The Statistical Power to Detect Temporal Trends in Catch per Unit Effort from Annual Gillnet Surveys for Walleye in Lake Erie

Prepared for:
The Ohio Department of Natural Resources
Division of Wildlife
Lake Erie Fisheries Unit
Sandusky Fisheries Research Unit
305 East Shoreline Drive
Sandusky, OH 44870

Prepared by:
Tyler Wagner, Ph.D.
Quantitative Fisheries Center
Department of Fisheries and Wildlife
Michigan State University
East Lansing, MI 48824

4 September 2007

Summary

1. Catch per unit effort data (CPUE) from annual walleye gillnet surveys were analyzed to examine the statistical power to detect temporal trends as a function of (1) trend magnitude, and (2) the number of fixed sample sites sampled each year.
2. Overall, walleye CPUE exhibited a negative trend over time, with an annual average decrease of 3.4%.
3. Increasing the number of sites sampled per year, the trend magnitude, and the sample duration increased the statistical power to detect temporal trends in walleye CPUE.
4. Power analysis suggests that the power to detect trends over the short-term (e.g., 10 years) is low regardless of the number of sites sampled.
5. With an actual trend magnitude similar to that estimated from the walleye CPUE analysis (i.e., 3% decline per year) and with sampling 50 sites per year, power was only 0.55 after sampling for 25 years.
6. Overall, the power to detect relatively small changes in walleye CPUE (e.g., a 3% decrease per year) remains low even after an extended sampling duration (> 25 years).

Introduction

The Ohio Department of Natural Resources (ODNR), Division of Wildlife, Lake Erie Fisheries Unit is interested in the statistical power to detect temporal trends in catch per unit effort (CPUE) of walleye sampled in annual gillnet surveys. Specifically, the ODNR is interested in the following questions:

(1) Is the current number of sites (i.e., grids) sampled per year sufficient to detect a decrease in CPUE?

(2) Given a pre-specified number of sites sampled per year, what trend magnitude can be detected with relatively high power (i.e., power ≥ 0.8)?

(3) Is the current number of nets set in each management unit (Figure 1) sufficient to detect a change in walleye CPUE?

(4) Is sample depth an important driver of walleye CPUE?

(5) Is management unit an important driver of walleye CPUE?

Because of small sample sizes, I was unable to address questions 4 and 5. The dataset I was given only had sample depth and management unit codes for the years 2004 – 2006. This small sample size prohibited examining the importance of these factors on walleye CPUE. Preliminary examination of available data suggests that sample depth and management unit may be important drivers of walleye CPUE given observed differences among depth categories and management units. The collection of additional data over the next several years will allow these questions to be examined in more detail. Alternatively, if pre-2004 data currently are available, questions 3 and 4 could be addressed now.

I used statistical and simulation modeling to address questions 1, 2, and 3. It should be noted, however, that under the current sampling protocol (ODNR usually setting one net per grid (i.e., sample site)), questions 1 and 3 are actually the same questions (if we pool across management units): “is the number of sites currently sampled sufficient to detect a change in walleye CPUE?” Briefly, our approach consisted of using mixed models to partition the total variability in log-total catch into several spatial and temporal components. Variance estimates were then used to generate simulated data to examine the effects of the number of sample sites and trend magnitude on the statistical power to detect temporal trends in walleye CPUE. For additional details see Wagner et al. (2007).

It should be noted that this analysis addresses sampling questions related to the statistical power to detect temporal trends in walleye CPUE. However, walleye surveys are used to fulfill multiple objectives (e.g., stock assessment models, determine age composition), so an 'optimal' design for one objective (i.e., trend detection) may not be the optimal design for other objectives. Thus, these results should be interpreted in the light of multiple, and sometimes competing objectives, and so the question of what is the optimal sampling design becomes much more complicated.

Components of variance background

A components of variance approach has been advocated to address the issue of variability in ecological data when evaluating regional temporal trends and monitoring of ecological systems (Urquhart et al. 1998; Larsen et al. 2001; Kincaid et al. 2004). Under this framework, total variance is partitioned into four components including, (1) site-to-site (spatial) variation, (2) coherent (year-to-year) variation affecting all sites (e.g., grids) in a similar manner, (3) ephemeral temporal variation (e.g., site-by-year interaction) corresponding to independent yearly variation at each site, and (4) residual variation (Larsen et al. 2001; Kincaid et al. 2004). A fifth component can be included in this framework in which each site has its own trend (i.e., trend variation: allowing the slope of the response variable versus time at each site to be a draw from a distribution and estimate the variance of the distribution of slopes; VanLeeuwen et al. 1996).

