data.

Model	Random effects	Variance of random intercept (SD)	Fixed effects	Fixed effect coefficients (SE)	Z value	P value
$(1\|Lake)$ $+ year\ 2018$ $+ Day\ of\ year$ $+ Day\ of\ year^2$ $+ Weekend$ $+ Hour\ of\ day$ $+ Hour\ of\ day^2$	$Lake$	1.271 (1.127)	$Intercept$	-0.41 (0.21)	-1.99	0.047*
			$Day\ of\ year$	1.24 (0.78)	1.58	0.11
			$Day\ of\ year^2$	-1.32 (0.77)	-1.71	0.087
			$Hour\ of\ day$	0.55 (0.37)	1.50	0.13
			$Hour\ of\ day^2$	-0.37 (0.36)	-1.04	0.30
			$Year\ 2018$	-0.36 (0.10)	-3.74	0.0002*
			$Weekend\ or\ holiday$	0.27 (0.11)	2.52	0.012*

Table S10: Parameter estimates for best-fit model to extensive creel and aerial survey fishing effort count data.

Model	Random effects	Variance of random intercept (SD)	Fixed effects	Fixed effect coefficients (SE)	Z value	P value
$(1\|Lake) + year\ 2018 + Day\ of\ year + Day\ of\ year^2 + Weekend + Hour\ of\ day + Hour\ of\ day^2$	$Lake$	0.98 (0.98)	$Intercept$	-0.31 (0.17)	-1.80	0.072
			$Day\ of\ year$	0.74 (0.69)	1.07	0.28
			$Day\ of\ year^2$	-0.86 (0.68)	-1.26	0.21
			$Hour\ of\ day$	0.48 (0.31)	1.52	0.13
			$Hour\ of\ day^2$	-0.32 (0.31)	-1.06	0.29
			$Year\ 2018$	-0.26 (0.07)	-3.91	<0.0001*
			$Weekend\ or\ holiday$	0.20 (0.088)	2.217	0.027*

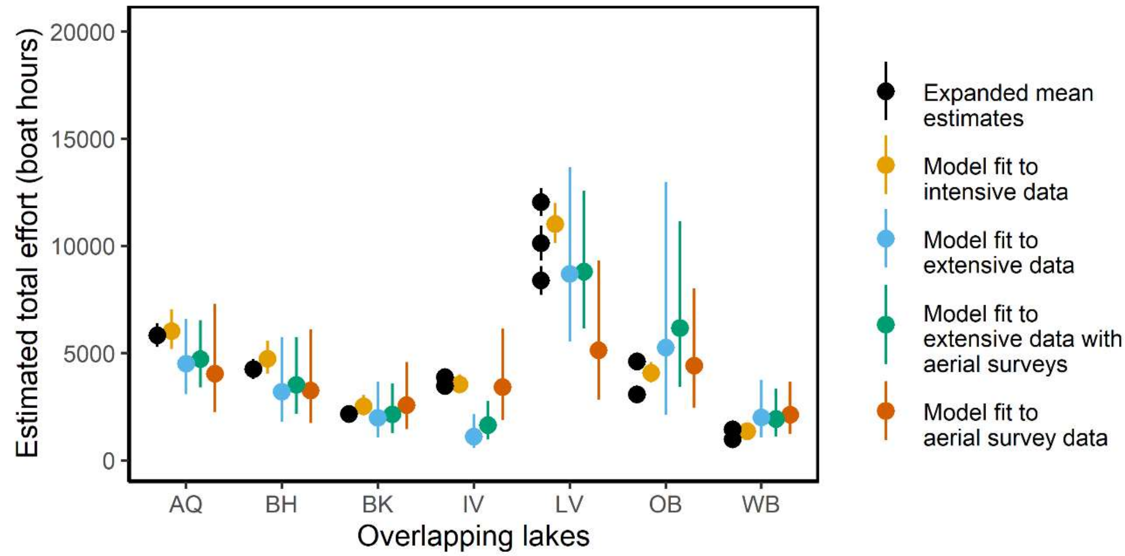

Figure S3: Comparison of total summer fishing effort estimates between traditional mean expansion of intensive data (black), and GLMM-based estimates (colors). Lakes that were intensively surveyed multiple years by the WDNR have multiple estimates depicted along with their 95% confidence intervals. Each dataset was fit to a simple GLMM containing only a lake-specific random intercept. Aerial survey data alone produced similar fishing effort estimates as larger datasets.

