Chapter A6: Uncertainty

Introduction

This chapter discusses sources of uncertainty in EPA's impingement and entrainment (I&E) analyses, and presents the preliminary results of an uncertainty analysis of the yield model used by EPA to estimate the benefits of reducing I&E of commercial and recreational fishery species. Section A6-1 discusses major uncertainties in EPA's I&E assessments, Section A6-2 briefly describes Monte Carlo analysis as a tool for quantifying uncertainty, Section A6-3 provides preliminary results of an uncertainty analysis by EPA of winter flounder yield estimates, and Section A6-4 discusses results of the uncertainty analysis.

Chapter Contents

A6-1 Types of Uncertainty

Despite following sound scientific practice throughout, it was impossible to avoid numerous sources of uncertainty that may cause EPA's I&E estimates in the regional analysis to be imprecise or to carry potential statistical bias. Uncertainty of this nature is not unique to EPA's I&E analysis.

Uncertainty may be classified into two general types (Finkel, 1990). One type, referred to as structural uncertainty, reflects the limits of the conceptual formulation of a model and relationships among model parameters. The other general type is parameter uncertainty, which flows from uncertainty about any of the specific numeric values of model parameters. The following discussion considers these two types of uncertainty in relation to EPA's I&E analysis.

A6-1.1 Structural Uncertainty

The models used by EPA to evaluate I&E simplify a very complex process. The degree of simplification is substantial but necessary because of the limited availability of empirical data. Table A6-1 provides examples of some potentially important considerations that are not captured by the models used. EPA believes that these structural uncertainties will generally lead to inaccuracies, rather than imprecision, in the final results.

Table A6-1: Uncertainties Associated with Model Structure

Type	General Treatment in Model	Specific Treatment in Model
Generally simple structure	Species lost to I&E treated independently	Fish species grouped into two categories: harvested or not harvested (forage for harvested species)
Biological submodels	No dynamic elements	Life history parameters constant (i.e., growth and survival did not vary through time); growth and survival rates did not change in response to possible compensatory effects
Economic submodels	No dynamic elements	Ratio of direct to indirect benefits was static through time; market values of harvested species were inelastic (i.e., were fixed and thus not responsive to market changes that may occur due to increased supply when yield is higher)
	Fish stock	Landings of commercial and recreational fish associated with I&E losses assumed to be within the State where facility is located
	Angler experience	I&E losses at a facility assumed to be relevant to angler experience (or perception) and Random Utility Model (RUM) models of sport fishery economics.

A6-1.2 Parameter Uncertainty

Uncertainty about the numeric values of model parameters arises for two general reasons. The first source of parameter uncertainty is imperfect precision and accuracy of impingement and entrainment data reported by facilities and growth and mortality rates obtained from the scientific literature. This results from unavoidable sampling and measurement errors. The second major source of parameter uncertainty is the applicability of parameter estimates obtained from I&E or life history studies conducted at other locations or under different conditions.

Table A6-2 presents some examples of parameter uncertainty. In all of these cases, increasing uncertainty about specific parameters implies increasing uncertainty about EPA's point estimates of I&E losses. The point estimates are biased only insofar as the input parameters are biased in aggregate (i.e., inaccuracies in multiple parameter values that are above the "actual" values but below the "actual" values in other cases may tend to counteract). In this context, EPA believes that parameter uncertainty will generally lead to imprecision, rather than inaccuracies, in the final results.

Table A6-2: Parameters Included in EPA's I&E Analysis that Are Subject to Uncertainty

Type	Factors	Examples of Uncertainties in Model
I&E monitoring /loss rate estimates	Sampling regimes	Sampling regimes subject to numerous plant-specific details; no established guidelines or performance standards for how to design and conduct sampling regimes
	Extrapolation assumptions	Extrapolation of monitoring data to annual I&E rates requires numerous assumptions regarding diurnal/seasonal/annual cycles in fish presence and vulnerability and various technical factors (e.g., net collection efficiency; hydrological factors affecting I&E rates); no established guidelines or consistency in sampling regimes
	Species selection	Criteria for selection of species to evaluate not well-defined or uniform across facilities
	Sensitivity of fish to I&E	Through-plant entrainment mortality assumed by EPA to be 100 percent; some back-calculations required in cases where facilities had reported entrainment rates that assumed <100 percent mortality. Impingement survival included if presented in facility documents.
Biological/life history	Natural mortality rates	Natural mortality rates (M) difficult to estimate; model results highly sensitive to M
	Growth rates	Simple exponential growth rates or simple size-at-age parameters used
	Geographic considerations	Migration patterns; I&E occurring during spawning runs or larval out-migration; location of harvestable adults; intermingling with other stocks
	Forage valuation	Harvested species assumed to be food limited; trophic transfer efficiency to harvested species estimated by EPA based on general models; no consideration of trophic transfer to species not impinged and entrained.
Stock characteristics	Fishery yield	For harvest species, used only one species-specific value for fishing mortality rate (F) for all stages subject to harvest; used stage-specific constants for fraction vulnerable to fishery
	Harvest behavior	No assumed dynamics among harvesters to alter fishing rates or preferences in response to changes in stock size; recreational access assumed constant (no changes in angler preferences or effort)
	Stock interactions	I&E losses assumed to be part of reported fishery yield rates on a statewide basis; no consideration of possible substock harvest rates or interactions
	Compensatory growth	None
	Compensatory mortality	None
Ecological system	Fish community	Long-term trends in fish community composition or abundance not considered (general food webs assumed to be static); used constant value for trophic transfer efficiency; specific trophic interactions not considered. Trophic transfer to organisms not impinged and entrained is not considered.
	Spawning dynamics	Sampled years assumed to be typical with respect to choice of spawning areas and timing of migrations that could affect vulnerability to I&E (e.g., presence of larvae in vicinity of intake structure)
	Hydrology	Sampled years assumed to be typical with respect to flow regimes and tidal cycles that could affect vulnerability to I&E (e.g., presence of larvae in vicinity of CWIS)
	Meteorology	Sampled years assumed to be typical with respect to vulnerability to I&E (e.g., presence of larvae in vicinity of intake structure)