Methods

A mixed model was used to assess the presence of any trend in walleye CPUE and to obtain estimates of variance components for use in simulation modeling. Because sample sizes did not allow elucidation of the importance of sample depth and management unit on walleye CPUE, I used data from all depths, years, and management units to quantify the magnitude of spatial and temporal components of variance. Although only data from management units 1 and 2 are currently used in stock assessment models, I argue that the question of trend detection is best addressed by using all available data. Tables 1 and 2 summarize the data used in the analysis.

The mixed model used for the analyses was

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

where Y_{ijk} is the log total catch for sample *k* at site *i* in year *j*, μ and λ are the fixed intercept and slope (fixed trend), respectively. The random effect a_i is a random effect for site *i*, representing site-to-site variability, iid as $N(0, \sigma_a^2)$, b_j is a random effect for the j^{th} year (coherent temporal variability), iid as $N(0, \sigma_b^2)$, t_i is a random effect for the trend for site *i*, iid as $N(0, \sigma_t^2)$, c_{ij} is the site$\times$year interaction (ephemeral temporal variability), iid as $N(0, \sigma_c^2)$, and e_{ijk} is the unexplained error (residual error), iid as $N(0, \sigma_e^2)$. The year covariate (y) is the j^{th} year minus the mean year used in the analysis. This standardization of year was performed to provide numerical stability. I estimated variance components using restricted maximum likelihood and *P*-values using a likelihood ratio test (Self and Liang 1987; Littell et al. 1996). I considered all analyses significant at $P < 0.05$.

Power analysis

I investigated the extent to which the following factors affected the ability to detect a temporal trend in walleye CPUE (1) increasing trend magnitude (λ ranged from a 3 – 20 % decrease per year), and (2) increasing the number of sites sampled (including sampling 10, 25, or 50 fixed sites each year). I used a simulation approach to examine the statistical power to detect temporal trends using the variance components estimated from equation 1. For each simulation, one thousand datasets were generated containing CPUE data for a population of sites. I ran the simulations using two population sizes from which potential sites were sampled from. First, the population of sites was set at 305 because this corresponds to the total number of grids that could be potentially sampled (102 grids in management unit 1, 130 grids in management unit 2, and 73 grids in management unit 3). Secondly, I used a population of sites set at 100 to represent sampling from a single management unit to examine the effects of sampling one management unit intensively. After a time series of CPUE was generated for each site over a 25 year time period, a trend of known magnitude (e.g., a decrease of 3% per year) was incorporated into the dataset. From these 1000 datasets, a user-specified number of sites (10, 25, 50) were then randomly sampled from the population. Sites (grids) were randomly sampled at the start of each simulation and those sites were 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.

The model specified in equation 1 was used to test the null hypothesis that $\hat{\lambda} = 0$ for each dataset and the test statistic was calculated and compared to a critical value ($\alpha = 0.05$). Because the data generated depict a situation in which we know the null hypothesis is false (i.e., a trend of known magnitude was incorporated into the data), power was estimated as the percentage of trials (out of 1000) that rejected the null hypothesis.

Results

Trends in walleye gillnet CPUE

Walleye CPUE exhibited a significant (P = 0.02) negative trend, with an average annual percent decrease of 3.4% (Table 1; Figure 2).

Variance components

All variance components were significantly different from zero, except trend variation which was estimated near zero. The nonsignificance of the trend variation suggests that each sample site has a similar trend over time, equal to the average trend, of an average annual percent decrease of 3.4% (Table 1). Site-to-site variation comprised 26% of the total variation, whereas, coherent temporal an ephemeral temporal variation comprised 15 and 43% respectively. The significance of the coherent temporal variation can be interpreted as, that in a given year all sites tended to either have higher or lower than average CPUE. The significant ephemeral temporal variation can be interpreted as, that in addition to coherent variation where all sites respond similarly in a given year, that all sites also deviated independently from one another (e.g., in a given year one site may have higher than average CPUE, while another may have lower than average CPUE). The unexplained error (residual variation) was 16% of the total variation (Figure 3).