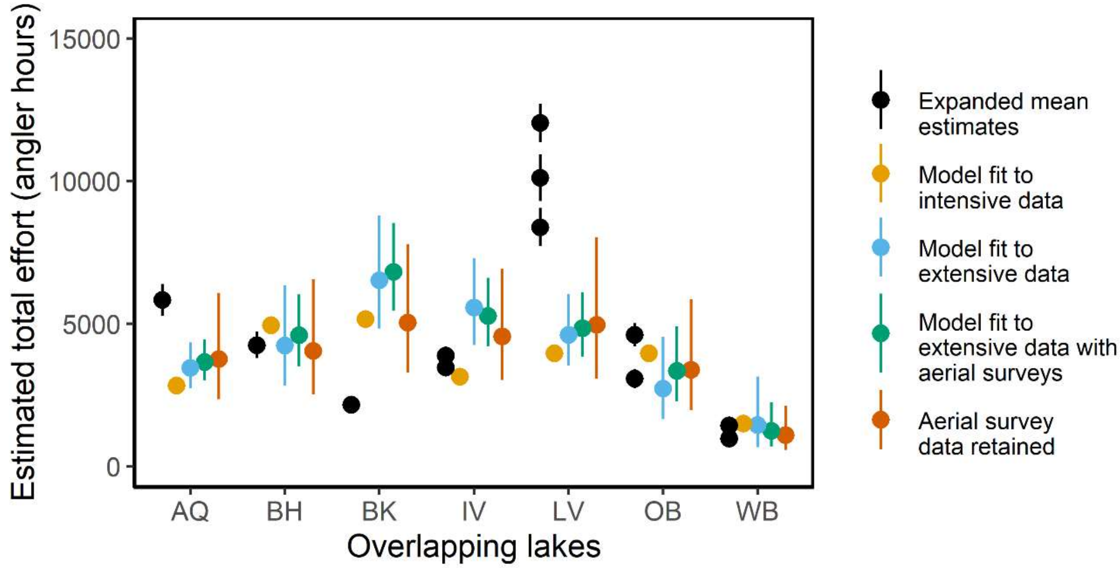

Figure S4: Total fishing effort estimates for out-of-sample lakes obtained by LOGO cross validation. Fishing effort estimates labeled "Aerial survey data retained" were obtained by leaving out on-site observations from the model fit but retaining three aerial survey observations.

Table S11: Random intercept values for all lake year combinations from a GLMM fit to the combined intensive and extensive datasets.

Lake name	WBIC	Random intercept: Mean hourly fishing effort	Surface area (hectares)	Years surveyed
Lac Vieux Desert	1631900	12.36316	1626.885	2013, 2006
Little Saint Germain Lake	1596300	10.61531	393.66	1997, 2015
Kentuck Lake	716800	10.05069	405.405	1998
Big Arbor Vitae Lake	1545600	9.223575	433.35	2008, 1998, 2005, 2011, 2014, 2017
Twin Lakes	1623800	8.957309	1162.755	2007, 1996, 2017
Big Saint Germain Lake	1591100	7.693141	656.91	2011
Catfish Lake	1603700	5.212021	396.09	2000, 2013
Upper Gresham Lake	2330800	5.078966	146.61	2019, 2015
Big Lake	2963800	5.056214	315.9	2008
Little Arbor Vitae Lake	1545300	5.03757	194.4	2019, 2015, 1996, 2007, 2017
Upper Buckatabon Lake	1621800	4.004153	199.665	2010
Lost Lake	1593400	3.95955	218.295	2019, 2015
Trout Lake	2331600	3.804271	1564.92	2001, 2004, 2007, 2010, 2013, 2016, 2019
Plum Lake	1592400	3.498538	428.085	1995, 2003, 2006, 2009, 2012, 2015, 2018
Eagle Lake	1600200	3.280705	232.875	2000, 2013
Star Lake	1593100	3.073703	493.695	1997, 2005
Big Muskellunge Lake	1835300	3.013383	363.285	1996
Palmer Lake	2962900	2.740117	260.82	2019, 2009
Found Lake	1593800	2.629966	136.08	2018, 2019, 2015
Allequash Lake	2332400	2.565629	164.43	2018, 2019, 2015, 2010
Spectacle Lake	717400	2.544978	67.23	2019, 2015
Clear Lake	2329000	2.415536	208.575	1999, 2004
Ballard Lake	2340700	2.318722	203.715	2019, 2015, 2001, 2011

