Evaluating the Power to Detect Temporal Trends in Fishery-Independent Surveys: A Case Study Based on Gill Nets Set in the Ohio Waters of Lake Erie for Walleyes

Tyler Wagner*1

Quantitative Fisheries Center, Department of Fisheries and Wildlife, Michigan State University, 153 Giltner Hall, East Lansing, Michigan 48824-1101

CHRISTOPHER S. VANDERGOOT AND JEFF TYSON

Ohio Department of Natural Resources, Division of Wildlife, Sandusky Fisheries Research Station, 305 East Shoreline Drive, Sandusky, Ohio 44870, USA

Abstract.—Fishery-independent (FI) surveys provide critical information used for the sustainable management and conservation of fish populations. Because fisheries management often requires the effects of management actions to be evaluated and detected within a relatively short time frame, it is important that research be directed toward FI survey evaluation, especially with respect to the ability to detect temporal trends. Using annual FI gill-net survey data for Lake Erie walleyes Sander vitreus collected from 1978 to 2006 as a case study, our goals were to (1) highlight the usefulness of hierarchical models for estimating spatial and temporal sources of variation in catch per effort (CPE); (2) demonstrate how the resulting variance estimates can be used to examine the statistical power to detect temporal trends in CPE in relation to sample size, duration of sampling, and decisions regarding what data are most appropriate for analysis; and (3) discuss recommendations for evaluating FI surveys and analyzing the resulting data to support fisheries management. This case study illustrated that the statistical power to detect temporal trends was low over relatively short sampling periods (e.g., 5-10 years) unless the annual decline in CPE reached 10-20%. For example, if 50 sites were sampled each year, a 10% annual decline in CPE would not be detected with more than 0.80 power until 15 years of sampling, and a 5\% annual decline would not be detected with more than 0.8 power for approximately 22 years. Because the evaluation of FI surveys is essential for ensuring that trends in fish populations can be detected over management-relevant time periods, we suggest using a metaanalysis-type approach across systems to quantify sources of spatial and temporal variation. This approach can be used to evaluate and identify sampling designs that increase the ability of managers to make inferences about trends in fish stocks.

Fishery-independent (FI) surveys help fulfill a variety of fisheries management objectives, including assessing stock status and trends, setting annual harvest levels, and evaluating the effects of natural and anthropogenic stressors on the spatiotemporal dynamics of populations (Quinn and Deriso 1999; Sitar et al. 1999; Allen et al. 2007). Although data derived from FI surveys are used to fulfill many objectives, one of the most common uses of these data is to infer changes in indices of relative abundance over time (Pennington and Strømme 1998; Stobutzki et al. 2006; Corradin et al. 2008). Knowledge of rates of population decline or increase is critical to maintain sustainable management of exploited species and to assess recovery efforts for

Much research has been conducted on various aspects of sampling methodology and on the efficiency of different gear types used during FI surveys. For example, the design of statistically valid FI surveys for monitoring fish populations has been acknowledged as essential to ensure that inferences regarding status and trends can be generalized to a population of interest

Received September 4, 2008; accepted December 11, 2008 Published online May 28, 2009

species of high conservation priority (Zwieten et al. 2002; Maxwell and Jennings 2005). In addition, fisheries management often requires the effects of management actions to be evaluated and detected within a relatively short time frame, such as 5 years (Zwieten et al. 2002). Therefore, quantifying the magnitude of change in the relative abundance of a population that can be detected from FI surveys has important implications for the time scale needed for managers to evaluate management actions, detect changes attributable to sources other than management (e.g., invasive species), and—to avoid undesirable changes in abundance—respond and take appropriate action.

^{*} Corresponding author: txw19@psu.edu

¹ Current address: U. S. Geological Survey, Pennsylvania Cooperative Fish & Wildlife Research Unit, Pennsylvania State University, 402 Forest Resources Building, University Park, Pennsylvania 16802, USA.

(Ault et al. 1999; Hayes et al. 2003; Smith and Tremblay 2003). This has led to a switch from nonprobability-based sampling to an increased use of probability-based sampling designs (Olsen et al. 1999). In addition, important research has compared the efficiency of different gear types and evaluates factors affecting the efficiency and representativeness of different gear types used during FI surveys (McInerny and Cross 2006; Niles and Hartman 2007; Yule et al. 2008). However, given the fiscal resources devoted to performing FI surveys, relatively little research has been focused on detecting temporal trends (especially for freshwater systems; see Maxwell and Jennings 2005 and Blanchard et al. 2008 for marine examples); in particular, little has been done with quantifying sources of variability that influence the power to detect changes over time in indices such as catch per effort (CPE). The inherent variability of fish populations is what leads to questions related to how to best sample (e.g., allocate resources) to maximize trend detection capabilities. For example, common questions posed by fishery managers include "given the current number of sites we sample each year, what magnitude of decline in CPE can we detect within 5 or 10 years?", or "how many sites do we need to sample each year to detect a prespecified decline in CPE within 10 years?" The answers to these and other monitoring questions are largely influenced not only by the total variability in fish CPE, but also by how it is partitioned among various spatial and temporal components (see Methods).

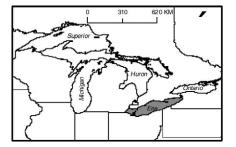
As in many freshwater systems, FI surveys are used in the Laurentian Great Lakes by state, provincial, tribal, and federal fish and wildlife agencies to assess the status and trends of a variety of fish species (Sitar et al. 1999; Stapanian et al. 2007; Pothoven and Madenjian 2008; Ohio Department of Natural Resources [ODNR] 2008). The data from these surveys represent an important source of information for managing both exploited and unexploited fish species in a socially and ecologically complex ecosystem. Specifically, FI surveys often are used to provide time series of relative abundance, size and age composition of the population, and indices of fish condition and sex ratios. For this case study, we focused on FI survey data for CPE of walleye Sander vitreus in the Ohio waters of Lake Erie. Walleyes, an ecologically important species native to the Great Lakes, support important recreational and commercial fisheries in the Great Lakes, particularly Lake Erie. For example, the walleye fishery in the Ohio waters of Lake Erie attracted over 1 million trips by sport anglers in 2007, most of the private boat effort (63%) targeting walleyes (ODNR 2008).