Power analysis

Given that residual variance was low relative to ephemeral temporal variance I only considered sampling schemes with one sample per site each year. The power to detect temporal trends in walleye CPUE was dependent on the number of years sampled, the number of sites sampled per year, and the magnitude of the trend (Figure 4). The power to detect temporal trends was not dependent, however, on the initial population of sites assumed in the simulation (i.e., 305 or 100). In fact the power curves were nearly identical to those illustrated in Figure 4 (data not shown). As expected, regardless of the number of sites sampled per year or the trend magnitude, the power to detect a trend increases with increasing sampling duration. However, how rapidly power increased over time depended on the number of sites and trend magnitude. In addition to power increasing with sampling duration, it increased with increasing trend magnitude and with increasing number of sites sampled each year.

The power to detect temporal trends remains fairly low for all sample sizes and trend magnitudes over a short sampling duration. For example, if 10 sites are sampled each year, for 10 years, the power to detect a temporal trend does not approach 0.8 until the magnitude of the decline is 20% per year (power = 0.78). However, if 50 sites are sampled each year a 10% annual decline can be detected with > 0.80 power in 15 years, but a 5% annual decline will still not be detected with > 0.8 power for approximately 22 years (Figure 4).

Discussion

The statistical power to detect changes in walleye CPUE is low over short sample durations (e.g., 10 years) unless the magnitude of change in CPUE is large. Although this analysis demonstrated that increasing the number of sample sites sample each year increased power, the power to detect smaller changes (e.g., 3%) remains low, even over a relatively long sampling duration (e.g., 25 years). Previous studies have demonstrated that significant coherent temporal variation has a large effect on reducing the power to detect temporal trends (Urquhart et al. 1998; Wagner et al. 2007), and coherent temporal variation was a significant source of variation for the walleye CPUE data. Unfortunately, the influence of coherent temporal variation on power can not be reduced by changing aspects of the sampling design. This is in contrast with ephemeral temporal variation, which can be reduced by adding more sample sites to the monitoring program. Therefore, it is likely that the relatively large coherent temporal variation is contributing to the low power to detect trends over short- to moderate timescales.

The power to detect trends in walleye CPUE was not sensitive to the number of sites assumed available for sampling each year (i.e., 305 vs. 100). This implies that even if a single management unit is intensively sampled, the power to detect trends will be relatively low over the short-term. However, this analysis did not account for potential differences in variance components among management units; therefore, the power to detect temporal trends may vary among management units based on management unit specific patterns in variability.

The power analysis reported here is the statistical power to detect temporal trends in walleye CPUE from annual gillnets surveys. It does not provide information on, for example, the power to detect if average CPUE in years 1 – 5 differ from the average CPUE in years 6 – 10. Nor does this analysis provide information on what the best sampling design is for collecting data for stock assessment models. Therefore, caution should be used if attempting to extrapolate these results to other analyses that utilize the gillnet survey data.

Acknowledgments

I thank Drs. Jim Bence, Travis Brenden, and Mike Jones, and members of the Quantitative Fisheries Center at Michigan State University for helpful comments and contributions that improved this report. This is technical report T2007-01 of the Quantitative Fisheries Center.

References

Kincaid, T. M., Larsen, D. P., and Urquhart, N. S. 2004. The structure of variation and its influence on the estimation of status: indicators of condition of lakes in the northeast, U.S.A. Environ. Monit. Assess. 98:1-21.

Larsen, D. P., Kincaid, T. M., Jacobs, S. E., and Urquhart, N. S. 2001. Designs for evaluating local and regional scale trends. Bioscience 51:1069-1078.

Littell, R. C., Milliken, G. A., Stroup, W. W., and Wolfinger, R. S. 1996. SAS System for Mixed Models. SAS Institute Inc., Cary, North Carolina.

Self, S., and Liang, K-Y. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J. Amer. Stat. Assoc. 82:605-610.

Thomas, M., and 9 coauthors. 2006. Report for 2005 by the Lake Erie Walleye Task Group, March, 2006.

