

## CHAPTER FOURTEEN

### MODELING

*Models should be as simple as possible, but not more so.*

*Albert Einstein, quoted in May (2004)*

*I would not give a fig for the simplicity on this side of complexity, but I would give my life for the simplicity on the other side of complexity.*

*Attributed to Oliver Wendell Homes*

*Because they offer insights into the population dynamics of salmon, it is inevitable that models such as ours will be used to assist in management or in the determination of policy questions. This will happen despite the protestations of statisticians who at times are forced to explain that the available data and scientific understanding of the issues under discussion might not be an adequate basis for answering such questions. A far more satisfactory situation is when the model is used in conjunction with a programme of research into the mechanisms underlying the empirical relations in the field.*

*Terry Speed (1993:291)*

In this section I review selected aspects of modeling that seem particularly relevant for modeling salmon in the Central Valley, and offer some recommendations. I then briefly discuss instream flow models. It is far from a comprehensive review of biological modeling,<sup>1</sup> which is a huge topic, and many important topics are not treated at all. Other models are discussed in other chapters. Generally, I take the view that the most important uses of modeling in salmon management are to provide clear summaries of the empirical relations among variables, as exemplified by the various models in Elliott (1994), and to promote clear thinking. As stated by Walters (1986:45): “The value of modeling in fields like biology has not been to make precise predictions, but rather to provide clear caricatures of nature against which to test and expand experiences.”

Proponents of adaptive management sometimes observe that all management is experimental, whether it is intended to be or not, and that much can be gained by accepting this and treating management explicitly as experimental (Hennessey 1994). Similarly, all salmon management is based on some kind of model, even if only a qualitative conceptual model, and clarity in thinking can be gained by making this modeling explicit. For whatever reason, however, modeling and data analysis have not become parts of the culture of salmon management in the Central Valley in that same way that they have in fisheries management and science elsewhere. This is unfortunate, but it offers those concerned with managing Central Valley salmon the opportunity to learn from the experience of others.

---

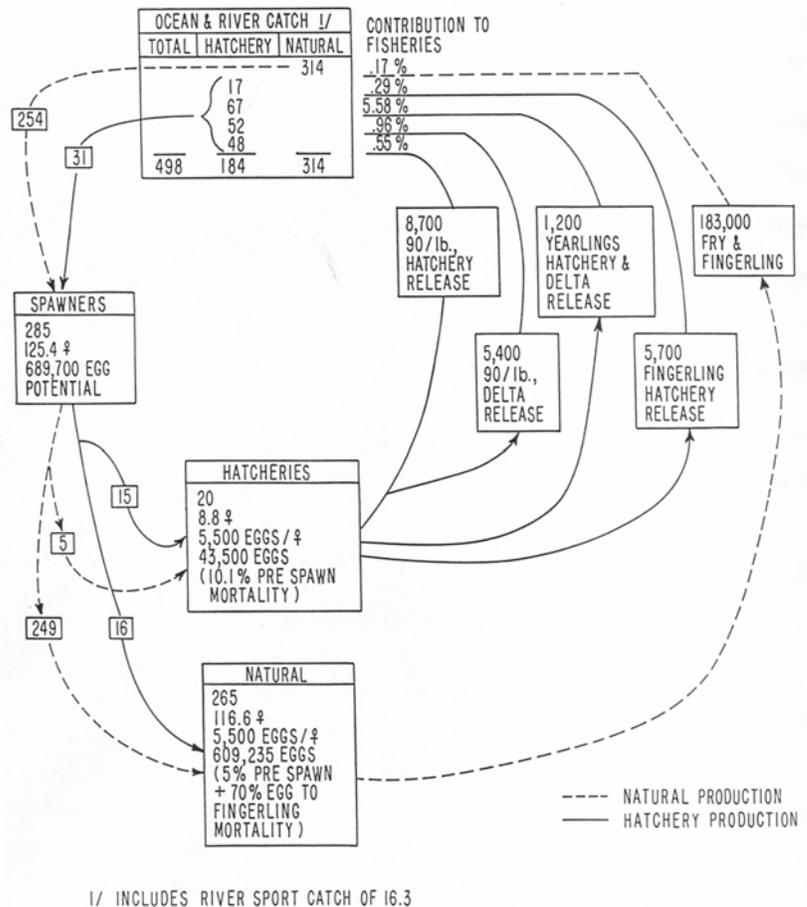
<sup>1</sup> For a good brief discussion of biological modeling generally, see Ch. 2 in Hilborn and Mangel (1997).

The term “modeling” describes a range of activities that can be categorized in various ways. For example, Hilborn and Mangel (1997), with an emphasis on ecological modeling, distinguish deterministic and stochastic models; statistical and scientific models; static and dynamic models; quantitative and qualitative models; and models used for understanding, prediction, and decision. Other sources use similar but not identical categories. The main dichotomy used in this chapter distinguishes simulation models, for which parameter values are taken from the literature or other sources or simply made up, and estimation models, for which parameter values are estimated directly by fitting the model to the data available on the subject at hand. On close inspection his distinction turns out to be somewhat blurred (Schnute 2003), but it emphasizes an important point regarding the use of models for prediction, as explained below.

### Simulation models

A simple but often effective use of models is to explore the consistency of existing information or current assumptions. For example, Hallock (undated) developed a simple conceptual model for Sacramento River Chinook using averaged data or estimates for 1971-1975 (Figure 14-1). The model implied that hatchery Chinook were much more vulnerable to the ocean fishery than naturally produced Chinook. Hallock thought this unlikely, and instead inferred that CDFG was seriously underestimating the proportion of naturally spawning fish that came from hatcheries. Similarly, a quantified conceptual model of a proposed restoration project might be used to assess the plausible size of the effect of the project on the target population.

Figure 14-1. Quantified conceptual model of the hatchery and naturally reproducing components of Sacramento River Chinook. Numbers are thousands of fish. Copied from Hallock (undated).



Adding rules for stepping through time and coding the conceptual model in some computer language turns the conceptual model into a simulation model. The model could also be elaborated, so that, for example, the runs in different Sacramento River tributaries might be treated separately, the juvenile life stage might be divided into phases that are treated separately, juvenile survival might be made to vary as a function of flow, pumping in the Delta, etc. The result is a complex population model. If the population is further divided into individuals or classes of individuals with different attributes that are tracked through time, the result is an individual-based model (IBM).

Several complex population or individual-based models have been or are being developed for Central Valley Chinook: CPOP (Kimmerer et al. 1989); EACH (EA 1991f); SALMOD (Bartholow et al. 1993); ORCM (Jager et al. 1997) SRCSIBM (Kimmerer and Jones & Stokes 1999; and WRIMF (Cramer et al. 2003). The schematic for juvenile Chinook in freshwater from SRCSIBM illustrates the detail in which some of these models describe fish behavior (Figure 14-2). It should be clear from other chapters in the review that many of these details are not well understood, as is recognized by the developers of the models. For example, Kimmerer and Jones & Stokes (1999:17) note that “Rules for the occupancy of rearing habitat are not well known.”

Not all simulation models focus on the salmon life cycle. For example, the simple metapopulation model developed by Levins (1969) has been very influential in ecology (Hanski and Gilpin 1997) and has also been applied to salmon. The basic conceptual model assumes a number of equally suitable patches of habitat, some of which are occupied by populations that have the same probability of extinction, with the unoccupied patches all having the same probability of colonization. The model is characterized by the rate  $m$  at which vacant patches are colonized and the rate  $e$  at which populations in patches go extinct. If the rate of colonization is proportional to both the number of occupied patches and the number of empty patches, and the rate of extinction is proportional to the number of occupied patches, then the rate of change in the proportion  $p$  of occupied patches of habitat is  $dp/dt = mp(1-p) - ep$ .