Crab Lake	2953500	2.209385	368.145	2000, 2002
Gunlock Lake	1539700	2.20889	106.92	2002
Lower Buckatabon Lake	1621000	2.201147	153.09	2010
Island Lake	2334400	2.161185	350.325	1999, 2004
Scattering Rice Lake	1600300	2.145238	106.515	2000, 2013
Yellow Birch Lake	1599600	2.097776	77.76	2000, 2013
Pioneer Lake	1623400	2.026316	173.745	2019, 2015
Tenderfoot Lake	2962400	1.952217	183.465	2009
South Turtle Lake	2310200	1.905668	188.73	2010
Wildcat Lake	2336800	1.901813	118.665	2018, 2019, 2015
Oxbow Lake	2954800	1.863795	211.815	2018, 2019, 2015, 2008
Pickerel Lake	1619700	1.854283	109.35	2019, 2015
Boot Lake	1619100	1.828323	115.83	2019, 2015
Van Vliet Lake	2956800	1.821236	93.15	2015, 2012
Voyageur Lake	1603400	1.813109	57.915	2013
Amik Lake	2268600	1.799903	57.105	1998, 2005, 2018
Duck Lake	1599900	1.724904	42.93	2000, 2013
Muskellunge Lake	1595600	1.70102	116.235	2018, 2019, 2015
Anvil Lake	968800	1.649552	152.685	2019, 2015
Big Lake	2334700	1.631623	334.935	1995
Birch Lake	2311100	1.620301	204.93	2018, 2019, 2015, 1997
Little Spider Lake	1540400	1.551719	90.315	2018, 2019, 2015
Little John Lake	2332300	1.477845	61.155	2019
Wild Rice Lake	2329800	1.472612	155.52	1999, 2004
Rest Lake	2327500	1.471456	265.275	1999, 2004
Stone Lake	2328800	1.435453	55.89	2004, 1999
Big Portage Lake	1629500	1.407541	237.33	2006
Deerskin Lake	1601300	1.401943	121.905	2019, 2015
Manitowish Lake	2329400	1.386719	200.88	2016, 1999, 2004
Otter Lake	1600100	1.365467	70.47	2000, 2013
Harris Lake	2958500	1.312162	216.27	1997, 2019
Towanda Lake	1022900	1.2963	56.295	2018, 2019, 2015
Mamie Lake	2964100	1.288534	136.485	2008
Landing Lake	1630700	1.285178	82.215	2019
Presque Isle Lake	2956500	1.276282	471.825	2012

Fawn Lake	2328900	1.258813	28.35	2004
Irving Lake	2340900	1.235148	169.695	2019, 2015, 2001, 2011
Lynx Lake	2954500	1.174211	124.335	1998
Brandy Lake	1541300	1.144568	45.765	2019, 2015
Little Crooked Lake	2335500	1.046359	62.37	2018, 2019, 2015
Arrowhead Lake	1541500	1.011037	38.88	2018, 2019, 2015
Rainbow Lake	2310800	1.000412	59.94	2019, 2015
Big Kitten Lake	2336700	0.972799	20.25	2019
Black Oak Lake	1630100	0.969767	228.42	2018, 2019, 2015, 2011
Papoose Lake	2328700	0.969182	170.91	1997, 2012
North Turtle Lake	2310400	0.939383	145.395	2010
Johnson Lake	1541100	0.898655	34.425	2018, 2019, 2015
Lake Laura	995200	0.897062	254.34	1998
Alder Lake	2329600	0.853639	106.92	1999, 2004
Lynx Lake	1600000	0.821144	12.555	2000, 2013
Boulder Lake	2338300	0.794834	208.98	1999, 1995
Spider Lake	2329300	0.781624	112.59	1999, 2004
Silver Lake	1599800	0.681033	23.085	2019, 2015
Stormy Lake	1020300	0.671659	211.815	2019, 2015
Erickson Lake	983600	0.642838	44.55	2019, 2015
Partridge Lake	2341500	0.587026	95.175	2019, 2015
Annabelle Lake	2953800	0.577419	78.57	1996, 2019
Hunter Lake	991700	0.52965	70.875	2018, 2019, 2015
Snipe Lake	1018500	0.529036	87.48	1995, 2000, 2003, 2006, 2009, 2012, 2015, 2018
Rock Lake	2311700	0.504688	48.6	2010
White Birch Lake	2340500	0.486677	45.765	2019, 2015, 2001, 2011
Day Lake	1843500	0.483598	44.55	2019, 2015
Lone Tree Lake	1000400	0.437012	52.65	2019, 2015
Street Lake	1884200	0.404973	18.63	2019, 2015
Camp Lake	1839100	0.390038	15.39	2018, 2019, 2015
Lake of the Hills	1620500	0.385468	24.705	2018, 2019, 2015
Little Star Lake	2334300	0.331636	105.3	1999, 2004