Urquhart, N. S., Paulsen, S. G., and Larsen, D. P. 1998. Monitoring for policy-relevant regional trends over time. Ecol. Apps. 8:246-257.

VanLeeuwen, D. M., Murray, L. W., and Urquhart, N. S. 1996. A mixed model with both fixed and random trend components across time. J. Agricult. Biol. Environ. Stats. 4:435-453.

Wagner, T., Bence, J. R., Bremigan, M. T., Hayes, D. B., and Wilberg, M. J. 2007. Regional trends in fish mean length at age: components of variance and the power to detect trends. Can. J. Fish. Aquat. Sci. 64:968-978.

Table 1. The number of samples (gillnet sets; n) per year.

Year	n
1978	4
1979	4
1980	4
1981	4
1982	4
1983	4
1984	7
1985	7
1986	7
1987	20
1988	20
1989	14
1990	31
1991	35
1992	37
1993	7
1994	7
1995	8
1996	14
1997	14
1998	12
1999	14
2000	17
2001	9
2002	13
2003	19
2004	37
2005	39
2006	51

Table 2. The number of samples (gillnet sets; n) per management unit. Note that these numbers are for 2004, 2005, and 2006 only. Management unit numbers for the years 1978 – 2003 were unavailable.

Management unit	n
1	34
2	63
3	30

Table 3. Parameter estimates, standard errors, and P-values for the fixed intercept and slope, and random effects of site, coherent temporal, slope variation, ephemeral temporal, and residual error for gillnet catch per unit effort for walleye in Lake Erie. n.e. = not estimable. See equation 1 for explanation of model parameters.

Parameter	Estimate	Standard error	P-value
Fixed effects			
Intercept $(\hat{\mu})$	4.02	0.14	<0.0001
Slope $(\hat{\lambda})$	-0.03	0.015	0.022
Random effects			
Site $(\hat{\sigma}_a)$	0.44	0.13	0.0004
Coherent temporal $(\hat{\sigma}_b)$	0.25	0.09	0.004
Slope $(\hat{\sigma}_t)$	0.0	n.e.	n.e.
Ephemeral temporal $(\hat{\sigma}_c)$	0.74	0.11	<0.0001
Residual $(\hat{\sigma}_e)$	0.28	0.07	<0.0001

