# Lambda Review Workshop Completion Report

Presented by the Standing Technical Committee to the
Lake Erie Committee
03/07/2007

## Introduction

The Lake Erie Committee (LEC) is an international committee consisting of state and provincial fisheries management agencies which operate with support from the Great Lakes Fishery Commission (GLFC). The LEC consists of senior staff members from the Michigan Department of Natural Resources (MDNR), New York State Department of Environmental Conservation (NYSDEC), Ohio Department of Natural Resources (ODNR), Ontario Ministry of Natural Resources (OMNR), and the Pennsylvania Fish and Boat Commission (PFBC), all agencies which administer their respective fisheries. The Committee's purposes are to:

- Consider issues pertinent to, or referred by, the Great Lakes Fishery Commission;
- Consider issues and problems of common concern to member agencies;
- Develop and coordinate joint programs and research projects; and
- Serve as a forum for state, provincial, tribal, and federal agencies.

One of the primary responsibilities of the LEC and associated sub-committees is to annually estimate percid (walleye and yellow perch) populations in Lake Erie, and establish safe harvest levels. As a function of this responsibility, the LEC partner agencies conduct independent and joint fisheries monitoring programs to better understand fish stock dynamics and for use in fish population abundance models.

In 2001, the Lake Erie Committee (LEC) instituted the Coordinated Percid Management Strategy (CPMS) in response to significant percid population declines (LEC 2004). The CPMS had two objectives: 1) to reverse declines and rebuild percid stocks to achieve broad distribution of benefits throughout the lake; and 2) to improve approaches used to estimate percid abundance and determine sustainable harvest levels. In an effort to address Objective 2 (to continually improve population abundance models for Lake Erie percids), in 2001, the Walleye Task Group (WTG) and Yellow Perch Task Group (YPTG) replaced the use of Catch-at-Age Analysis (CAGEAN) software with Auto-Differentiated Model Builder software (ADMB) as the tool of choice for estimation of abundance and mortality trends in Lake Erie percid populations. ADMB is a modeling package based upon C++ platform that estimates model parameters using optimization via maximum likelihood estimators. One big advantage of ADMB is the software contains automatic differentiation routines that minimize processing time significantly. Other reasons for adopting ADMB software for estimating percid abundance included greater flexibility with data including series length, multiple gear types including survey gear, accommodation of unequal years/ages between series; the absence of a terminal F input requirement, and customized reporting.

As stated above, with the ADMB software, one attractive option is the ability to include auxiliary data series in the model to estimate population size. Most experts on statistical catch-at-age analysis recommend using auxiliary data series, in addition to traditional fishery catch and effort information (Deriso et al. 1985; Quinn and Deriso 1999), particularly for models used to estimate absolute abundance. Auxiliary survey gear inclusion in the modeling framework has been shown to provide less biased estimates of abundance, as opposed to strictly modeling fishery-dependent data. Most agencies have extensive fishery independent surveys that are administered annually to collect relative catch-at-age information on percids, as well as for quantification of population parameters. For example, since 1989 OMNR has annually administered a fall gillnet survey, sampling 54-188 stations in the west and central basins of Lake Erie. This survey is used to estimate relative catch-at-age for walleye and yellow perch in Ontario waters of Lake Erie. For yellow perch, since 1969 ODNR has annually sampled 10-40 stations in the fall using a bottom trawl. This survey is used primarily to estimate relative catch-at-age for yellow perch in Ohio's western basin of Lake Erie.

When modeling catch-at-age data using CAGEAN or ADMB, it is usually necessary to include a data weighting factor ($\lambda$) for each data series included in the model. This data weighting factor is a pre-specified weighting term to govern how strongly the data series should influence the overall fit of the population model. Quinn and Deriso (1999) note that "the choice of $\lambda$ and other weighting terms for auxiliary information is one of the most critical aspects of catch-age analysis." These weighting factors essentially represent the confidence we have in the dataset relative to the true population trends, and allow for each data series to affect the outcome of the population model based upon this confidence.