Because walleyes in Lake Erie are important ecologically and economically, obtaining FI survey data for input into population models to assist in the development of harvest programs and to track trends in abundance over time is a critical component of effective management. For instance, the Walleye Management Plan developed for Lake Erie establishes fishery sustainability and quality objectives for walleye management (Locke et al. 2005). Included in this plan is an exploitation policy designed to help meet specific management objectives, and the effects of this policy are required to be evaluated on a 5-year basis. Simulation modeling is one method currently used to evaluate harvest policies; however, because of the relatively long time scale the simulation models operates over and the high temporal variability in walleye abundance over shorter time scales, both these sources of information (long-term simulations and short-term monitoring) should be considered when determining the success of the harvest policy in meeting objectives.

The goal of this study was to use FI surveys for walleyes sampled from the Ohio waters of Lake Erie as a case study to illustrate the utility of quantifying sources of variability for evaluating capabilities of FI surveys to detect temporal trends under different sampling scenarios and to establish realistic expectations for what can be detected over relatively short but management-relevant time periods (e.g., 5–10 years). Although Lake Erie walleyes are used as a case study, the methods discussed can be applied to different species and systems. The specific objectives of this study were (1) to quantify the spatial and temporal sources of variation in CPE from annual gill-net surveys of walleyes conducted by ODNR in Ohio waters of Lake Erie; (2) using the estimates of spatial and temporal variability, to examine the statistical power for detecting trends with regard to (a) the number of net sets, (b) the magnitude of the trends, (c) the number of years over which FI surveys were conducted (i.e., the sampling duration), (d) the assumptions made about what data should be included in the analysis (specifically, how do results change if years before zebra mussels invasion, which were characterized by lower water clarity compared with that in postinvasion years, are excluded from the analysis); (3) to discuss the results in terms of setting realistic expectations for FI surveys; and to provide recommendations for future analyses to help evaluate and inform the design of FI surveys.

Methods

Annual gill-net surveys.—Experimental gill-net surveys were conducted in the Ohio waters of Lake Erie



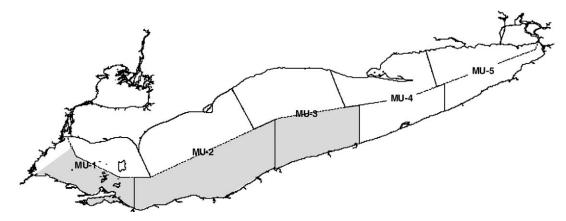


FIGURE 1.—Map of Lake Erie showing fishery management units. The outlined areas represent the five management units (MUs) for walleyes recognized by the Lake Erie Committee; the shaded areas represent the extent of Ohio waters and the MUs used in the analysis.

by ODNR, Division of Wildlife, to ascertain the relative abundance of walleyes. The gill-net survey was initiated in 1978, and although the survey has changed through the years in terms of effort expended (i.e., the number of nets set), the same sampling gear has been utilized throughout. Gill nets used by the ODNR are experimental mesh multifilament gill nets; each net consists of a gang of 13 randomly ordered panels (each panel is 1.8 m tall × 30.5 m wide) with mesh sizes ranging between 51- and 127-mm stretched mesh in 6-mm increments. Annually, the gill-net survey is conducted during the fall (late Septemberearly November) with "canned" gill-net sets (i.e., the nets are suspended from the surface to a depth of 1.8 m with individual floats at each panel). Lake Erie is divided into five management units (MUs); gill-net surveys used for this study were conducted in Ohio waters of MUs 1-3 (Figure 1). Total counts of age-2 and older walleyes were obtained from gill-net catches as an index of relative abundance. Younger ages of walleyes were not included in this analysis because these age-classes of walleyes are not fully recruited to the ODNR gill nets. Thus, CPE of walleyes in ODNR gill nets was defined as the total number of age-2 and older walleyes collected in each individual net set.

Variance components.—It is not just the total variance that influences the statistical power to detect temporal trends; rather, how the total variance is partitioned among spatial and temporal sources has a critical influence on power (Urquhart et al. 1998). Therefore, quantifying multiple sources of spatial and temporal variation has been advocated as an approach to address the issue of variability in ecological data when evaluating temporal trends and monitoring aquatic ecosystems (Urquhart et al. 1998; Larsen et al. 2001; Kincaid et al. 2004). This approach, known as a variance components framework, has been applied successfully to indicators of water quality (Urquhart et al. 1998), fish habitat indicators (Larsen et al. 2004), and more recently to fish mean size at age (Wagner et al. 2007). Under this variance components framework, total variance is partitioned into several components, including (1) site-to-site (spatial) variation; (2) coherent (year-to-year) variation affecting all sites in a similar manner; (3) ephemeral temporal variation (i.e., site × year interaction) corresponding to independent yearly variation at each site; (4) trend variation, in which each

site is allowed to have its own trend; and (5) residual variation (VanLeeuwen et al. 1996; Larsen et al. 2001; Kincaid et al. 2004). More complex error structures may also need to be considered beyond that previously described. For example, the temporal analysis of survey catches can potentially suffer from a lack of independence across annual observations because individual cohorts are captured during multiple years. The progression of a strong cohort over time would probably result in more similar catches during those years. Such may be the case with percids, where variation in abundances over time is strongly influenced by year-class strength. Thus, investigations into temporal autocorrelation in year effects are also warranted. Once variance components have been estimated, they can be used to evaluate the statistical power of different fishery indicators, such as CPE. Here, we apply and demonstrate the utility of this approach for examining sampling questions related to FI surveys.