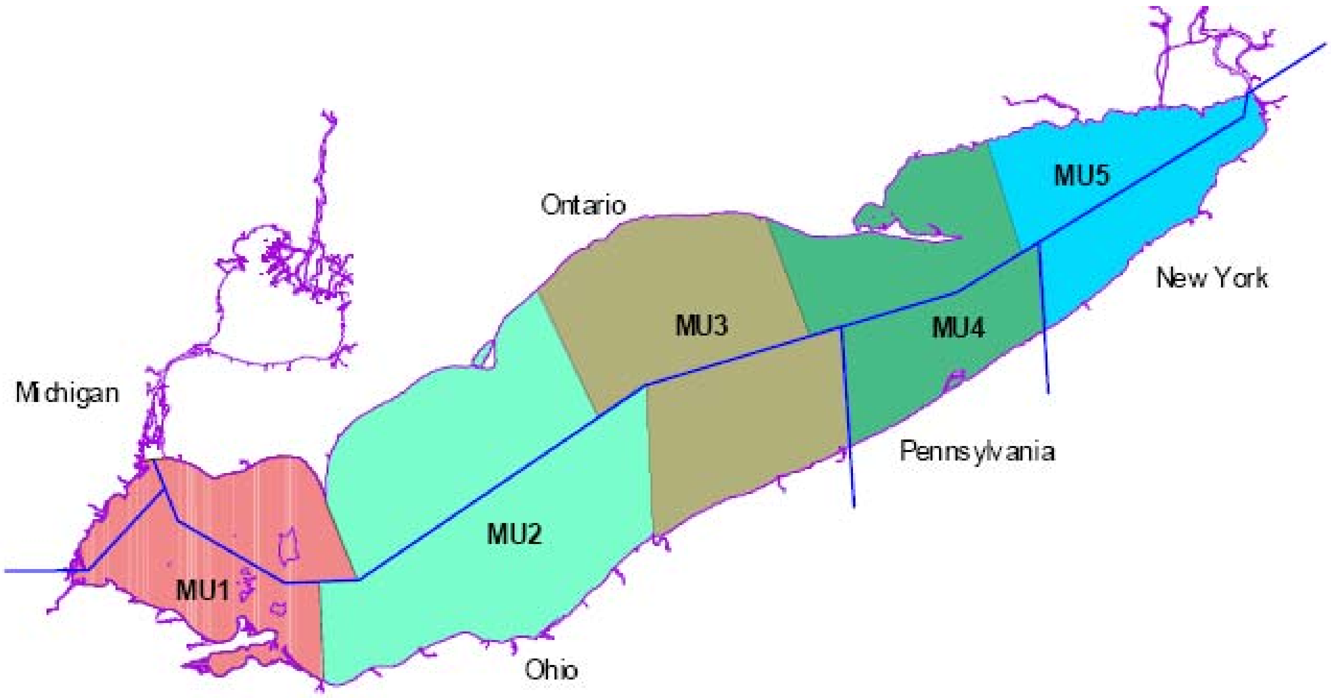

Figure 1. Map of Lake Erie with management units (MU) recognized by the Walleye Task Group (from Thomas et al. 2006).