Traditionally, the WTG and YPTG have calculated catch, effort, and survey catch-per-unit-effort (CPUE) data weighting factors by the variance ratio method as suggested by Deriso et al. (1985) or by an alternative method suggested by Quinn and Deriso (1999). The former method involves pre-specifying the effort lambda ($\lambda_E$) based upon the ratio of variances method (e.g. variance of the set of annual observed log (Catch) divided by the variance of the set of annual observed log (Effort)). Then, for all other data series, the $\lambda$ values were iteratively solved in the model.

One other problem which has never been addressed in the model is the relative importance or weighting between the three data categories in the population model (i.e.; fishery effort, fishery catch and survey catch rates data types). Currently (and historically) a single data source within each data category was assigned a lambda weighting of 1. Consequently, general data types were weighted equally, but with variable weighting within data categories.

However, after further review (LEC 2005; Wright et al. 2005) and discussion within the task groups, external reviewers (Myers and Bence 2002), the LEC and task groups recommended that the method used by the WTG and YPTG to calculate lambda weighting values be re-evaluated. Due to these recommendations and the apparent critical nature of $\lambda$ values with respect to the population models themselves, in 2005, the WTG, YPTG, and Standing Technical Committee (STC) were all assigned charges by the LEC to address the Lambda weighting issue.

In response to this charge, representatives from the WTG, YPTG, and STC arranged for the review of current methods and thinking about possible changes to those methods by the Michigan State University's Quantitative Fishery Center (QFC) (Bence 2006), and based on that review the WTG, YPTG and STC arranged for a workshop directed by the QFC to explore what we are doing when we weight various data sources, and develop a defensible process for weighting each data source. The agenda of the workshop included: 1) discussion of the theoretical meaning of λs; 2) discussion of auxiliary data sources available and examination of precision of these data sources; 3) discussion of combining/separating data sources and an *a priori* process for developing lambdas; 4) an example of estimating lambdas and fitting the model; and 5) discussion of other modeling issues including error structures, and density-dependent catchability.

## Workshop Results

### Process Overview

In initial discussion with the QFC's Dr. Mike Jones and Dr. Jim Bence, the STC and task group modelers agreed a new method for setting λs for the data sets should be developed. We felt that the λ values for fishery catch and survey catch rates should be based upon how well we measure harvest, effort, and abundance; therefore, we used prior information from sampling to guide the λ weighting. It was decided that the data sources should be weighted based upon the observed variability in the data sources themselves. As such, all λ values will be pre-specified prior to population model runs. Effort λs describe the degree to which observed fishing effort is related to fishing mortality. Therefore, although our ability to measure fishery effort would set an upper bound for these λs, other sources of variation (e.g., temporal variation in fishing catchability) may be more important and are not directly observed. Consequently we would probably be guided by the best fit of catch and survey data as we tried alternative relative values of λs for effort versus catch series.

Because the initial intent of the workshop was to develop a process for prescribing λ values to each data source, the QFC/modeling group recommended that the process be developed for one model (the walleye model) and then the same process could be applied for estimating λs in the other catch-at-age models after the workshop (i.e. yellow perch MU 1-4).

The data sources used in the west-central catch-age analysis walleye model are presented in Table 1. It was decided during the workshop that, although additional sources of historic data exist, the model would continue to use the same sources of data as in the most current accepted assessment since the methodologies are understood well and decisions about using historic index gear data that were not used previously requires further investigation. The spatial extent of data sources was also discussed. There was interest in splitting survey and fishery data further spatially as opposed to lumping due to concerns with respect to differences in gear catchability and selectivity between areas (west vs. central basin). The WTG is preparing alternate model configurations with data partitioned into west and central basin sources. Improved model fit would provide the basis for adopting this approach.