ODNR biologists had concerns about the potential effects of changes in water clarity over time on catchability with gill nets. For example, fall Secchi measurements in the late 1980s were steady at approximately 0.75 m and then increased substantially in the mid-1990s to near 1.6 m. This increase in water clarity coincided with colonization by dreissenid mussels and led to concerns that the spatial and temporal dynamics of walleyes in the 1970s and 1980s were not representative of current walleye dynamics (ODNR unpublished data). To address this concern, we performed two separate analyses on the Lake Erie walleye FI survey data. First, we analyzed the entire time series extending from 1978 to 2006 and included data sampled from MUs 1-3 (hereafter referred to as the "full analysis"). We performed this analysis to make use of the entire ensemble of data when quantifying spatial and temporal patterns in CPE. In a second analysis, to address the concern regarding a change in gill net catchability, we restricted the analysis to a subset of data ranging from 1996 to 2006 (hereafter referred to as the "restricted analysis"). The ODNR biologists also felt that the most appropriate data to include in the restricted analysis were those currently used in stock assessment models. Thus, the restricted analysis was also restricted to data collected only from MUs 1 and 2. Although using a cutoff date of 1996 and restricting the analysis to MUs 1 and 2 is somewhat arbitrary, this represents one of many possible choices managers can make with respect to which data are most appropriate to incorporate in such an analysis. Thus, we use this as an opportunity to evaluate the sensitivity of the results to assumptions made about what data are most appropriate for characterizing the spatial and temporal dynamics of walleye populations in Ohio waters of Lake Erie.

Statistical model.—For both the full and the restricted analyses separately, we used a hierarchical model to assess the presence of any temporal trend in walleye CPE and to obtain estimates of variance components for use in simulation modeling. Although this model does provide an estimate of average temporal trend, the variance estimates are of primary interest for evaluation procedures outlined below.

The mixed model used for the analyses was

$$Y_{ijk} = \mu + a_i + y(\lambda + t_i) + b_j + c_{ij} + e_{ijk}$$
 (1)

Where Y_{ijk} is the $\log_e(\text{CPE})$ for sample k at site i in year j, and μ and λ are the fixed intercept and slope, respectively. The random effect, a_i , represents site-tosite variability, which is independent and indentically distributed (iid) as $N(0, \sigma_a^2)$; b_i , is a random effect for the jth year (coherent temporal variability) that is iid as $N(0, \sigma_b^2)$; t_i is a random effect for the trend (i.e., slope) for site i that is iid as $N(0, \sigma_t^2)$; c_{ii} is the site \times year interaction (ephemeral temporal variability), which is iid as $N(0, \sigma_c^2)$; and e_{ijk} is the unexplained error (residual error), which is iid as $N(0, \sigma_e^2)$. The year covariate (y) is the jth year minus the mean year used in the analysis. Thus, the parameters estimated by using equation (1) are a fixed slope and intercept and the five variances estimates $(\hat{\sigma}_n)$ associated with the random effects. This standardization of year was performed to provide numerical stability. To investigate whether the sequential relative abundance values were correlated over time, we also modeled the coherent year effect by using a first-order autoregressive model. Year effects were modeled as

$$\tau_{t+1} = \rho \tau_t + \varepsilon_t; \varepsilon_t \sim N(0, \sigma_s^2), \tag{2}$$

where τ is the year effect, ρ is the first-order autoregressive parameter, and ϵ is a random error term. We used a likelihood ratio test to determine whether the more complicated error structure was warranted. Variance components were estimated using restricted maximum likelihood and *P*-values with a likelihood ratio test (Self and Liang 1987; Littell et al. 1996). We considered all analyses significant at P < 0.05. All means are shown \pm 1 standard error (SE).

Simulations.—We investigated the extent to which the following factors affected the ability to detect a temporal trend in walleye CPE: (1) the data set used in the analysis (full versus restricted), (2) changes in the number of sites sampled (i.e., the number of net sets, sampling 10, 25, 50, or 100 fixed sites each year), (3) trend magnitude (λ decreases ranging from 3% to 20% per year), and (4) sample duration (range, 5 to 25

years). In this study we limited the analysis to the detection of linear trends; however, if a monotonic increase or decrease in CPE occurs, then a linear trend will be present (Urguhart and Kincaid 1999). We used a simulation approach to examine the statistical power to detect temporal trends, using the variance components estimated from equation (1), following methods outlined in Wagner et al. (2007). For each simulation, 1,000 data sets were generated containing CPE data for a population of sites. We ran the simulations for two population sizes from which potential sites were sampled. For the full analysis the population of sites was set at 305, corresponding to the total number of sites identified by the ODNR that could be potentially sampled (102 sites in MU 1, 130 sites in MU 2, and 73 sites in MU 3). For the restricted analysis, the population of sites was set at 232, corresponding to the total number of sites that could potentially be sampled in MU 1 and 2. Because the current protocol for annual gill-net surveys is to set one net per sample site, we considered only sampling schemes with one sample (net set) per site each year (no resampling).

After a time series of CPE was generated for each site over a 25-year time period, we incorporated into the data set a negative trend of known magnitude (λ) . From these 1,000 data sets, a user-specified number of sites (10, 25, 50, or 100) were then randomly sampled from the population of sites. Selected sites were randomly sampled at the start of each simulation; those sites were then considered fixed and sampled throughout the 25-year sampling period. All sites were available for sampling at the start of each simulation. Data were analyzed for different sampling durations from 5 up to 25 years and analyzed for the presence of a trend. A model similar to equation (1) was used to test the null hypothesis that $\hat{\lambda} = 0$ for each data set and the test statistic was calculated and compared against a critical value ($\alpha = 0.05$). We chose $\alpha = 0.05$ for our simulations; however, the magnitude of α is a choice for decision makers to make in consideration for the costs associated with type I (i.e., concluding a trend is present when in fact one does not exist) or type II errors (i.e., concluding a trend is absent when in fact a trend does exist). The decision as to what is an acceptable error rate will vary depending on the species of interest and management goals. The model differed from equation (1) by excluding the random effect for the site × year interaction because sites were not sampled multiple times per year in the simulations. Because the data generated depict a situation in which we know the null hypothesis was false (H_0 : $\hat{\lambda} = 0$), power was estimated as the percentage of trials (out of 1,000) that rejected the null hypothesis (Wagner et al. 2007). All

simulations were performed by using the Statistical Analysis System (SAS Institute 2004).