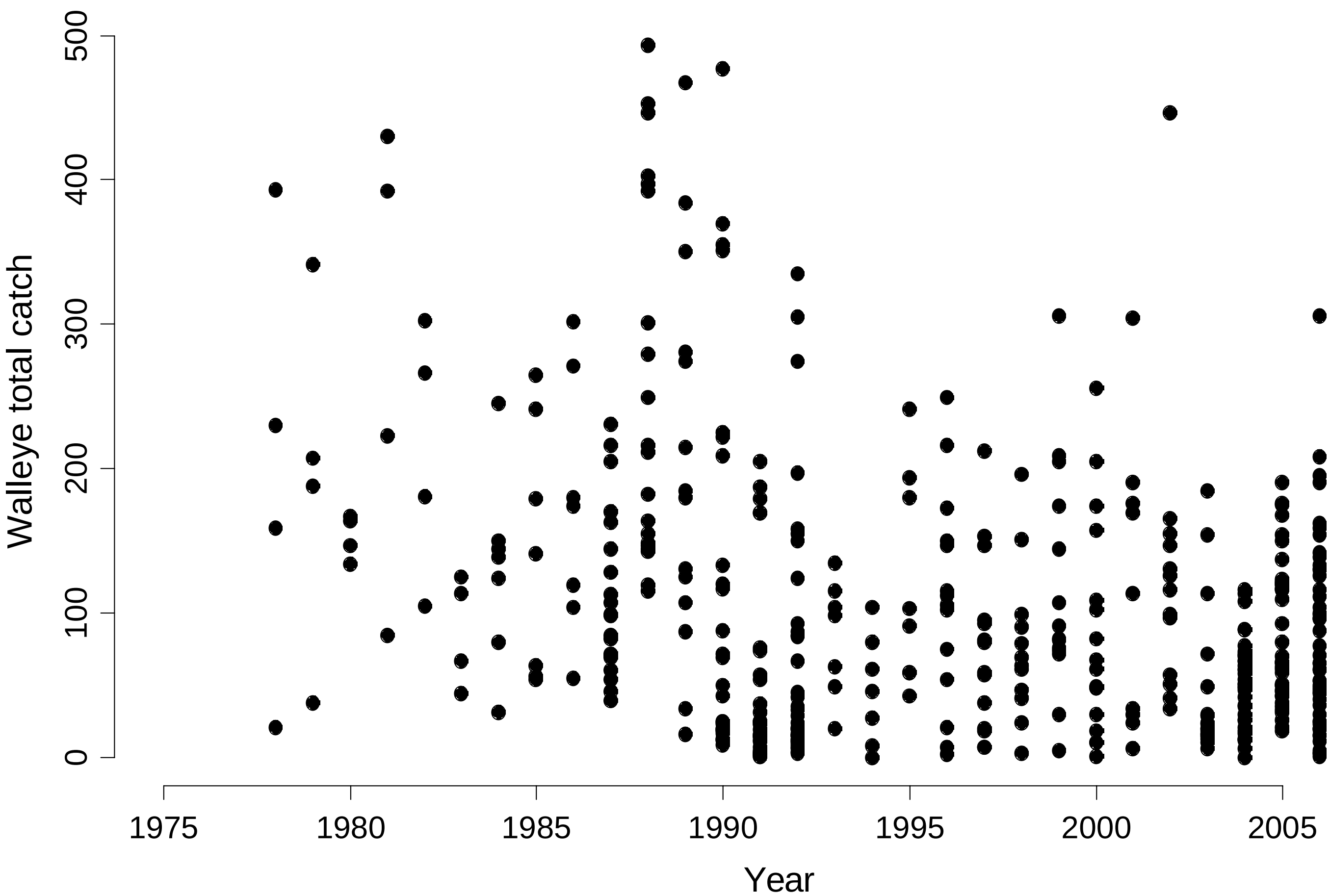

Figure 2. Plot of CPUE for walleye gillnet surveys versus sample year (1978 – 2006). Each dot represents a gillnet sample (net set).

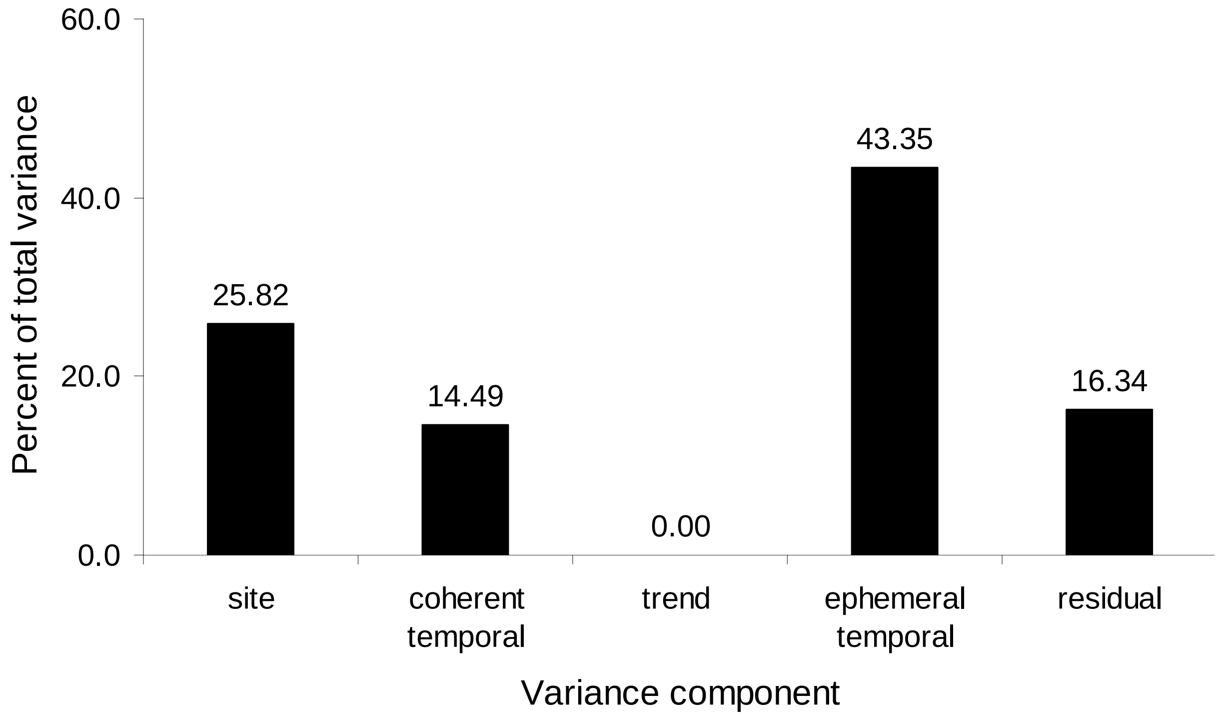

Figure 3. Estimated percent of total variation attributed to site, coherent temporal, ephemeral temporal, trend variation, and residual variance. Estimates are from a mixed model for log (total walleye catch) versus time.