Even though this seems obvious once it is stated, Hanski and Gilpin (1991) argue that the model was an important advance in ecological thinking because it focused attention on the dynamics of a group of populations rather than individual populations, and because it clarified the importance for those dynamics of the proportion of suitable habitats that are occupied. It also emphasized that suitable patches may be temporarily unoccupied, so that unoccupied habitats are not necessarily unsuitable. Again, the conceptual model can be turned into a simulation model by developing rules for stepping through time. The Levins model is clearly general and unrealistic, but extensions of the model that include spatial structure and variation in habitat size and colonization and extinction rates have been developed (Hanski 1997). These need not be complex. Cooper and Mangel (1999) used a simple metapopulation model with source and sink populations to show that abundance data could mislead managers if the relations between source and sink populations were not recognized and taken into account. Young (1999) used another simple metapopulation model to consider the circumstances in which it makes sense to increase the colonization rate of threatened salmon by rearing eggs taken from fish occupying one stream

and planting them in nearby and similar unoccupied streams. It is worth noting that while Levins's model may seem abstract and academic, it was developed in an article on insect pest control.

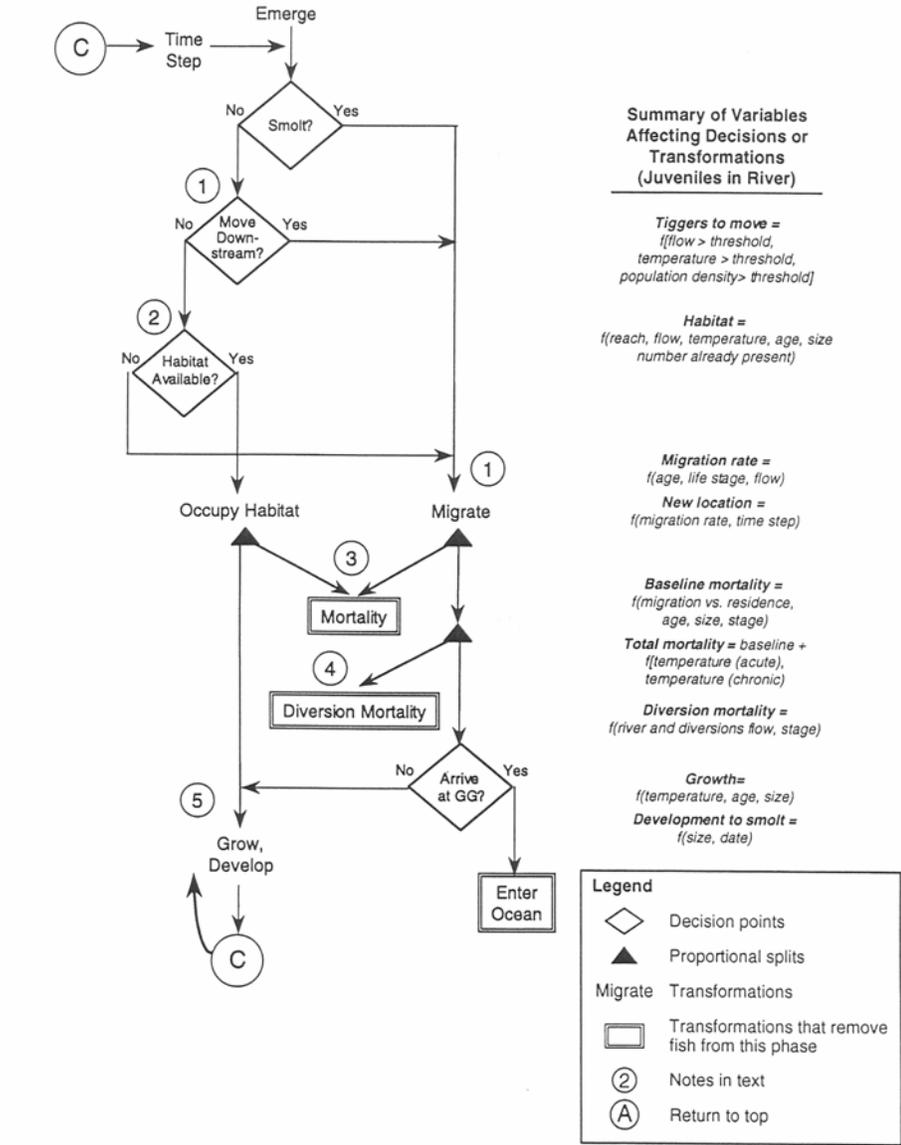


Figure 6  
Flow Diagram for Juveniles in River

Figure 14-2. Schematic for the simulation of the juvenile life stage in SRCSIBM, copied from Kimmerer and Jones & Stokes (1999).

Essentially, simulations are thought experiments (Starfield 1997; Schnute 2003), in which parameter estimates may be more or less well founded in data. For example, Hard (2004) used data from a breeding experiment with ocean-type hatchery Chinook to estimate parameters for simulations of the effects of size-selective harvest over five generations on size at age (and found generally a modest response). For other purposes, such as simulating monitoring programs, it may be enough that the parameter values be plausible. In either case, the basic objective is to clarify thinking, or to make qualitative predictions, rather than quantitative predictions.

### **Estimation models**

With estimation models, as the term is used here, parameter values are obtained by fitting the model to data. Estimation models are essentially statistical, although the terms of the models may be intended to have biological meaning. Stock-recruitment models are usually used as estimation models, and provide a useful example of this kind of model. They have also been important in salmon management, although not in the Central Valley, and deserve attention on that account as well.

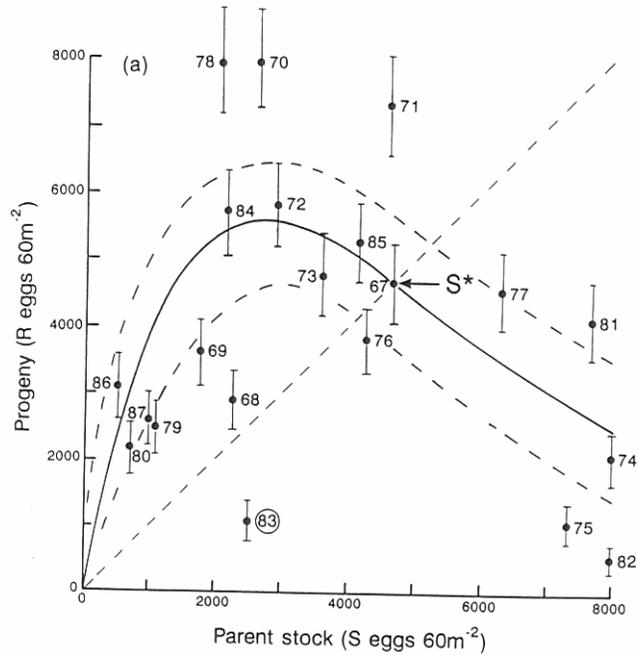
Stock-recruitment modeling relates the abundance of progeny to the abundance of parents, and involves the concepts of density-dependent and density-independent mortality, discussed in Ch. 1. A number of models, briefly reviewed by Elliott (1994), has been proposed to describe this relationship. Applied to fisheries, these are called stock-recruitment relationships, where the parental generation is the stock ( $S$ ) and the filial generation is the recruits ( $R$ ). These can be written in a general form as  $R = aSf(S)$ , where the parameter  $a$  describes the rate of increase in the absence of density-dependent mortality (say when the population is small), and the function  $f(S)$ , which varies from model to model, relates survival to population density (Elliott 1994).