Results

Annual Gill-Net Surveys

For the full analysis, the number of gill-net sets per year ranged from 4 during the years 1978–1983 to 51 in 2006, and the average annual CPE ranged from 43 \pm 10.3 in 1991 to 282 \pm 79.8 in 1981. For the restricted time series (1996–2006), the number of gill-net sets ranged from 9 in 2001 to 51 in 2006, with average CPE ranging from 44 \pm 13.0 in 2003 to 128 \pm 29.2 in 2002 (Figure 2).

Trends in Observed Walleye CPE

For both the full and restricted analysis, modeling the coherent year effect as a first-order autoregressive process did not improve the fit of the model (for both analyses, P>0.20). Thus, the model specified in equation (1) was deemed most parsimonious and used to examine temporal trends and estimate variance components. For the years and sites included in the full analysis, walleye CPE in ODNR gill nets exhibited a significant negative temporal trend (fixed slope estimate $\hat{\lambda}=-0.03\pm0.02$; P=0.02) from 1978 to 2006, with an average annual percent decrease of 3.4% (Table 1). However, when the analysis was restricted to the years 1996–2006 and MUs 1 and 2, CPE did not exhibit a significant temporal trend ($\hat{\lambda}=0.00\pm0.04$, P=0.91; Table 2).

Variance Components

For the full analysis, all variance components were significantly different from zero, except trend variation, which was estimated to be near zero. Site-to-site variation accounted for 26% of the total variation, whereas coherent temporal and ephemeral temporal variation contributed 15% and 43%, respectively. The significance of the coherent temporal variation can be interpreted as, in a given year, all sites tended to either have higher or lower than average CPE. The significant ephemeral temporal variation can be interpreted as follows: in addition to coherent variation, where all sites respond similarly in a given year, all sites also deviated independently from one another (e.g., in a given year, one site may have higher than average CPE, whereas another may have lower than average CPE). The unexplained error (residual variation) was 16\% of the total variation (Figure 3A).

For the restricted analysis, all variance components except ephemeral temporal variation (P = 0.18) were significantly different from zero. However, the variance estimate for ephemeral temporal variation made up 32% of the total estimated variation (Figure 3B). Because it is

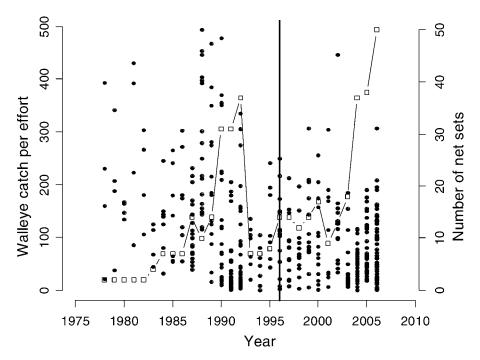


FIGURE 2.—Plot of annual site-specific catches for age-2 and older walleyes (circles; CPE) and total number of net sets (squares) from annual gill-net surveys on Lake Erie from 1978 to 2006. The vertical solid line indicates the cutoff date (1996) for the full and restricted analyses.

unlikely that sample sites did not exhibit independent yearly variation each year in CPE, and because the nonsignificance of the estimate probably resulted from the limited sample size used in this analysis, we included the ephemeral temporal variance estimate in the power analyses. Site-to-site variation accounted for 13% of the total variation, and coherent temporal variation 10%. Trend variation was 0.9% and unexplained error (residual variation) was 43% of the total variation.

TABLE 1.—Parameter estimates, standard errors, and *P*-values for the fixed intercept and slope and the random effects of site, coherent temporal, slope variation, ephemeral temporal, and residual error for gill-net catch per unit effort for walleyes in Lake Erie based on a time series from 1978 to 2006. Estimates are for the 'full' analysis; see equation 1 for explanation of model parameters.

Parameter	Estimate	SE	P-value
Fixed effects			
Intercept (µ̂)	4.02	0.14	< 0.001
Slope $(\hat{\lambda})$	-0.03	0.02	0.022
Random effects			
Site $(\hat{\sigma}_a)$	0.44	0.13	< 0.001
Coherent temporal $(\hat{\sigma}_b)$	0.25	0.09	0.004
Slope $(\hat{\sigma}_t)$	0.0	a	a
Ephemeral temporal $(\hat{\sigma}_c)$	0.74	0.11	< 0.001
Residual $(\hat{\sigma}_e)$	0.28	0.07	< 0.001

^a Not estimable.

Simulations

Power curves of walleye CPE for both analyses (full and restricted) demonstrated expected patterns with respect to changes in sampling duration, the number of sites sampled each year, and trend magnitude, power increasing as sampling duration, the number of sites sampled each year, and trend magnitude increased (Figures 4, 5). However, how rapidly power increased was dependent on the data set analyzed (the full or restricted analysis). Few consistent patterns were evident when data sets were compared, with the patterns depending on the sampling scenario considered. However, the restricted analysis had greater power to detect a trend of any magnitude considered, regardless of sampling duration, when 50 or 100 sites were sampled each year. In addition, in most cases the observed increase in power from increasing the number of sites sampled each year from 10 to 20, 50, and 100 was greater for the restricted analysis. This pattern was most noticeable for smaller trend magnitudes scenarios (i.e., 3\% and 5\% trend) after 10-15 years of sampling (Figures 4, 5). For example, for a trend magnitude of 5% decline per year, increasing the number of sites sampled each year from 10 to 100 for the full analysis resulted in an increase in power of 0.24 at year 15 (Figure 4B). For the same scenario the restricted

TABLE 2.—Parameter estimates, standard errors, and *P*-values for the fixed intercept and slope and the random effects of site, coherent temporal, slope variation, ephemeral temporal, and residual error for gill-net catch per unit effort for walleyes in Lake Erie based on management units 1 and 2 and data from 1996 to 2006. Estimates are for the 'restricted' analysis; see equation 1 for explanation of model parameters.