A6-1.3 Uncertainties Related to Engineering

EPA's evaluation of I&E was also affected by uncertainty about the engineering and operating characteristics of the study facilities. It is unlikely that plant operating characteristics (e.g., seasonal, diurnal, or intermittent changes in intake water flow rates) were constant throughout any particular year, which therefore introduces the possibility of bias in the loss rates reported by the facilities. EPA assumed that the facilities' loss estimates were provided in good faith and did not include any intentional biases, omissions, or other kinds of misrepresentations.

A6-2 Monte Carlo Analysis as a Tool for Quantifying Uncertainty

Stochastic simulation is among a class of statistical procedures commonly known as Monte Carlo modeling methods. Monte Carlo methods allow investigators to quantify uncertainty in model results based on knowledge or assumptions about the amount of uncertainty in each of the various input parameters. The Monte Carlo approach also allows investigators to conduct sensitivity analyses to elucidate the relative contribution of the uncertainty in each input parameter to overall uncertainty. Monte Carlo methods are particularly useful for assessing models where analytic (i.e., purely mathematical) methods are cumbersome or otherwise unsuitable. A thorough introduction to the statistical reasoning that underlies Monte Carlo methods, and their application in risk assessment frameworks, is provided in an EPA document "Guiding Principles for Monte Carlo Analysis" (U.S. EPA, 1997).

The characteristic feature of Monte Carlo methods is the generation of artificial variance through the use of pseudorandom numbers. The solution to the model of interest is recalculated many times, each time adding perturbations to the values of the model parameters. The types of perturbations are selected to reflect the actual uncertainty in knowledge of those parameters. Recalculations are conducted thousands of times, and the variation in the resulting solution is assessed and interpreted as an indicator of the aggregate uncertainty in the basic result.

A6-3 EPA's Uncertainty Analysis of Yield Estimates

A6-3.1 Overview of Analysis

As described in detail in Chapter A5 of this report, EPA estimated foregone yield using the Thompson-Bell equilibrium yield model (Ricker, 1975). The Thompson-Bell model is based on the same general principles that are used to estimate the expected yield in any harvested fish population (Hilborn and Walters, 1992; Quinn and Deriso, 1999). The general procedure involves multiplying age-specific weights by age-specific harvest rates to calculate an age-specific expected yield (in pounds). The lifetime expected yield for a cohort of fish is then the sum of all age-specific expected yields.

(Equation 1)

$$Y_k = \sum_j \sum_a L_{jk} S_{ja} W_a (F_a / Z_a)\left(1 - e^{-Z_a}\right)$$

where:

Y_k = foregone yield (pounds) due to I&E losses in year k
L_{jk} = losses of individual fish of stage j in the year k
S_{ja} = cumulative survival fraction from stage j to age a
W_a = average weight (pounds) of fish at age a
F_a = instantaneous annual fishing mortality rate for fish of age a
Z_a = instantaneous annual total mortality rate for fish of age a

Quantifying the variance in yield estimates resulting from uncertainty in the numeric values of L, S, W, F, and Z assists in the interpretation of results, gives a sense of the precision in yield estimates, provides insight into the sensitivity of predictions to particular parameter values, and indicates the contribution of particular parameters to overall uncertainty.

EPA evaluated uncertainty in yield estimates for winter flounder using I&E data for a facility located on a North Atlantic estuary. The I&E loss records and winter flounder life history parameters that were used are provided in the Phase II proposal docket as DCN # 4-2037.

EPA developed a custom program written in the S language to conduct the Monte Carlo analysis. Wherever possible, the simulation tool re-used the same code that was used to calculate yield for the original assessment. Graphical displays were used to confirm the behavior of random number generation and to examine results.