The stock-recruitment model applied most commonly to salmon probably is the Ricker model, which can be written as  $R = aSe^{-bS}$ . This specifies a dome-shaped relationship in which  $R$  peaks when  $S = 1/b$ , and is equal to  $S$  when  $S = \ln(a)/b$  (Figure 14-3). Thus, when the model is fit to data, the value of  $b$  depends on the size of the population, and values of  $b$  for different populations must be adjusted to a common equilibrium size to allow comparisons. Hankin and Healey (1986:1751) suggested that “pessimistic, reasonable(?) and optimistic” values of  $a$  are 5, 10, and 15. Corresponding values for  $b$  for early maturing populations such as those in the Central Valley are 0.00183, 0.00292, and 0.00357, given a few other assumptions and adjusted to give equilibrium populations of 1,000. For Central Valley rivers, Reisenbichler (1986) found values of  $a$  generally in the neighborhood of 10, and Rein (1993) estimated  $a$  for San Joaquin River Chinook at 5.01 or 6.91, depending upon the assumptions made about errors in escapement estimates.

Stock-recruitment models traditionally are applied to salmon as part of harvest management, to estimate the number of fish that should be allowed to escape the fishery and reproduce, but they can be applied to non-exploited populations as well. Elliott (1985; 1994) found that the Ricker model gave the best fit of six models to his data on the population of anadromous brown trout in a section of a small stream in England (Figure 14-3). In this case, Elliott’s interest was in

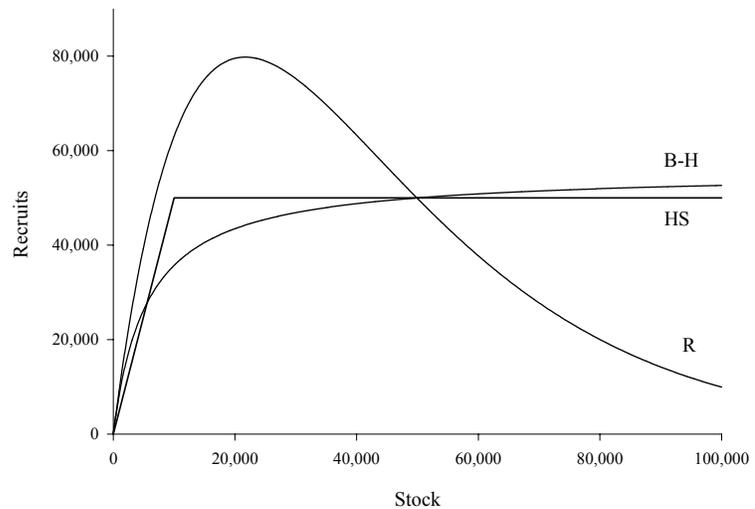
understanding the nature of the density-dependent processes in the study population. Although there are now better methods for comparing models than the ones used by Elliott (1985), his basic approach was sound. Hilborn and Mangel (1997) emphasize the importance of evaluating models against other models, rather than against some abstract criterion of “validation” or “verification.” There are strong arguments that these terms cannot usefully be applied to models (Oreskes et al. 1994; Oreskes and Belitz 2001), as suggested by the aphorism “All models are wrong, but some are useful” (Box 1979).

Figure 14-3. The Ricker model fit to data on brown trout in a section of a small stream in England. Dashed lines around the fit show 95% confidence intervals. Error bars around points show two standard errors; numbers give brood years. Numbers of fish in both generations are given in eggs per area, based on measured lengths of spawners and a length fecundity relationship. The outlier year 1983 was a drought year, and was not used in fitting the model. S\* is the equilibrium population. Reproduced from Elliott (1994), courtesy of Oxford University Press.



The Beverton-Holt model,  $R = aS/(1 + bS)$ , which gives an asymptotic spawner-recruit relationship (Figure 14-4), is also frequently applied to salmon. The Ricker and Beverton-Holt models predict very different consequences of large numbers of spawners, but both assume that, below an equilibrium level, survival increases as population decreases, even when populations are very low. This does not matter much for management of healthy stocks, but can lead to dubiously optimistic expectations when the models are applied to very small populations (Myers et al. 1999). This prompted development of another model, the “hockey stick model,” in which survival is constant up to some level of escapement, above which recruitment is constant (Barrowman and Myers 2000). Since a piecewise linear model seems implausible and introduces numerical problems in fitting the model to data, Barrowman and Myers (2000) also describe generalizations of the hockey-stick, in which the sharp bend is rounded off. In another approach, Munch et al. (2005) apply a Bayesian nonparametric approach to estimating the stock-recruitment relationship. For purposes of this discussion, however, the main point is that the parameter values for these models usually are estimated by fitting the model to data directly from the population of interest.

Figure 14-4. Comparison of Ricker and Beverton-Holt models, with  $a = 10$  and  $b$  set to give an equilibrium population of 50,000, and a hockey stick model.

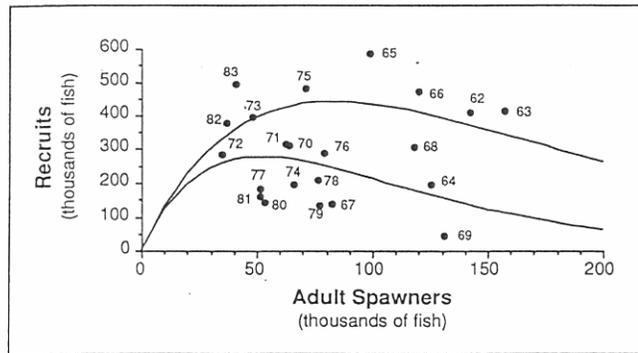


Although the parameters of the Ricker model are reasonably well defined with Elliott's data, there is a great deal of scatter among the data for individual years. Thus, the model does not generate good predictions of the population in a particular future year, based on previous escapement. This has implications for proper use of the model. It would be one thing to set a target escapement based on the model fit to these data (assuming that this were a harvested population), but it would be quite another thing to set a target harvest for a particular year by subtracting the desired future escapement from the predicted recruitment, even though predicting recruitment from escapement is what the model seems to do. Put differently, the model estimates the average escapement for a given stock, and it is necessary to know something about the variability in recruitment to understand how the model should or should not be used. Proper use of models requires an understanding of what is being modeled, as well as of the strengths and weaknesses of the model itself.

In Elliott's data, there is a great deal of scatter in the number of recruits over the range of stock values that is predicted to produce many recruits ( $> \sim 4,500$ ). Hence, data on recruits in this range are not very informative, and Elliott was able to get a good fit because he had data on recruits from both small and large values of stock. The need for such data for estimating parameters well creates a dilemma for fishery managers. If the fishery is successfully managed to allow approximately the desired escapement, as determined from an initial estimate of the parameters of the model, then subsequent recruitment will be of little use in refining the parameter estimates. The idea of adaptive management was developed in the context of estimating the value for fisheries management of data on recruits from high or low values of stock (Walters and Hilborn 1976); that is, whether the fishery should forego some harvest in current years in order that future management could be based on better parameter estimates. This point generalizes to other kinds of adaptive (experimental) management: unless the experiment covers a broad range of conditions, the response will be obscured by normal variation.