Parameter	Estimate	SE	P-value
Fixed effects			
Intercept $(\hat{\mu})$	4.15	0.13	< 0.001
Slope $(\hat{\lambda})$	0.00	0.04	0.910
Random effects			
Site $(\hat{\sigma}_a)$	0.12	0.06	0.002
Coherent temporal $(\hat{\sigma}_b)$	0.09	0.06	0.002
Slope $(\hat{\sigma}_t)$	0.01	0.01	0.003
Ephemeral temporal $(\hat{\sigma}_c)$	0.28	0.26	0.180
Residual $(\hat{\sigma}_e)$	0.38	0.25	< 0.001

analysis resulted in an increase in power of 0.66 at year 15 (Figure 5B).

The power to detect temporal trends of a 3-5%annual decline in CPE over a relatively short sampling duration (5-10 years) remained low regardless of the number of sites sampled each year or which data set was used for the analysis. A trend in CPE could be detected within 5-10 years only for relatively large trend magnitudes. For example, in the full analysis, if 10 sites were sampled each year for 10 years, the power to detect a temporal trend did not approach 0.8 until the magnitude of the decline was 20\% per year (power = 0.78). However, if 50 sites were sampled each year, a 10% annual decline could be detected with greater than 0.80 power in 15 years; a 5% annual decline, on the other hand, would still not be detected with power greater than 0.8 for approximately 22 years (Figure 4). Similarly, the power of the restricted analysis to detect temporal trends remained low for a trend magnitude of 3\%, regardless of the sample duration and the number of sites sampled. For instance, when 100 sites were sampled, the power to detect a temporal tend after 25 years of sampling was only 0.62. Sampling 100 sites and assuming a 5% annual decline, the power to detect trends did not exceed 0.80 until 17 years of sampling (Figure 5). However, power was relatively high for detecting trends of large magnitude (e.g., 20% decline per year) for the restricted analysis. For instance, if 100 sites were sampled each year, power exceeded 0.80 after 10 and 6 years for trend magnitudes of 10% and 20%, respectively, for the restricted analysis.

Discussion

The percent declines we considered for this analysis are within the range of values observed in walleyes and other important fisheries, including marine and freshwater fisheries (Wilberg et al. 2005; Campana et al.

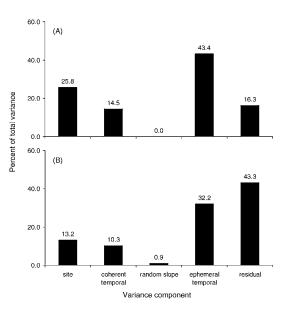


FIGURE 3.—Estimated percent of total variation for (A) the full analysis and (B) the restricted analysis attributed to site, coherent temporal, ephemeral temporal, trend (random slope), and residual variance. Estimates are from a mixed model for $\log_e(\text{total walleye catch})$ versus time. The time series ranged from 1978 to 2006 and from 1996 to 2006 for the full and restricted analyses, respectively.

2006; Irwin et al. 2008). However, given the sampling scenarios we investigated, the statistical power to detect temporal trends in walleye CPE was low over short to moderate sampling durations (e.g., 5–10 years) unless the magnitude of annual change in CPE was relatively large (e.g., 10–20% per year decline). Although these analyses demonstrated that increasing the number of sites sampled each year increased power, even under a scenario where 100 sites were sampled each year, the power to detect small changes (e.g., 3\% per year decline) remained low over relatively long sampling durations. Zwieten et al. (2002) also concluded that the ability to detect trends in catch rates was low over relatively short time periods, finding that it would take 30 years of observations to detect a 1.6% per year decrease in total catch rates in an industrial pelagic purse-seine fishery in northern Lake Tanganyika with 90% power. When assessing the power of the English bottom trawl survey, Maxwell and Jennings (2005) found that even an annual decrease of 50% in adult abundance for the less abundant species caught in the surveys, such as the thornback ray Raja clavata, spurdog Squalus acanthias, and spotted ray R. montagui, would have low power to be detected after 5 years of sampling. However, for abundant species, such as dab Limanda

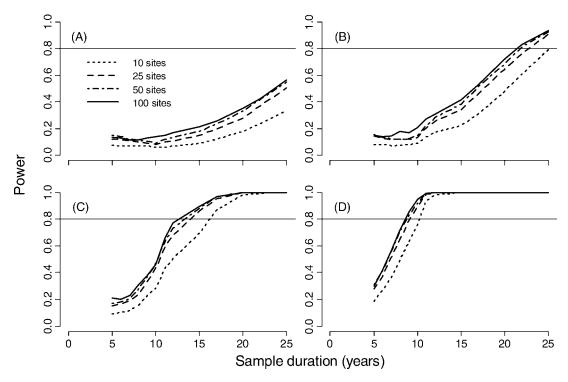


FIGURE 4.—Power curves for the full analysis (years 1978–2006) for detecting temporal trends in gill-net CPE for walleyes in Lake Erie with increasing numbers of fixed sites sampled per year (10, 25, 50, or 100) and increasing trend magnitude: (A) 3% decline per year; (B) 5% decline per year; (C) 10% decline per year; and (D) 20% decline per year. The solid horizontal line denotes power = 0.80.

limanda and long rough dab Hippoglossoides plates-soides, the power to detect a decline of more than 20% after 5 years was relatively high (Maxwell and Jennings 2005). Taken together with our analysis, these results suggest that FI surveys often have relatively low power to detect small temporal changes in CPE over relatively short sampling durations. This is important to quantify because low to modest changes in CPE or other indices may not be detectable over management-relevant time scales. However, the ability of a survey to detect trends probably will also vary among species and systems according to species-specific spatial and temporal dynamics; thus, broadly generalizing the results of our analysis to all FI surveys is not prudent.