In the past, a "standard" data source was identified within each of the three data categories, with other data sources weighted relative to these. The emerging perspective involves weighting data sources from all three categories relative to a single standard. The standard data source was selected based upon expert opinion and the precision of the data series. The relative λ weighing values were calculated as the ratio of the "standard" variance/other variance. At the workshop, there was an attempt to determine the optimum relative commercial effort weighting that would provide the best fit for the commercial catch data. This was done by iteratively changing relative effort weight in the model. Below follows a description of the process used to develop λ values for the workshop, recommendations for changes in the data structure subsequent to discussions at the workshop, and results for the walleye catch-at-age models.

Table 1. Data series in the west-central catch-age analysis walleye model.

| Data Series | Agency |
|---|---|
| Ontario Commercial Effort | OMNR |
| Ohio+Michigan Sport Effort | ODNR, MDNR |
| Ontario Commercial Catch/age | OMNR |
| Ohio+Michigan Sport Catch/age | ODNR, MDNR |
| Ontario Gillnet Survey CPUE (MU1-3) | OMNR |
| Ohio Gillnet Survey CPUE (MU1) | ODNR |
| Michigan Gillnet Survey CPUE (MU1) | MDNR |

The Ontario commercial catch data series was selected as the "standard". The reason for selecting this data series as the standard is the experts had the most confidence in this data series and felt that the variance associated with it was the lowest in all the data series due to the fact that commercial harvest at age estimates are generated from mandatory reporting (or a complete census). Because the Ontario commercial catch data series was selected as the "standard", other catch and survey data lambdas were determined based on the ratio of the "standard variance" to catch or survey variance. The standard data series is the numerator because lambda weightings are inversely proportional to variance. Input fishery effort lambdas were determined in a similar manner, using the commercial effort as the "standard". Effort and catch lambdas were scaled relative to each other by means of a multiplier. The multiplier was changed in an iterative process in order to optimize the model fit to commercial catch data. Although we took some care in defining which data source was the "standard", we emphasize that results are sensitive only to the relative weights and not to the choice of the standard.

The process below describes the method used for estimating λ values:

1. Estimate the annual variability of each catch (fishery and survey data) or effort data series used in the model.
    a. For catch time series (commercial and sport catch and survey CPUE), age-specific Coefficient of Variations (CVs) were estimated as:

$$\frac{(std / \mu)}{\sqrt{n}} \tag{1}$$

Note that here *std* represents the standard deviation on a log scale of deviations between estimated and "true" values. True values for catch are, of course the actual amount caught. True values for survey CPUE are what would be caught if survey CPUE were in fact exactly proportional to true population abundance.

b. An average over ages of the age-specific CVs calculated by equation 1 was used as the estimate of variability for each data source, and are presented in Table 2. Commercial and sport fishery CVs were adjusted based on a priori information that was not accounted for in the initial CV calculations. In particular it was recognized that errors in commercial catch reporting and variation in all extractions probably would lead to a greater CV such that the sport catch CV was assumed to be the same as the commercial catch CV.

c. Catch and survey λ values were estimated based on their relative variance, with the commercial catch used as the standard.

$$\left(CV_{s\tan dard}\right)^2 \Big/ \left(CV_{other}\right)^2 \qquad (2)$$

This was considered reasonable because CV is a very close approximation of the log scale standard deviation for the range of variances we considered.

Table 2. Average of age-specific coefficient of variation values and resulting lambdas by walleye data series.

| Data Series | Agency | CV | Lambda Value |
|---|---|---|---|
| Ontario Commercial Catch at age | OMNR | 10%* | 1.00 |
| Ohio+Michigan Sport Catch at age | ODNR, MDNR | 10% | 1.00 |
| Ontario Gillnet Survey CPUE (MU1-3) | OMNR | 25% | 0.16 |
| Ohio Gillnet Survey CPUE (MU1) | ODNR | 50% | 0.04 |
| Michigan Gillnet Survey CPUE (MU1) | MDNR | 44% | 0.05 |

* Ontario commercial catch CV was originally determined to be ~1%. This value was increased due to additional variation in the data and model expectations.