Stock-recruitment models have uses beyond managing harvest. For example, Reisenbichler (1986) compared the parameters of the Ricker model fit to escapement data from California streams for different periods of time and found statistically significant changes, suggesting that management of the streams was making them less productive. Kope and Botsford (1990) found a similar change (Figure 14-5), probably resulting from the installation of the Red Bluff Diversion Dam. Speed (1993) developed a state-space model incorporating stock-recruitment models to analyze abundance data for Chinook in the San Joaquin River, and found (p. 290) “that there is a discernable effect of spring flow (or related variables) on the number of recruits over and above that resulting from the size of the parent stock.” Stock-recruitment models can also be used to generate plausible synthetic population data in simulations to evaluate monitoring programs (e.g., Korman and Higgins 1997; Williams 1999). In such cases the particular values used for the parameters of the stock-recruitment model may have little importance, although they should be plausible.

Figure 14-5. The Ricker Model fit to data for Sacramento River fall-run Chinook before and after construction of the Red Bluff Diversion Dam. Data points are labeled by year; upper curve was fit to 1962-66, lower curve to 1967-83. Copied from Kope and Botsford (1990)

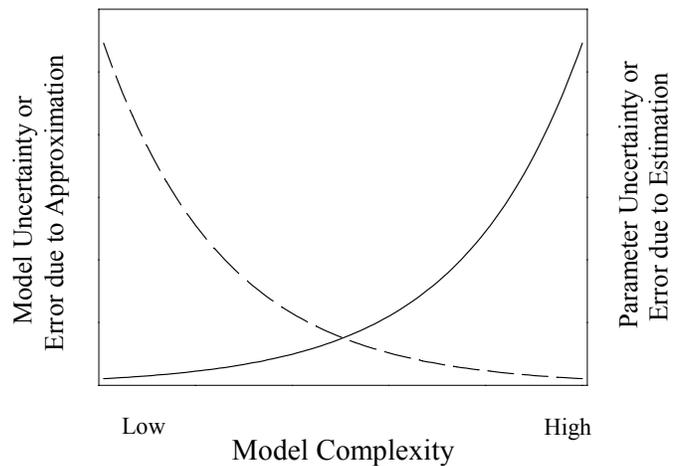


Stock-recruitment models are simple population models, although their implementation can be very sophisticated (e.g., Punt and Hilborn 1997; Schnute and Kronlund 2002), and more complex population models are in a sense stock-recruitment models that take environmental or other factors that influence mortality explicitly into account. Individual-based models (e.g., Jager et al. 1997; Kimmerer and Jones & Stokes 1999) take this elaboration a step farther by tracking variation among individuals, or groups of individuals, within the population. Such models are attractive because they seem more realistic, but many mathematically sophisticated modelers think that this advantage is illusory (e.g., May 2004), at least for models used for making predictions. One fundamental problem is the need to estimate parameters.

The Deriso model,  $R = aS(1-bcS)I/c$  (Deriso 1980), illustrates the point that more complex models may not perform as well as simpler models. Although the extra parameter allows the Deriso model to mimic the Ricker and Beverton-Holt models and to fit intermediate cases, it performs poorly in practice (Schnute 1985). When parameters must be estimated from data that are measured with error, as is almost always the case with fishery data, a simpler model may give better predictions. Ludwig and Walters (1985) demonstrated this by generating synthetic population data with an age-structured population model based on the Deriso model.

Measurement errors were simulated by adding modest random terms to the data. Different models were then fit to the simulated data to estimate optimal yield. A model based on the Ricker Model that did not account for age did as well or better than the model used to generate the data. There is a fundamental trade-off between model uncertainty, or error due to approximation, and parameter uncertainty, or error due to estimation (Figure 14-6); accordingly, “Effective management models cannot be realistic” (Ludwig 1994:516).

Figure 14-6. Conceptual model of the trade-off between model uncertainty (dashed line) and parameter or estimation uncertainty (solid line). In a good predictive model these two types of uncertainty are balanced. Redrawn from Ludwig (1994).



A common and seemingly persuasive response to arguments such as those just made is to show that the model in question can accurately reproduce historical data such as escapement, given historical data on flows, etc., as input. It seems, intuitively, that in such a case the model cannot be far wrong, but this is a case where intuition fails. Often, a model with many adjustable parameters can give good fits to data with more than one set of parameter values, a problem known as equifinality. Since all the sets of parameter values cannot be “right,” getting a good fit does not “validate” the parameter values.

An alternative defense of the use of more complex models is the plea that although the actual predictions are not reliable, the relative performance of simulated management actions can still be compared (e.g., Kimmerer et al. 2001). This argument depends on the assumptions that the aspects of the actions that matter to salmon are accurately reflected in the model, and that the rest of the system is more or less linear. These are both brave assumptions.

In evaluating models used for prediction, then, the number of parameters must be taken into account. Several statistical criteria can be used to do this. One criterion is the adjusted residual squared error, which is the residual sum of squares divided by  $n-2p$ , where  $n$  is the sample size and  $p$  is the number of parameters (Efron and Tibshirani 1993). Another is the Akaike information criterion (AIC), advocated by Burnham and Anderson (1998); it is the negative log likelihood of the data given the model (a measure of how well the model fits the data, where more negative is better) plus twice the number of parameters in the model. Burnham and Anderson (1998) give a good discussion of model selection, and use of the AIC has become common in the ecological literature. However, some good statisticians think that the AIC tends

to select models with too many parameters (T. Speed, personal communication 2002), so it may be prudent to apply several criteria and proceed with caution if different methods rank models differently. Similarly, Burnham and Anderson (1998) suggest that if the ranking of two models with the AIC is close then the best course may be to use both, and this also seems good advice for cases where different criteria give different rankings. Moreover, any of these statistical criteria may select models that are biologically implausible, so models should be evaluated on that basis as well (Burnham and Anderson 1998), and even if a model is plausible and fits the existing data well, the biological idea expressed by the model may still prove wrong. For example, if flow and temperature are highly correlated, a model based on flow may give good predictions of an effect that is really a consequence of temperature.

### **Using models to evaluate restoration strategies**

Managers often want to use models to evaluate restoration strategies, in the sense of predicting whether a given set of restoration actions will allow a population to meet some abundance criterion. For such purposes, more statistical approaches are needed. The analyses of smolt survival by Newman and Rice (2002) and Newman (2003) described in Ch. 10 and Appendix B are examples. In another example, Lindley and Mohr (2003) used an elaboration of the Ricker model in a Bayesian framework to model the effects of striped bass predation on winter Chinook. This allowed them to assess whether resumption of stocking striped bass would significantly increase the risk that winter-run will go extinct. This approach uses the simplest models that describe the biology relevant to the problem at hand. In this case, the interest was on the effects of striped bass, so the model included a term for the effects of striped bass predation on the growth rate of the winter-run population, a term for the effect of density-dependent mortality on the population growth rate, a term for the mean population growth rate in the absence of bass predation and density-dependent mortality, a term for the possible change in the population growth rate resulting from conservation actions after 1989, and a “process error” with a mean of zero and a variance estimated, along with the parameters in the other terms, from the data. The data in this case are estimates of winter-run escapements and of the abundance of adult striped bass. Thus, the model includes five parameters to be estimated from data, and a few such as the proportion of winter-run that mature at age three and age four that were taken from the literature. Five is not a small number of parameters to estimate from data, but the model is much simpler than models such as SCSIFM and the others described above.

Models such as Lindley and Mohr’s have several important advantages over other types of models. They meet management needs for a basis upon which to make decisions, in this case whether to allow stocking of striped bass to resume, and they also provide assessments of the range that the parameters might plausibly take. With Bayesian models, these assessments come in the form of the posterior distributions (e.g., Figure 14-7). In this case, for example, the posterior distribution for  $\alpha$ , the per-bass predation rate, has a median that translates into about a 9% chance that a juvenile winter-run will be eaten by a bass (Lindley and Mohr 2003). The distribution shows that small values of  $\alpha$  are most likely, but large values are plausible. Generally, the posterior distributions summarize the information in the data, given the model, in a form that is useful for management.