The low power to detect certain trends as indicated by this case study does not imply reduced importance of these surveys for fulfilling many management objectives, including collecting information for population models. In addition, because the power to detect trends in CPE not only depends on the structure of variation but also is a function of the survey design (Urquhart et al. 1998; Larsen et al. 2004), FI surveys must be evaluated with regard to components of

variation in conjunction with survey design in an effort to maximize capabilities for trend detection. The statistical power of a test also depends on the type I and type II error rates deemed acceptable by managers. Because the costs associated with a type I error are often lower than those for a type II error, managers may decide to retain a higher type I error rate (e.g., $\alpha = 0.10$; Taylor and Gerrodette 1993; Strayer 1999; Maxwell and Jennings 2005). Examining the effects of changing the type I error rate is easily incorporated into the variance component approach.

We cannot determine whether the observed changes in variance structures between the full and restricted data sets are related to indirect effects of zebra mussels on sampling efficiency (i.e., water clarity effects) or to potential changes in population dynamics. However, the decision regarding what data to include in the analyses for trend detection affected how the spatial and temporal dynamics of the population were viewed through altering how the total variance was partitioned. For example, including the entire time series or just using a subset of the time series altered the estimated structure of variation, and this change in estimated variance structure resulted in different inferences being

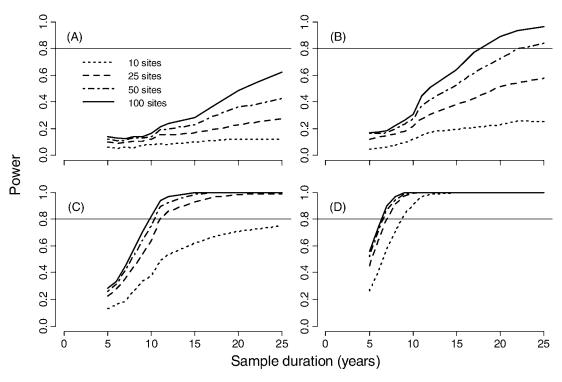


FIGURE 5.—Power curves for the restricted analysis (years 1996–2006) for detecting temporal trends in gill-net CPE for walleyes in Lake Erie with increasing numbers of fixed sites sampled per year (10, 25, 50, or 100) and increasing trend magnitude: (**A**) 3% decline per year; (**B**) 5% decline per year; (**C**) 10% decline per year; and (**D**) 20% decline per year. The solid horizontal denotes power = 0.80.

made about trends in population abundance. The nonsignificance of the trend variance component in the full analysis implied that each sample site had a similar trend over time, equal to an average annual percent decrease of 3.4%. However, the significant trend variation in the restricted analysis suggested that walleyes sampled at different sites were deviating from one another in their average trend, some increasing in CPE and others decreasing. The differing patterns between the full and restricted analyses pattern could be reflecting changes in fish distribution or habitat over time. We suspect that the magnitude of trend variation will often be sensitive to the length of time series used in the analysis. Over moderate time periods, abundance at different sites may often show different trends, even though over long time horizons each site follows the regional trend. If so, more careful definition is needed for what kind of trends are of interest for evaluating power. If moderate-term (e.g., 10 year) trends are of interest, then trend variation over this time frame should be considered in the analysis. If power to detect long-term trends is of primary interest, the shorter-term

trend variation might be better treated as correlation within the ephemeral (site × year interaction) variation.

The differences in variance estimates also had implications for the perceived ability to detect changes in abundance over time. For example, on average, the ability to detect a temporal trend after 10 years was lower for the full analysis than for the restricted analysis. This lower power was largely influenced by the larger percentage and magnitude of the coherent temporal variance component estimated for the full analysis ($\hat{\sigma}_b = 0.25 \pm 0.09$) than that for the restricted analysis ($\hat{\sigma}_b = 0.09 \pm 0.06$). The larger coherent temporal variance also influenced the observed pattern where, in the full analysis, the gain in power when the number of sample sites was increased was smaller than for the restricted analysis. This inability to gain power by changing aspects of the sampling design (i.e., adding more sample sites) when large coherent temporal variability is present reflects that the influence of coherent temporal variation on power cannot be reduced by changing specific design details of the sampling protocol. This is in contrast to other sources of variation examined in this study (see Urquhart et al.

1998 for additional details on specific variance components). Previous studies have demonstrated the large influence that coherent temporal variation has on reducing the power to detect temporal trends (Urguhart et al. 1998; Wagner et al. 2007). Those findings imply that, all else being equal, if a large proportion of the total variation in a fish populations is composed of coherent temporal variation, then that population will inherently be more difficult to detect temporal trends in than will fish populations with smaller coherent temporal patterns, regardless of what survey design is employed. To illustrate this point for the CPE data, we considered a scenario using the full data set where there was a 5% annual decline in CPE and 50 sites were sampled each year. We then reduced coherent temporal variation by half as well as setting it equal to zero and compared those results with those for the situation in which the estimated value for coherent temporal variance was used. Under this situation, the power to detect a 5% trend in 10 years increased from 0.15 for a situation using the estimated coherent temporal variance to over 0.8 when there was no coherent temporal variation. Because of the large influence of coherent temporal variability on trend detection capabilities, annual covariates can also be included in the modeling exercise in an attempt to explain this source of variation and to increase power. For example, for the restricted analysis we used as covariates annual average Secchi disk depth and water temperature data (averaged across sample sites within each year), which were taken at the time nets were set, in an attempt to explain coherent temporal variation. However, in our case neither Secchi disk depth nor water temperature explained coherent temporal variation (P = 0.35 for Secchi depth and P = 0.87 for water temperature). However, we recommend that monitoring programs measure physiochemical and additional biological data hypothesized to influence the coherent dynamics of the fish population being monitored. Including these data may not only allow for an improved estimate of power but also lend insight into important biological processes that influence all sample sites similarly within a given year.