Sparkling Lake	1881900	0.302415	63.585	1996, 2006
Dead Pike Lake	2316600	0.279634	125.145	2016, 2005
Nichols Lake	1870400	0.246555	14.985	2019, 2015
Lost Canoe Lake	2339800	0.242777	112.995	2015, 1995
Wabasso Lake	2045000	0.232163	21.06	2018, 2016
Indian Lake	2764400	0.209596	32.4	2019, 2015
Whitney Lake	2338100	0.203369	91.53	2019, 2015
Lake Adelaide	1831700	0.199687	23.085	2019, 2015
Little Rock Lake	1862100	0.08288	15.795	2019, 2015, 2008
Averill Lake	2956700	0.075635	27.54	2012

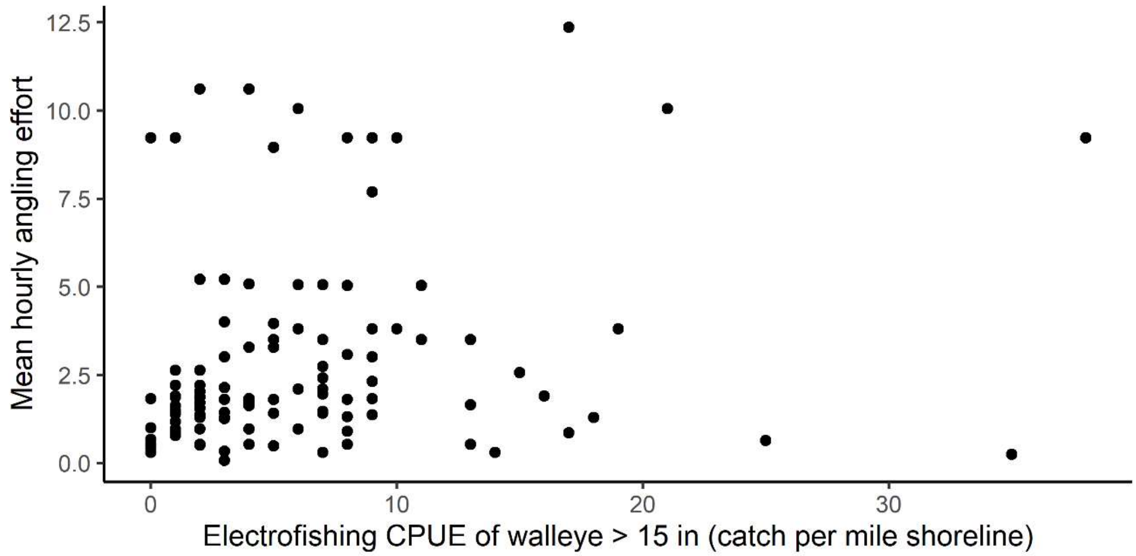

Figure S5: Lake-specific random intercepts estimated for all lakes surveyed versus electrofishing catch per unit effort of adult walleye.