Besides the estimate of bass predation, this model also provides an assessment of the post-1989 conservation measures. The median estimate and 0.9 confidence interval for the relevant parameter indicate that, based on this model and the data, more likely than not the post-1989 conservation measures did some good. Although this study was not published until 2002, the actual work was done years before when data were available only through 1996, before the recent increase in winter-run escapements (Figure 2-7). Even using the pre-1997 data, however, allowing for a change in the mean population growth rate resulted in an assessment that winter-run are less likely to go extinct, compared to the results of Botsford and Brittnacher (1998), who assumed a constant mean population growth rate.

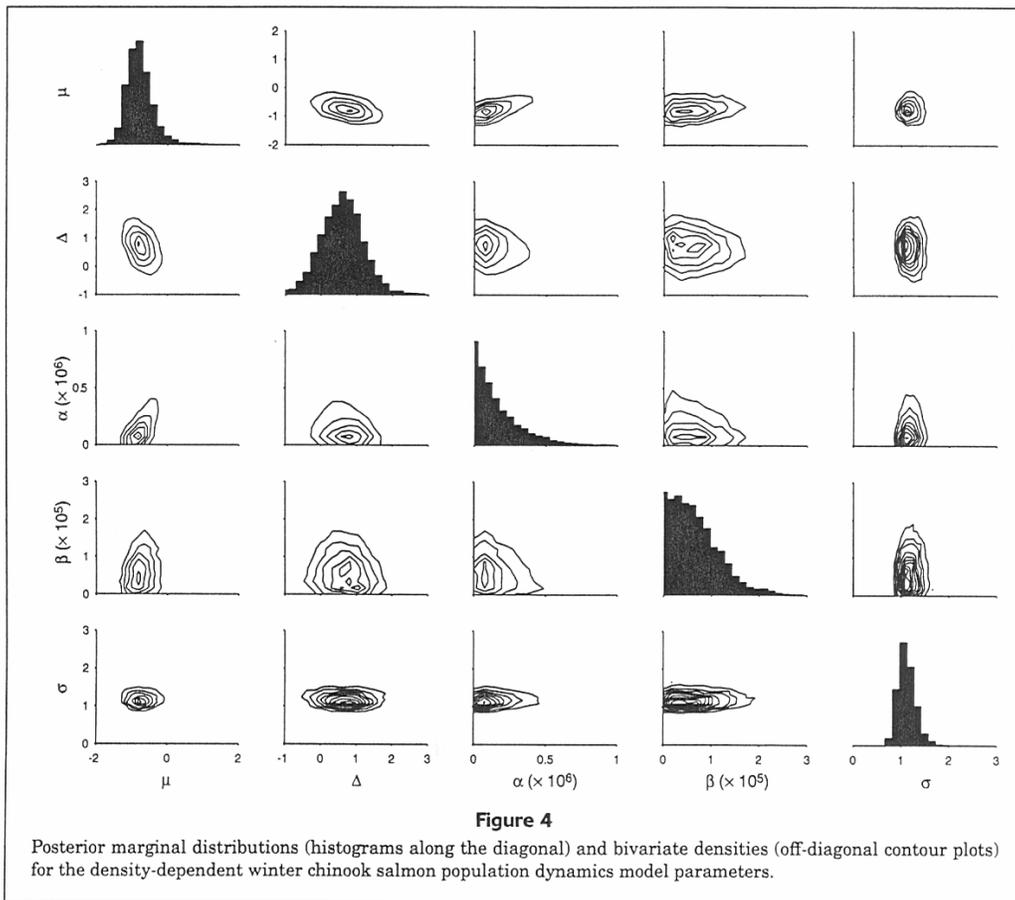


Figure 14-7. Posterior distributions (main diagonal) and bivariate densities for the parameters of the models used to assess the effects of stocking striped bass on the population of winter-run Chinook:  $\mu$  (growth rate),  $\Delta$  (growth rate change),  $\alpha$  (predation rate),  $\beta$  (density dependence), and  $\sigma$  (process error). These show, for example, that the bass predation rate is probably small but large values are not inconsistent with the data. Note that values of  $\alpha$  and  $\beta$  are very small and are multiplied here to make the axes legible. Copied from Lindley and Mohr (2002).

One potential problem with this model is that it does not account explicitly for changes in ocean conditions, which could inflate the apparent success of post-1989 conservation measures and lead to an overly optimistic assessment of the condition of winter-run, since conservation

measures can be made permanent but changes in ocean conditions cannot. This could be dealt with by comparing the results of this model with the result of a model that included a parameter for ocean conditions. However, if the interest is in the post-1989 conservation measures rather than the effects of stocking striped bass, then probably the model should be revised by leaving out the striped bass parameter.

The main disadvantage of models such as Lindley and Mohr's (2003) is that most biologists and managers, including the author of this review, have a limited grasp of the statistics and mathematics involved. This is even more true of "hidden process models," described by Newman et al. with an application to winter-run (In press). Although models such as SRCSIBM are complex (e.g., Figure 14-2), the individual steps are easy enough for people with limited quantitative background to understand, and this allows for a level of comfort that does not extend to modeling that in a sense has to be accepted on faith. However, the basic ideas of this kind of analysis arguably are more intuitive than those of conventional frequentist statistics, so the problem is largely one of language and familiarity. This problem should ease over time, as people become more accustomed to hearing about prior and posterior distributions, conditional probability densities, etc. It will also help if managers provide appropriate training opportunities for their staffs, and if more people with strong training in these methods become involved in Central Valley salmon studies. Recruiting such people should be a priority for agencies concerned with salmon management in the Central Valley.

When too few data are available to fit an estimation model, simulation models are the only available option for exploring the plausible outcomes of some management action. The question again arises, how complex should the model be? Two basically different approaches are advocated in the literature. Some modelers, for example Starfield (1997), advocate the use of simple simulation models tailored to the particular problem at hand. Others, for example Jager et al. (1997) and Kimmerer and Jones & Stokes (1999), propose the use of highly complex simulation IBMs that seem too expensive to develop except as general purpose models, intended to address a whole range of questions. The CALFED Environmental Water Account (EWA) Review Panel has recently recommended development of such a model for Central Valley Chinook (2002 EWA Review Panel 2002).

This recommendation can be questioned, for several reasons. One is that many aspects of Chinook life history remain poorly understood. It is possible to define rules that generate what seems like reasonable approximations to observed behavior, but this does not show that the model has "got it right." On the one hand, important aspects of the behavior may not be well observed, and on the other hand, other sets of rules may produce equally reasonable behavior. Even where the rules or relationships are relatively well known, there can still be the need to select values for parameters, as discussed above and elaborated below.

Another objection to complex simulation models is based on experience with human nature. Schnute (2003:409) remarked that "Everyone recognizes that simulation provides a thought experiment, not an accurate representation of nature." Unfortunately, experience shows that

many people do not recognize this, and even people who do recognize it intellectually may forget it in the often difficult process of making decisions. There is less danger that a simple model will be confused with “an accurate representation of nature.”