It was not the purpose of this study to perform a thorough evaluation of survey designs and how they interact with components of variation to influence trend detection power for walleyes in Lake Erie. Rather, we wanted to illustrate the need to evaluate FI surveys and the utility of quantifying various spatial and temporal sources of variation to address management questions related to FI survey design and expectations. However, the evaluation of survey design is a logical next step once the most precise variance components are estimated based on the available data.

Recommendations and Future Directions

Our analysis focused on time series data for walleyes collected by the ODNR collected during 1978-2006. We realize that other freshwater systems may not have the benefit of long-term data sets such as those available for many species in the Great Lakes. However, because of the widespread use of FI surveys to monitor fish populations, analyses can be performed that take advantage of multiple time series already collected for different species within a given region. For instance, our analysis was limited to CPE data for Ohio waters, and including time series from other parts of Lake Erie (and other Great Lakes) would provide further insight into the structure of variation, which ultimately drives the results and interpretations of the power analyses. Such time series, even if they are of differing lengths and quality and are taken from multiple systems, can be analyzed in combination to estimate variance components. For example, by utilizing methods similar to those used in this study, a meta-analysis type approach could be used across systems, where multiple time series for a given species are analyzed simultaneously. Because FI surveys collected from different agencies contain data that were collected at different frequencies and some may contain missing data, some FI surveys may contain more or less information about the spatial and temporal dynamics of fish populations. A meta-analysis type of an approach will allow surveys containing more information to help inform those surveys or systems that contain less information. This type of analysis will result in more precise variance estimates, which can then be used in a comprehensive evaluation of FI surveys for commercially and ecologically important fish species. The variance components so obtained can then be used to establish expectations for the trend detection capabilities of FI surveys and ultimately to evaluate and improve FI survey designs. A metaanalysis type of an approach will also benefit both relatively data-rich systems, such as the Great Lakes, and relatively data-poor systems, which may include some inland lakes and streams. In addition, such an analysis will help determine whether sampling designs deemed "optimal" vary among systems, species, or life stages (juvenile or adult).

Lastly, this approach lends itself to the development and comparison of fishery indicators. Managers should evaluate multiple indicators, when appropriate, to ensure that they are monitoring an indicator that represents specific management objectives, particularly given the time and expense involved in monitoring. For example, composite indicators that reflect trends in CPE for multiple species can be developed and

evaluated (Maxwell and Jennings 2005). Composite indicators or indicators of the whole fish community may exhibit lower variability than for individual fish species, and thus may represent more sensitive indicators for some management objectives (Smokorowski and Kelso 2002).

Before performing such analyses, however, specific objectives of the monitoring program must be specified, along with the changes in CPE that are desired to be detectable. Although we focused on detecting temporal trends in this study, we recognize that FI surveys are often used to fulfill multiple objectives. Thus, explicitly stating the goals and objectives of a monitoring program will help ensure that the survey design utilized will maximize the ability to meet multiple objectives. The evaluation of FI surveys and potential improvements to existing survey design is important not only to ensure that management actions can be confidently evaluated over the short-term but also to establish realistic expectations for what kind of changes can be detected.

Acknowledgments

We thank Jim Bence, Travis Brenden, Mike Jones, Brian Irwin, and Jon Deroba of the Quantitative Fisheries Center at Michigan State University for helpful comments and contributions that improved this manuscript. Gill-net data were collected by fisheries biologists stationed at the Sandusky and Fairport Fisheries Research Stations, Ohio Department of Natural Resources, and was completed under Federal Aid in Sport Fish Restoration Project F-69-P. This is manuscript 2009-08 of the Quantitative Fisheries Center.

References

- Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. *In B. N. Petrov and F. Csaki*, editors. Second International Symposium on Information Theory. Akademiai Kiado, Budapest.
- Allen, L. G., D. J. Pondella, II, and M. A. Shane. 2007. Fisheries-independent assessment of a returning fishery: abundance of juvenile white seabass (*Atractoscion nobilis*) in the shallow nearshore waters of the Southern California Bight, 1995–2005. Fisheries Research 88:24–32.
- Ault, J. S., G. A. Diaz, S. G. Smith, J. Luo, and J. E. Serafy. 1999. An efficient sampling survey design to estimate pink shrimp population abundance in Biscayne Bay, Florida. North American Journal of Fisheries Management 19:696–712.
- Blanchard, J. L., D. L. Maxwell, and S. Jennings. 2008. Power of monitoring surveys to detect abundance trends in depleted populations: the effects of density-dependent habitat use, patchiness, and climate change. ICES Journal of Marine Science 65:111–120.