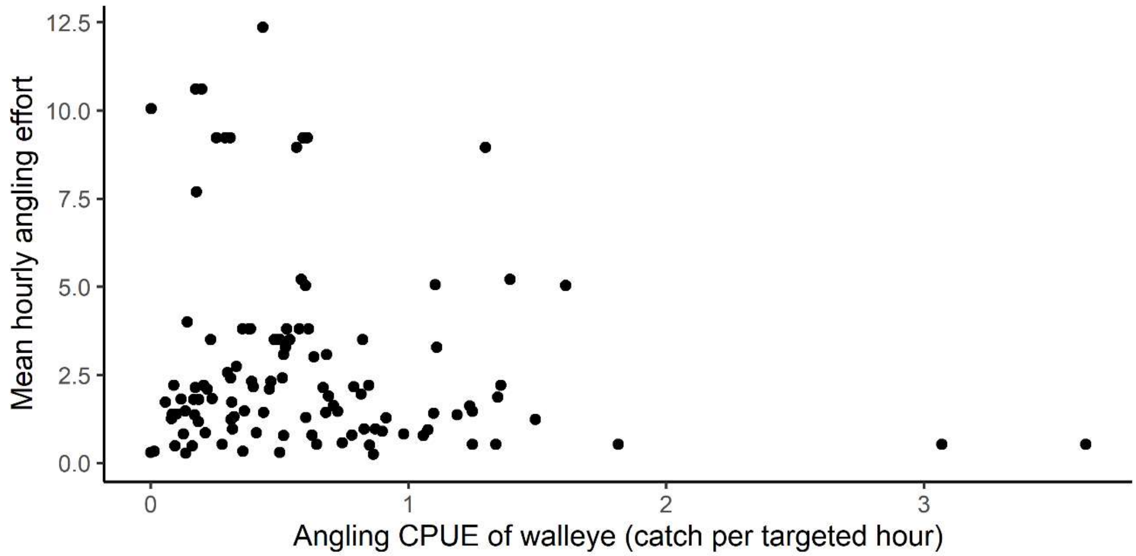

Figure S6: Lake-specific random intercepts estimated for all lakes surveyed versus angling catch per unit effort of walleye.

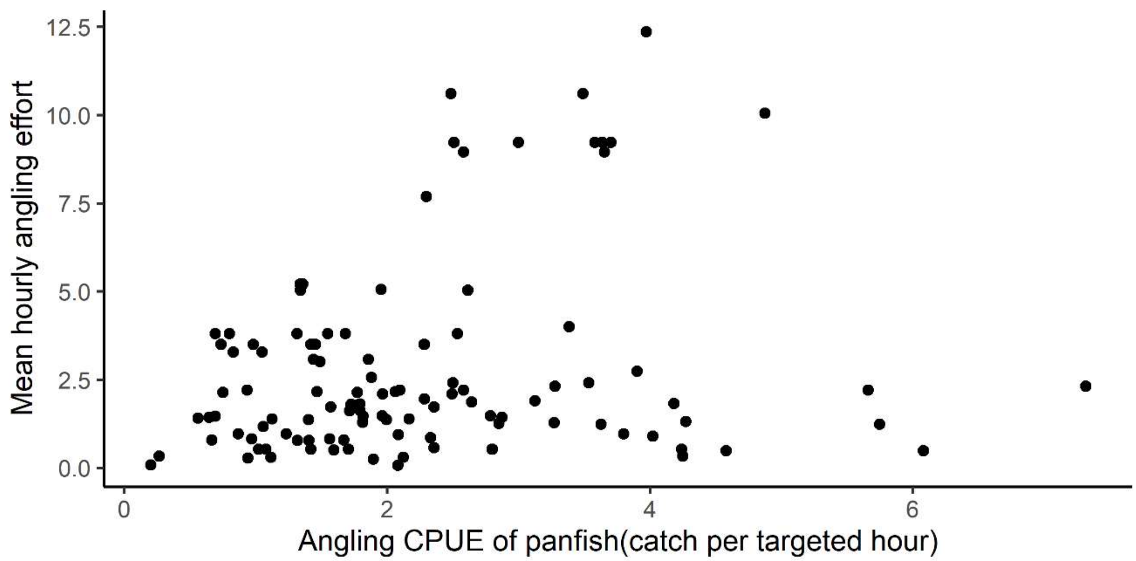

Figure S7: Lake-specific random intercepts estimated for all lakes surveyed versus angling catch per unit effort of panfish, including yellow perch, bluegill, pumpkinseed, and black crappie.