2. Again as for catch and survey values, for effort data sources, relative λ values were estimated based on the perceived relative variance in gear catchability (q) of the standard data series/variance of the other data series (Eq, 2) In order to calculate the relative sport fishing effort λ, we set the Ontario commercial effort λ=1.0, and based upon expected catchability deviations in sport effort CV=1.75 (sport catchability believed to vary 75% more than commercial); therefore, sport effort λ= $(1.0)^2/(1.75)^2=0.327$.

3. Next, we ran model iterations by scaling all catch and survey lambdas relative to the initial commercial effort lambda (scalars 1 – 2500) (Table 3). The plan was to accept the scalar that produced variation between observed and model predictions of commercial catch-at-age that had a desired CV corresponding to our prior estimate of the CV of commercial catch-at-age based on how well commercial catch is measured. The model was updated

to use the full negative log likelihood in the objective function, as opposed to the concentrated log likelihood that was used in the original model (this did not change parameter estimates but was desired for possible future uses of the model). Initially, we attempted to achieve a CV based on comparison of modeled and observed data of 1% (Figure 1). Later, a value of 10% was considered a more realistic target due to additional variation in the data.

Table 3. Standard deviation estimates and relative lambda values for each data set based on different input scalar values.

| | Scalar Value | | | | | |
|---|---|---|---|---|---|---|
| **Model Output and Input** | **1** | **100** | **1000** | **2500** | **3000*** | **7000** |
| Total f | 25 | 138 | 203 | 228 | 27168 | 257 |
| TSS | 433 | 438 | 439 | 439 | 77589 | 439 |
| Commercial effort Dev SD estimate | 0.30 | 2.94 | 9.30 | 14.70 | 54.60 | 24.60 |
| Sport effort Dev SD estimate | 0.52 | 5.14 | 16.26 | 25.71 | 95.48 | 43.02 |
| Com catch SD estimate | 0.30 | 0.29 | 0.29 | 0.29 | 1.00 | 0.29 |
| Sport catch SD estimate | 0.30 | 0.29 | 0.29 | 0.29 | 1.00 | 0.29 |
| Ontario survey SD estimate | 0.75 | 0.74 | 0.74 | 0.74 | 2.49 | 0.74 |
| Ohio survey SD estimate | 1.50 | 1.47 | 1.47 | 1.47 | 4.98 | 1.47 |
| Michigan survey SD estimate | 1.32 | 1.29 | 1.29 | 1.29 | 4.38 | 1.29 |
| SDLAM | 0.30 | 2.94 | 9.30 | 14.70 | 54.60 | 24.60 |
| Relative Commercial effort lambda | 1.00000 | 0.01000 | 0.00100 | 0.00040 | 0.00033 | 0.00014 |
| Relative Angler effort lambda | 0.32700 | 0.00327 | 0.00033 | 0.00013 | 0.00011 | 0.00005 |
| Relative Commercial catch lambda | 1 | 1 | 1 | 1 | 1 | 1 |
| Relative Angler catch lambda | 1 | 1 | 1 | 1 | 1 | 1 |
| Relative Ontario Survey lambda | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 | 0.16 |
| Relative Ohio Survey lambda | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| Relative Michigan Survey lambda | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |

*The model did not converge when the scalar value was between approximately 3000 – 6000.

Outcomes and recommendations from initial runs:

a. The relative effort λ value was varied across the entire scale range (1-2500) in an attempt to match the modeled catch variability with the observed catch variability.

b. We had difficulty in matching the modeled catch variability (~30%) in the catch deviations with the observed variability (1% or 10%) from the data sources that were used for λ weighting (Figure 2).

c. Our a priori estimates or educated guesses about CVs for catch were relatively small, while the variation implied by differences between model predicted and observed values implied much larger CVs and these implied CVs were relatively constant despite the value of effort lambda parameter. The disconnect between our prior information and outputs indicates a relatively poor model fit (i.e., to a large extent deviations between model predictions and observed catch may reflect real processes that the current model is not flexible enough to emulate, such as fishery selectivity might differ among ages and this may vary among surveys, which the current assessment model does not allow). It is not clear whether effort lambdas can be estimated reliably in the catch-age analysis, given our current model.