Selection of input distributions for parameters of interest are a key element of any Monte Carlo analysis. In the winter flounder test case, the input distributions were uniform distributions with a range defined as the initial, best estimate of the parameter +/- 15%. A uniform distribution was selected because of its simplicity and the 15% range was selected because this magnitude of variance is considered plausible.

EPA investigated sensitivity of the model to variations in parameters by grouping the parameters into five classes:

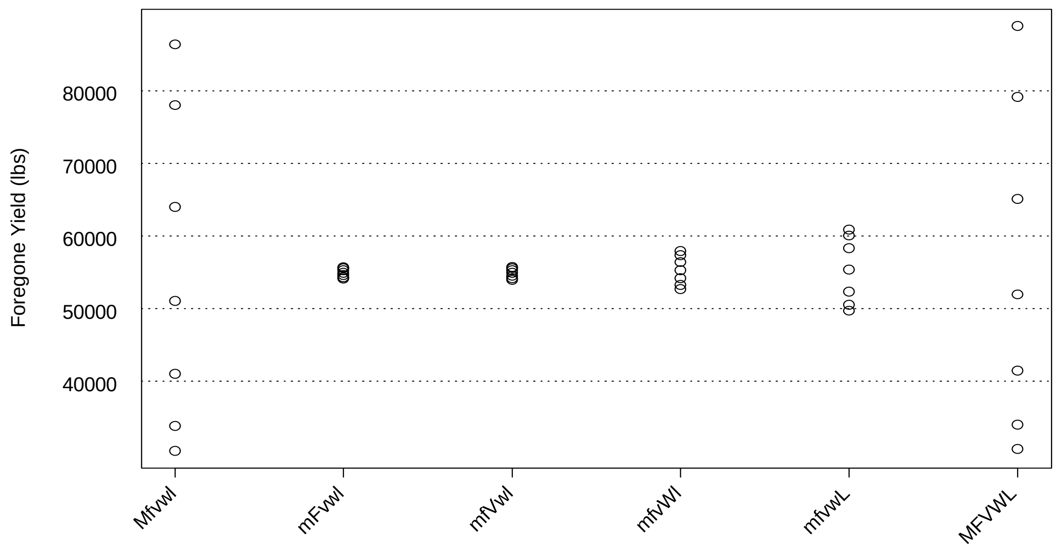
- natural mortality (M) at all life stages,
- fishing mortality (F) at all life stages,
- fraction vulnerable to fishing (V) at all life stages (i.e., age of recruitment to the fishery),
- weight at age (W), and
- the reported I&E loss rates (L).

The analysis consisted of repeating runs (n=10,000 in each run) of the model wherein each of the groups of parameters was either held constant at their best estimates or were varied stochastically according to the defined input distributions. The relative importance of these groups of parameters was assessed by comparing the relative amount of variation between each set of runs. Model sensitivity to individual parameters has not been examined.

A6-3.2 Preliminary Results

For entrainment losses, the analysis indicated that the yield model is most sensitive to uncertainty in natural mortality rates, followed by uncertainty in the I&E loss rates themselves (Figure A6-1). Age specific weights were the third most important group, followed by fishing mortality and age at recruitment, which were relatively insignificant sources of uncertainty.

Figure A6-1: Results of Preliminary Parameter Sensitivity Analysis of Estimates of Foregone Yield (pounds) of Winter Flounder Due to Entrainment by a Power Plant Located in a North Atlantic Estuary



Data points are plotted at the 5th percentile, 10th percentile, 25th percentile, median, 75th percentile, 90th percentile, and 95th percentile of 10,000 independent estimates of foregone yield within each parameter set. Groups are distinguished by uppercase letters designating which types of parameters were treated stochastically in the simulation and lowercase letters for types of parameters fixed at their best estimates. M = natural mortality rates; F = fishing mortality rates; V = age of recruitment to the fishery; W = weight at age; L = entrainment loss rates.

A6-4 CONCLUSIONS

This chapter includes a general discussion of uncertainty and describes a general approach that was tested by EPA as a way to quantify uncertainty associated with the yield model described in Chapter A5. Preliminary results of the uncertainty analysis suggest that uncertainty about natural mortality rates is a significant contributor to aggregate uncertainty in yield estimates. Unfortunately, as noted in a review article by Vetter (1987), “True rates of natural mortality, and their variability, are poorly known for even the great stocks of commercial fish in temperate regions that have been subject to continuous exploitation for decades” (Vetter, 1987, p. 39). As a result, the uncertainty in mortality parameters cannot be overcome. As Vetter (1987) noted, this is a difficulty shared by all models of fish stock dynamics. Nonetheless, through consultation with local fish biologists as well as the scientific literature, EPA expended considerable effort to identify reasonable mortality rates and other life history information for use in its yield analyses. These parameter values and data sources are presented in Appendix 1 of each regional study (Parts B-H of this report).