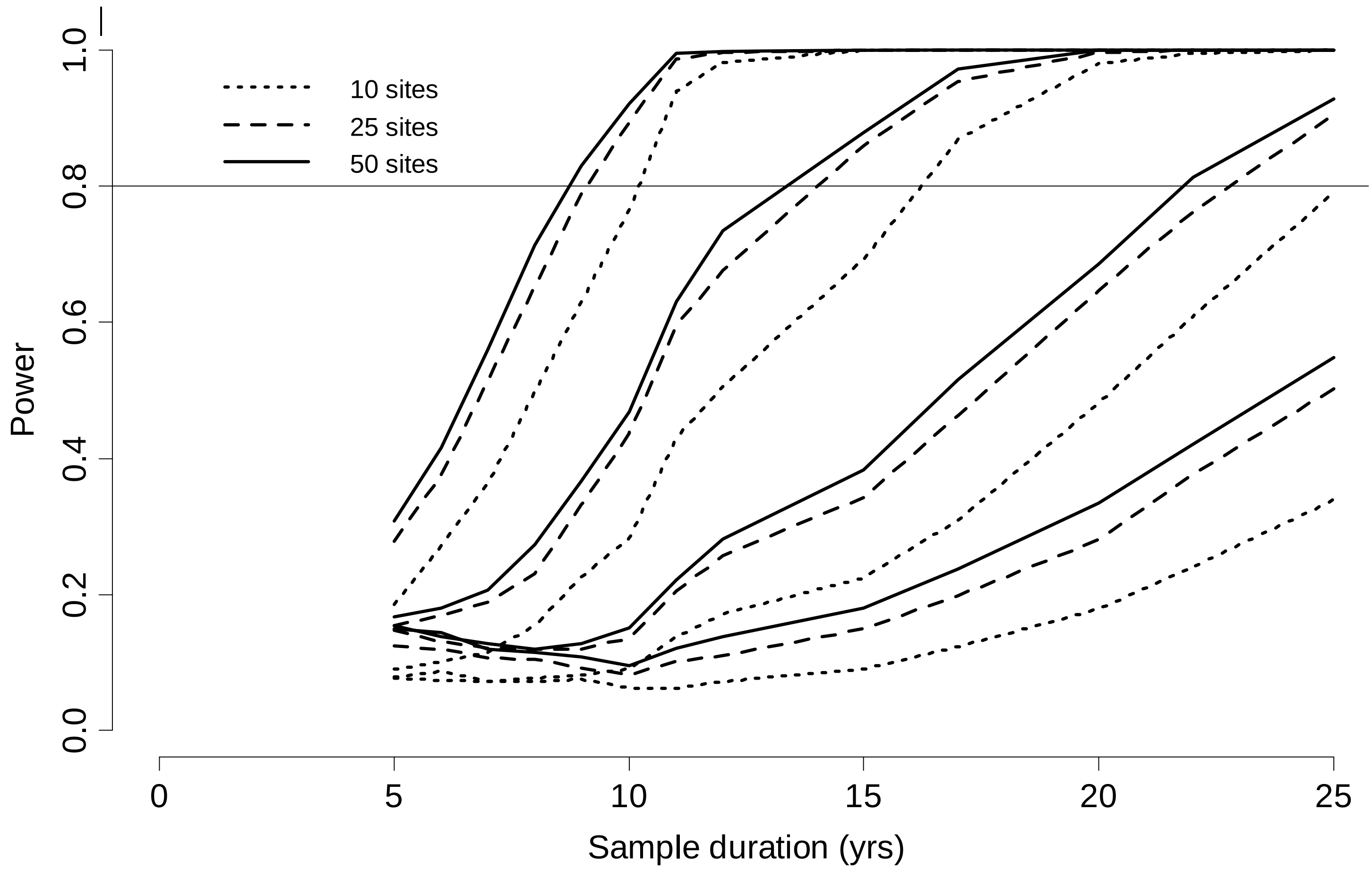

Figure 4. Power curves for detecting temporal trends in gillnet catch per unit effort for walleye in Lake Erie with increasing number of fixed sample sites sampled per year and increasing trend magnitude (bottom set of three curves = average annual percent change of 3%; middle-lower set of three curves = average annual percent change of 5%; middle-upper set of three curves = average annual percent change of 10%; and upper set of three curves = average annual percent change of 20%. This analysis was performed assuming the total population of sites from which to sample was equal to 305.