**Workshop Recommendations**

Recommendations from the QFC and modeling group included:

1. To try to improve the fit of the model, we should include fishery and survey data separately by basin in the model to account for spatial differences in abundance and age-structure (e.g., break the Ontario gillnet survey data up into two data sources – one for the western basin, one for the central basin) and other causes for differences in catchability and or selectivity.

2. Include selectivity parameters for each of the survey data sources to account for differences in spatial distribution and age distribution across the basins or other potential causes for different apparent selectivity evidenced by the different age compositions for the different surveys. These two proposed changes to the model should improve the fit of the model.

3. Re-run the model with re-estimated λ values for each of the parsed-out data sources (with the new basin-specific fishery and survey data sets) and selectivity parameters, and attempt to "match" the resulting modeled variability with the observed variability, by again varying lambdas

4. More work and thinking is needed with regard to catchability. While there is concern and evidence that fishery catchability is density dependent, to large measure this may already be captured by the use of catchability blocks. Given the use of catchability blocks, the remaining question would be whether there are substantial changes in catchability in response to changing density within blocks. Another alternative would be to model fishery catchability as a random walk. Even if such an approach is not adopted the results might provide a firmer foundation for choosing catchability blocks. Abrupt changes in the estimated random walk might suggest that using blocks is a better approach. Another alternative would be to model density dependent catchability using a power function, either with or without blocks (but not in conjunction with a random walk). However, this probably only offers modest benefits and then only if most of the temporal variability in catchability (overall or within blocks if they are used) is due to density dependence. While survey catchability is typically assumed to be relatively stable in comparison with fisheries, the assessment group is urged to continue to evaluate whether survey catchability might be changing. If it seems to be, this could be incorporated into future assessment models.

Other background material discussed at the workshop included:

- The question of "how much does fishery effort influence fishing mortality" may vary between fisheries. i.e: does anyone think that the amount of gill net fished does not influence fishing mortality? Is it possible that effort data is informative?
- Gill net effort duration is not considered in reported gill net effort.
- Aspects of harvest estimates and reporting addressed by the Blue Ribbon Panel for review of Procedures Used to Estimate Percid Harvest in Lake Erie.
- Variablity of sampling intensity within surveys (addressed somewhat by CV calculations)
- In addition to spatial considerations, changing skill levels and technologies of fishers, effects of water clarity was discussed in the context of changing catchability and effort lambdas.
- The current approach assumes that log scale variance is constant over all ages, and a prior value for this constant variance was estimated by averaging age-specific estimates. It was noted that the data analysis seems to suggest that the variance is not constant. While this is a second order issue, it was recommended that alternative model structures be considered that allow for unequal variances (e.g., use of multinomial likelihood for age compositions rather than the CAGEAN-like approach of using a lognormal distribution).

**Progress and Future Direction**

- The process for developing lambda weightings appears to be sound.
- The criteria for assessing overall model fit with respect to the variability in the data inputs appears logical
- The model output variance for commercial catches appears insensitive to varying lambda weighting values for the other input data series.
- The criteria used to assess lambda weighting will need to be re-evaluated
    - Varying lambda values did not improve model fit
    - Criteria for fit were never achieved (i.e. modeled variances never matched input variance for commercial catch data).
- The modeling subgroup will further explore these recommendations in '07 by
    - Exploring split versus combined fishery dependent/independent surveys
        - If the different series indicate consistent difference in age composition or cpe within a year the surveys will be split
        - If the different series indicate similarity with respect to age composition or cpe with a year the surveys will be combined
    - Sensitivity analyses of old model and new model to changes in lambda values
    - Evaluation of additional criteria for improving model fit
- The STC and task groups will work with the above recommendations to refine model results and use them to learn how to best implement the model. In the end, these recommendations may be excluded, included with modifications, or adopted as is to affect the best model implementation possible.