A third and related objection is that the level of detail in the models varies among life stages, depending upon the amount of information available. Less is known about the ocean phase of the salmon life cycle, for example, so of necessity this is modeled in less detail. To the extent that the model is intended to serve as a summary of available understanding, this is unobjectionable. However, if the model is intended to assess restoration or mitigation measures, then the model is only as strong as its weakest link, and the scientific purpose of detailed simulation in some aspects of the model while other aspects are generalized is unclear. May (2004:793) noted that “It makes no sense to convey a beguiling sense of ‘reality’ with irrelevant detail, when other equally important factors can only be guessed at.” The Recovery Science Review Panel, convened by NOAA Fisheries to provide guidance for salmon recovery efforts coast-wide, was sharply critical of the Ecosystem Diagnosis and Treatment (EDT) model, which is now widely used in Washington and Oregon, largely on that basis: “The inclusion of so much detail may create an unjustified sense of accuracy, but actually introduces sources of inaccuracy, uncertainty, and error propagation” RSRP (2000:6).

A final objection is based on the view that models should be used as aids to thought. When models become complex they become “black boxes” that are not well suited for this purpose, and tend to become substitutes for thought instead. This is particularly true when models appear to provide answers to important management questions. Instream flow models, discussed below, are a good example of this problem.

A simulation model that explores the consequences of density-dependent mortality and density-dependent migration on the potential utility of classes of restoration actions for ocean-type Chinook salmon in Puget Sound (Greene and Beechie 2004) seems to occupy a reasonable middle ground of complexity. This is a modified Leslie Matrix model that uses survival estimates at various life-stages to step the simulated population through multiple life cycles. The stages are: redds, streams, tidal deltas, nearshore habitats, and the ocean. It uses Beverton-Holt density-dependence, and allows for three different juvenile life history patterns: parr migrants, that rear in the stream, delta users, that rear largely in the tidally influenced area near the river mouth, and fry migrants, that migrate directly to nearshore areas in Puget Sound. Altogether, the model has 40 parameters for which values were taken from other sources. The model seems sufficiently generalized that to minimize the risk that it will be mistaken for an “accurate representation of nature,” but sufficiently detailed to characterize the main elements of the Chinook life cycle and their use of habitats.

In using their model, Greene and Beechie simulated a base case, in which density-dependent mortality occurs only at spawning, and ran the model until the simulated population stabilized, for four model watersheds with different areas of spawning, stream, delta, and nearshore habitat, and two real watersheds with contrasting distributions of habitat area. They also simulated cases

with density-dependent mortality and density dependent migration, again running the model until the population stabilized. Finally, they adjusted various parameters by 5% for each case to explore how the equilibrium population responded to the adjustment in the various cases.

The Greene and Beechie (2004) model seems most useful not for indicating what kinds of restoration actions would have the most effect on a particular population of Chinook, but rather for showing what one needs to believe about the density-dependent processes affecting the population in order to believe that a certain kind of action would have the most effect. That is (p. 598), “understanding how restoration actions may effect salmon populations depends upon knowledge of density-dependent mechanisms.” Unfortunately, little is known about density-dependent mechanisms or processes affecting Central Valley Chinook. The model also suggests, at least as parameterized for Puget Sound Chinook, that nearshore habitats in the sound are particularly important. We know relatively little about comparable habitats for Central Valley Chinook.

### **Other uses of models**

Models have other uses than making predictions, and complex models may serve such purposes well. Models may also help structure discussions or negotiations among different parties involved in natural resources management (Volkman and McConnaha 1993), and relatively simple simulation models play a central role in adaptive management process described by Walters (1986). Complex models may also be useful in this context. Despite the RSRP’s harsh assessment of the EDT, it is nevertheless being used for recovery planning in Washington and Oregon (PSTRT 2002; W/LCTRT 2003), apparently because it proved useful in the political processes that are involved in recovery planning and in other restoration activities (J. Anderson, UW, personal communication 2004). The spatially-explicit description of the system in the model allowed people from different areas to contribute their knowledge to the modeling process, and various parties could set up model runs that reflected their views of the main problems and favored remedies. When these produced implausible results, the parties were persuaded to become more open-minded to alternative views of the problems. These are real virtues, but effective management needs to be sound scientifically as well as politically.

The process of developing complex simulation models can be useful as a way of organizing existing knowledge and identifying gaps in understanding, even if the model predictions are unreliable. In this use, however, the main benefit may go to the people who develop the model, and there is a danger of confusing the ability to specify rules that give plausible results with real understanding.

A more fruitful use of complex simulation models is to treat them as experimental systems, distinct from the “real world,” in which the consequences of various sets of assumptions can be examined (Peck 2004). For example, complex simulation models can be used to assess other modeling approaches. Railsback et al. (2003) used a complex individual-based model (IBM) of trout to show that, given reasonable assumptions about trout behavior, the depth and velocity at foraging sites selected by simulated fish in a simulated stream should change with discharge.

This finding challenges a basic assumption of the Physical Habitat Simulation System (PHABSIM), which is discussed below.

Complex simulations that are focused on specific questions may also be useful. For example, Haefner and Bowen (2002) have developed an IBM for the behavior of fish near fish screens, which is intended to be used in designing a new facility. Physical models may also be useful for this purpose (e.g., Swanson et al. 2004), and the two types of modeling could be used in conjunction.

### **More on parameters**

With estimation models, as defined above, parameters are estimated by fitting the models to data, as described for example by Hilborn and Mangel (1997). With simulation models, parameter values are obtained independently. This can raise problems. The RSRP was particularly critical of using professional judgment for setting many parameter values in the EDT model. Alternatively, parameter values may be taken from the literature, although the values may be adjusted later to give a better fit to data (e.g., Jager et al. 1997). Obtaining parameter values from the literature has its own set of dangers, as detailed by Trudel et al (2004) in terms of parameters for bioenergetic models. They compared the predictions of four published models with observations and found that (p. 337):

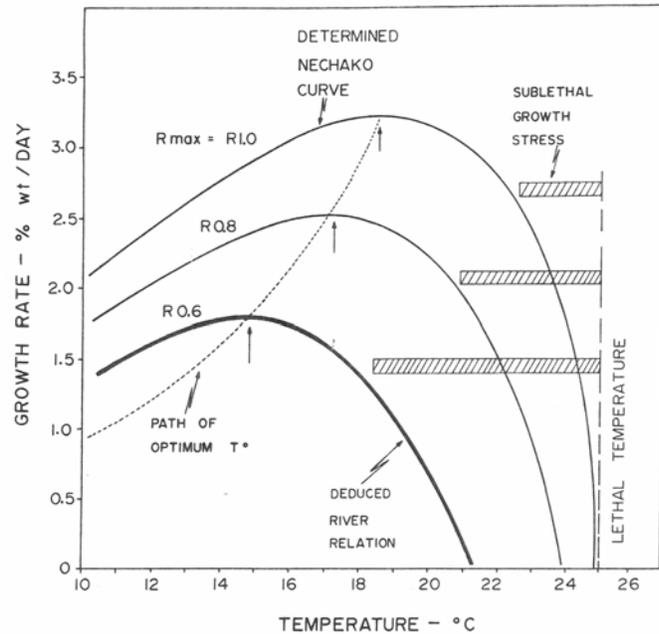
- (1) oxygen consumption rates for sockeye salmon and steelhead were not always accurately predicted by models that were derived specifically for these species, (2) metabolic rates of Pacific salmon and steelhead were not accurately predicted by models developed from closely related species, and (3) allometric equations of metabolic rates were not stable when measured over small ranges of sizes, and were inaccurate when extrapolated to smaller or larger fish.

The situation can be even worse for parameters that cannot be determined in laboratory experiments, as illustrated here by the example of the ration parameter, which appears in various salmon models. The ration parameter is the ratio of the food consumed by juvenile salmon in the circumstances modeled to the food that would be consumed were food readily available. The ration parameter matters because it strongly affects estimates of the relation between water temperature and growth (Figure 14-8), as demonstrated by extensive experiments with juvenile sockeye (Brett et al. 1969).