- Campana, S. E., L. Marks, J. Warren, and N. E. Kohler. 2006. Effects of recreational and commercial fishing on blue sharks (*Prionace glauca*) in Atlantic Canada, with inferences on the North Atlantic population. Canadian Journal of Fisheries and Aquatic Sciences 63:670–682.
- Corradin, L. M., M. J. Hansen, D. R. Schreiner, and M. J. Seider. 2008. Recruitment dynamics of lake trout in western Lake Superior during 1988–1995. North American Journal of Fisheries Management 28:663–677.
- Hayes, D., E. Baker, R. Bednarz, D. Borgeson, Jr., J. Braunscheidel, J. Breck, M. Bremigan, A. Harrington, R. Hay, R. Lockwood, A. Nuhfer, J. Schneider, P. Seelbach, J. Waybrant, and T. Zorn. 2003. Developing a standardized sampling program: the Michigan experience. Fisheries 28(7):18–25.
- Irwin, B. J., T. J. Treska, L. G. Rudstam, P. J. Sullivan, J. R. Jackson, A. J. VanDeValk, and J. L. Forney. 2008. Estimating walleye (*Sander vitreus*) density, gear catchability, and mortality using three fishery-independent data sets for Oneida Lake, New York. Canadian Journal of Fisheries and Aquatic Sciences 65:1366–1378.
- Kincaid, T. M., D. P. Larsen, and N. S. Urquhart. 2004. The structure of variation and its influence on the estimation of status: indicators of condition of lakes in the northeast U.S.A. Environmental Monitoring and Assessment 98:1– 21.
- Larsen, D. P., P. R. Kaufmann, T. M. Kincaid, and N. S. Urquhart. 2004. Detecting persistent change in the habitat of salmon-bearing streams in the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Sciences 61:283–291.
- Larsen, D. P., T. M. Kincaid, S. E. Jacobs, and N. S. Urquhart. 2001. Designs for evaluating local and regional scale trends. Bioscience 51:1069–1078.
- Littell, R. C., G. A. Milliken, W. W. Stroup, and R. S. Wolfinger. 1996. SAS system for mixed models. SAS Institute, Cary, North Carolina.
- Locke, B., M. Belore, A. Cook, D. Einhouse, R. Kenyon, R. Knight, K. Newman, P. Ryan, and E. Wright. 2005. Lake Erie walleye management plan. Great Lakes Fishery Commission, Ann Arbor, Michigan. Available: www.glfc.org/lakecom/lec/WTG.htm. (November 2008).
- Maxwell, D., and S. Jennings. 2005. Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. Journal of Applied Ecology 42:25–37.
- McInerny, M. C., and T. K. Cross. 2006. Factors affecting trap-net catchability of black crappies in natural Minnesota lakes. North American Journal of Fisheries Management 26:652–664.
- Niles, J. M., and K. J. Hartman. 2007. Comparison of three larval fish gears to sample shallow-water sites on a navigable river. North American Journal of Fisheries Management 27:1126–1138.
- ODNR (Ohio Department of Natural Resources). 2008. Ohio's Lake Erie fisheries 2007. ODNR, Sandusky.
- Olsen, A. R., J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Young. 1999. Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54:1–45.
- Pennington, M., and T. Strømme. 1998. Surveys as a research

tool for managing dynamic stocks. Fisheries Research 37:97–106.

- Pothoven, S. A., and C. P. Madenjian. 2008. Changes in consumption by alewives and lake whitefish after dreissenid mussel invasions in Lakes Michigan and Huron. North American Journal of Fisheries Management 28:308–320.
- Quinn, T. J., and R. B. Deriso. 1999. Quantitative fish dynamics. Oxford University Press, New York.
- SAS Institute. 2004. SAS/STAT 9.1 user's guide. SAS Institute, Cary, North Carolina.
- Self, S., and K.-Y. Liang. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. Journal of the American Statistical Association 82:605–610.
- Sitar, S. P., J. R. Bence, J. E. Johnson, M. P. Ebener, and W. W. Taylor. 1999. Lake trout mortality and abundance in southern Lake Huron. North American Journal of Fisheries Management 19:881–900.
- Smith, S. J., and M. J. Tremblay. 2003. Fishery-independent trap surveys of lobsters (*Homarus americanus*): design considerations. Fisheries Research 62:65–75.
- Smokorowski, K. E., and J. R. M. Kelso. 2002. Trends in fish community structure, biomass, and production in three Algoma, Ontario, lakes. Water, Air, and Soil Pollution: Focus 2:129–150.
- Stapanian, M. A., C. P. Madenjian, and J. Tost. 2007. Regional difference in size at age of the recovering burbot (*Lota lota*) population in Lake Erie. Journal of Great Lakes Research 33(Supplement 1):91–102.
- Stobutzki, I. C., G. T. Silvestre, A. Abu Talib, A. Krongprom, M. Supongpan, P. Khemakorn, N. Armada, and L. R. Garces. 2006. Decline of demersal coastal fisheries resources in three developing Asian countries. Fisheries Research 78:130–142.
- Strayer, D. 1999. Statistical power of presence-absence data

- to detect population declines. Conservation Biology 13:1034–1038.
- Taylor, B. L., and T. Gerrodette. 1993. The uses of statistical power in conservation biology: the vaquita and northern spotted owl. Conservation Biology 7:489–499.
- Urquhart, N. S., and T. M. Kincaid. 1999. Designs for detecting trend from repeated surveys of ecological resources. Journal of Agricultural, Biological, and Environmental Statistics 4:404–414.
- Urquhart, N. S., S. G. Paulsen, and D. P. Larsen. 1998. Monitoring for policy-relevant regional trends over time. Ecological Applications 8:246–257.
- VanLeeuwen, D. M., L. W. Murray, and N. S. Urquhart. 1996.
 A mixed model with both fixed and random trend components across time. Journal of Agricultural, Biological, and Environmental Statistics 4:435–453.
- Wagner, T., J. R. Bence, M. T. Bremigan, D. B. Hayes, and M. J. Wilberg. 2007. Regional trends in fish mean length at age: components of variance and the power to detect trends. Canadian Journal of Fisheries and Aquatic Sciences 64:968–978.
- Wilberg, M. J., J. R. Bence, B. T. Eggold, D. Makauskas, and D. F. Clapp. 2005. Yellow perch dynamics in southwestern Lake Michigan during 1986–2002. North American Journal of Fisheries Management 25:1130– 1152.
- Yule, D. L., J. V. Adams, J. D. Stockwell, and O. T. Gorman. 2008. Factors affecting bottom trawl catches: implications for monitoring the fishes of Lake Superior. North American Journal of Fisheries Management 28:109–122.
- Zwieten, P. A. M., F. C. Roest, M. A. M. Machiels, and W. L. T. van Densen. 2002. Effects of interannual variability, seasonality, and persistence on the perception of long-term trends in catch rates of the industrial pelagic purseseine fishery of northern Lake Tanganyika. Fisheries Research 54:329–348.