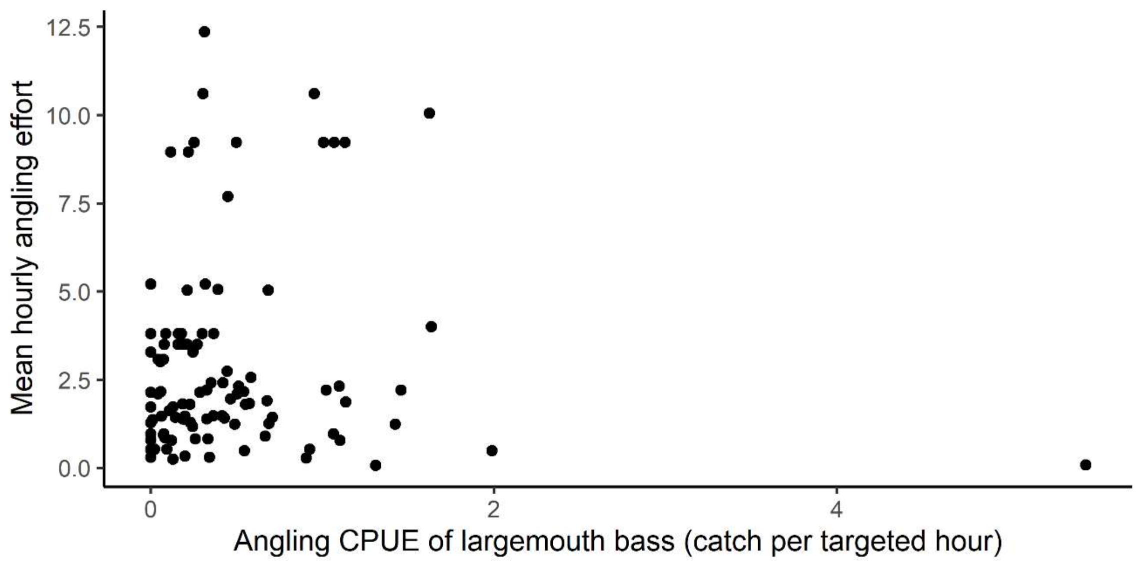

Figure S8: Lake-specific random intercepts estimated for all lakes surveyed versus angling catch per unit effort of largemouth bass.

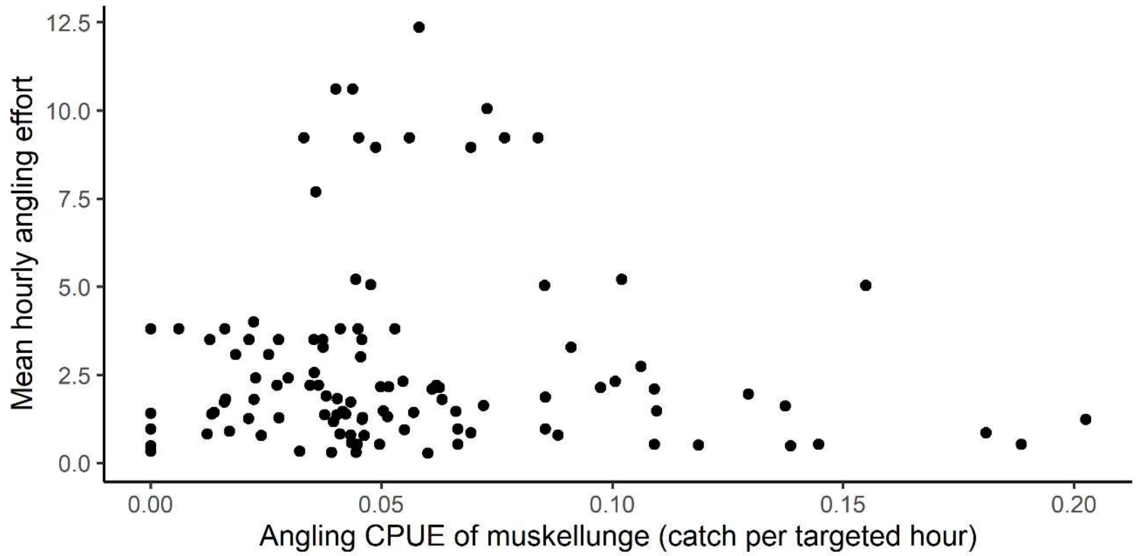

Figure S9: Lake-specific random intercepts estimated for all lakes surveyed versus angling catch per unit effort of muskellunge.