The Oak Ridge Chinook model (Jager et al. 1997; Jager and Rose 2003) has a parameter for the maximum ration at feeding stations, set at 0.6. Jager et al. (1997) do not cite an authority for the estimate, but the value 0.6 probably comes from Brett et al. (1982), a litigation-driven study of Chinook in the Nechako River, British Columbia, an interior tributary to the Fraser River where diversions for hydropower decreased flow and increased water temperature. Brett et al. (1982) was published as a Canadian Technical Report of Fisheries and Aquatic Sciences, that is, as grey literature, not subject to formal peer review. Moreover, Brett et al.'s objective was to estimate the effect of the increase in temperature on the growth of Nechako River Chinook, for

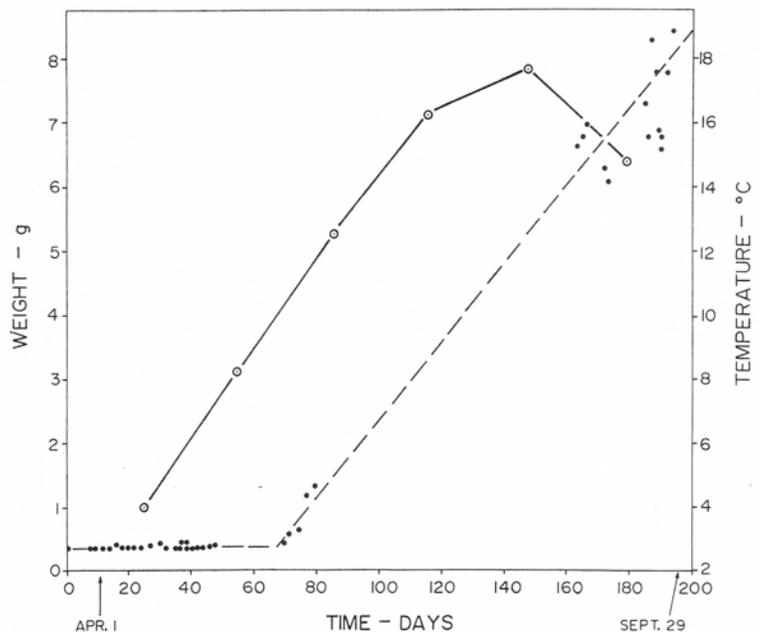
which they needed the estimate of ration level (Figure 14-8). In other words, estimating ration level was not the primary purpose of the study, but rather a step in the main process at hand, and the estimate of ration level actually came from application of a growth-ration model that is described only in theses (Stauffer 1973; McLean 1979) to an estimate of growth developed from 5 and 14 samples of 10 to 50 fish collected in early June and September in the Nechako River, and mean monthly temperature data (Figure 14-9).

Figure 14-8. Growth rate v. temperature curves for Nechako River Chinook. The top curve is derived from laboratory experiments. The lower curves are estimated based on data for juvenile sockeye and data on the size of juvenile Chinook in the Nechako River (see Fig. 5). Copied from Brett et al. 1982.



These data are sparse at best, and estimating growth rate from sequential samples entails strong assumptions, for example that neither mortality nor the capture efficiency of the gear used is size-selective. Brett was a prestigious salmon biologist, which probably accounts for the ready acceptance and wide application of the 0.6 value, but it is a weak reed to lean upon. Even if the estimate of ration level were accurate for the Nechako River in 1981, it is a bold step to apply it elsewhere, particularly since not that many data would be required to develop a comparable estimate for the stream in question, and the fit of the model to the data could be assessed. In that case, however, the citation would be to unpublished theses, rather than to an established expert in the field.

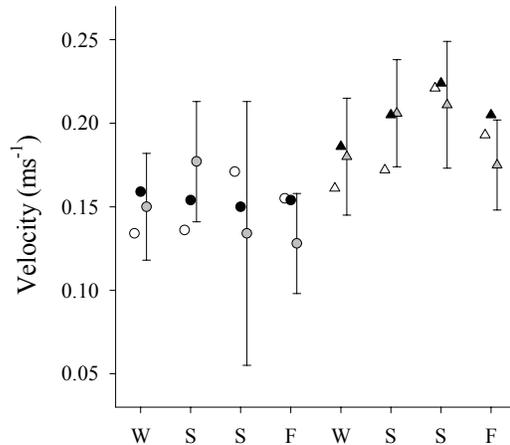
Figure 14-9. Data used by Brett et al. (1982) to develop the estimate that juvenile Chinook in the Nechako River in 1981 experienced a ration level of 0.6. The dashed line is the estimated growth rate. Fish were collected by seining and by trapping, but no specifics were given. Copied from Brett et al. (1982).



Although it is easy to imagine improving on the estimates developed by Brett et al. (1982), there are practical limits to what can be done with wild living resources. It may be useful to contrast the use of complex models in fisheries management with, say, detailed models of the growth and development of crop plants that are used in conjunction with field and laboratory studies. Unlike wild Chinook and steelhead, crop plants have been selected for uniformity in life-history patterns, at least within varieties, and their environments are more or less controlled. More importantly, the growth and development of crop plants can be observed in detail, so intermediate predictions of the model can be tested, and parameter values can be obtained from replicated field or growth chamber experiments. Wild fish are more variable and enormously more difficult to observe, let alone experiment upon, and effective management models for them therefore must take statistical considerations regarding the estimation of parameters into account. These considerations limit the optimal complexity of the model.

A final example may illustrate the benefits of simplicity in modeling for making predictions. Hill and Grossman (1993) used a bioenergetics model to develop good predictions of the water velocity at positions selected by rainbow trout and rosyside dace of different sizes in different seasons. However, they got better predictions from a simpler model based only on prey capture success rate (Figure 14-10), which is a decreasing function of water velocity (as velocity increases beyond some threshold, the fish miss drifting prey more often). In this case, the results of the bioenergetic model clarify why the simpler model works better. The energetic cost of holding position at realistic water velocities increases only slowly with velocity, and is small compared to the energy gained from successful foraging. Thus, the energy balance is determined mainly by the rate at which prey are captured. The more complex and mechanistic bioenergetics simulation may be satisfying intellectually, but serves less well for predictions.

Figure 14-10. Predicted and observed water velocity of positions selected by 52-7 mm rainbow trout (circles) and by 71-125 mm rainbow trout (triangles), for winter, spring, summer and fall. Observed mean values have error bars showing 95% confidence intervals; open symbols show predictions of the bioenergetic model, filled symbols of the capture-rate based model. Data from Hill and Grossman 1993.



### Instream flow models

In the complex population models discussed above, a salient weak link is the need to relate the habitat value of streams to the rate of flow. How to do this is an unsolved problem (Castleberry et al. 1996). The model most commonly used is the Physical Habitat Simulation System (PHABSIM). This model, developed by an interagency group in the late 1970s, combines hydraulic and simple biological models to estimate a parameter called “weighted usable area” (WUA) as a function of flow (Bovee et al. 1998). Other “numerical habitat models” that combine hydraulic and simple biological models have also been proposed (Guay et al. 2000), as have others with more complex biological models (Van Winkle et al. 1996; EPRI 2000), but PHABSIM is still the standard method. At least two of the salmon population models used in the Central Valley (SALMOD and ORCM) are built on PHABSIM, and SRCSIBM uses WUA as an input, at least as an option.

WUA is intended to be a metric of the habitat value of a reach of stream. In most applications of PHABSIM, patches of habitat are characterized in terms of three “microhabitat” variables: depth, water velocity, and substrate size, at different rates of flow, for the species and life stage in question. Within each patch or “cell,” substrate size is estimated from field data, and either one or two-dimensional hydraulic models are used to estimate mean depth and water velocity. At each discharge, the habitat value of each patch is evaluated according to “suitability curves” ranging from zero to one for each variable, and the product of these values and the area of the patch produces an estimate of WUA. These are summed over the patches, producing a curve of WUA over discharge for the reach. PHABSIM is normally used as a spatially-explicit model, because the estimates of WUA are made patch by patch.

Since the 1980s, PHABSIM has been applied to many Central Valley streams to estimate the relation between flow and habitat for Chinook or steelhead. In some cases, however, flow recommendations were based on considerations such as water temperature rather than WUA (e.g., Aceituno 1993). Recent or ongoing applications include the studies of spawning habitat in

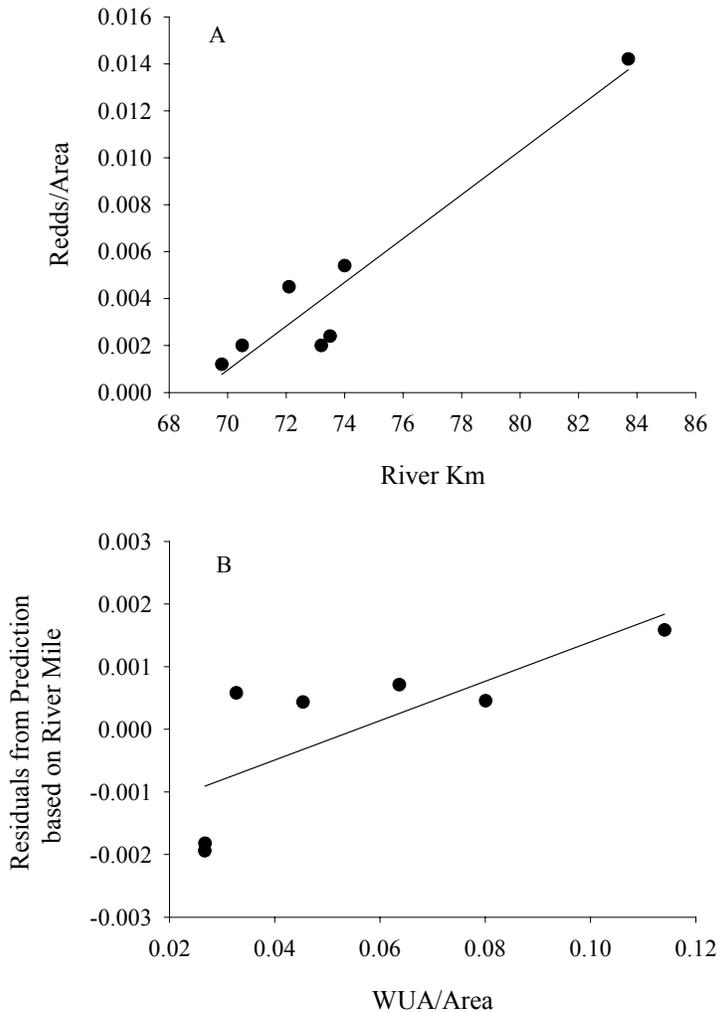
the Merced and American rivers (Gallagher and Gard 1998), the Feather River (Thomas R. Payne & Assoc. 2002), and on Butte Creek and the Sacramento River by USFWS..

Although PHABSIM is widely used, it has long been sharply criticized on biological, hydraulic, and statistical grounds (Marthur et al. 1985; Shirvell 1986; 1994; Scott and Shirvell 1987). These criticisms have recently been extended (e.g., Castleberry et al. 1996; Williams 1996; Bult et al. 1999; Kondolf et al. 2000; Bradford and Higgins 2001; Williams 2001b; Holm et al. 2001, Heggenes 2002; Railsback et al. 2003; Kemp et al. 2003), and the method has much more support among agency scientists and consultants than among academic scientists.

In the Central Valley, some biologists who are critical of PHABSIM nevertheless think that it “works” for Chinook spawning. Support for this view is provided by Gallagher and Gard (1999), who applied PHABSIM to sites “known to be heavily used for spawning” on the Merced and American rivers, and found a moderate but highly significant correlation between WUA and the density of Chinook redds in PHABSIM cells in the Merced River (cell by cell results for the American River were not reported). However, it is not clear what they would have found if they had applied PHABSIM to riffle areas generally, rather than only to areas that were heavily used and presumably had good hyporheic conditions. More seriously, based in the data in tables 1 and 5 of Gallagher and Gard (1999), position along the river (River km) explains most of the variation in the density of the redds considered in the Merced River (Figure 11-15a; similar data were not given for the American River). Spawning habitat quality, as estimated by PHABSIM, is related to the residuals from estimates based on river kilometer, but the residuals are not large compared to the reported range in spawning density, and the relationship is not strong (Figure 11-15b).

Gallagher and Gard (1999) found stronger relationships between WUA and redds in both rivers when the data were aggregated by site. Since PHABSIM is calculated cell by cell, however, the meaning of this finding is unclear. The belief that the aggregated findings strengthen the cell by cell findings seems to be an example of the “ecological fallacy” (Freedman 1999), that is, the belief that relationships observed for groups necessarily hold for individuals. One problem in this case is that the density of redds increases with distance upstream. Hence, the site-level analyses are confounded by a location variable, even if stream and hyporheic conditions are otherwise equal in all sites. This holds even if the most upstream site is excluded, as it was in some of the reported analyses.

Figure 14-11. A. Density of spawning in the Merced River is determined primarily by position along the river, or distance from Crocker-Huffman Dam ( $r^2 = 0.91$ ). B. Spawning habitat quality as estimated by PHABSIM helps to explain the residuals in A ( $r^2 = 0.58$ ), but the residuals are not large and the relationship is not quite significant by the standard criterion ( $p = 0.054$ ). Data from tables 1 and 5 in Gallagher and Gard (1999).



Castleberry et al. (1996), who called PHABSIM “not defensible scientifically,” recommended that an adaptive approach be applied instead, pointing to the example of the Hodge Decision (Hodge 1990) in the case of EDF et al. v. EBMUD. Judge Hodge set flows that conditioned EBMUD’s ability to divert from the American River at Nimbus Dam, based on the evidence presented. However, he also recognized the great uncertainty in the evidence, so he retained jurisdiction and ordered the parties to cooperate in studies intended to determine whether changes should be made in the standards. Probably the best work of this sort is being conducted by BC Hydro (e.g., Higgins 2001; McDaniels and Gregory 2004; Gregory et al. 2006) as part of the British Columbian equivalent of a FERC relicensing process.

PHABSIM seems to provide information that managers and decision-makers want on the relation between the flow in a stream and its value as habitat. Probably on that account, it continues to be used, despite the unanswered scientific objections to it cited above. At the least,

however, managers and decision-makers should be aware of these objections, and should require the following from parties presenting PHABSIM results for their consideration:

1. Scatter plots comparing the predictions of the hydraulic model used with measured data at randomly located sites in the study area, as described in Williams (2001b);
2. Scatter plots of WUA predicted by PHABSIM with WUA calculated from the measured data at the same locations.
3. Scatter plots of the data used to develop the habitat suitability criteria used, overlain with line plots of the criteria, together with bootstrap confidence intervals for the criteria (see Williams et al. 1999, for an example of such confidence intervals).
4. Plots comparing WUA calculated with the selected suitability criteria with WUA calculated using selected 5<sup>th</sup> and 95<sup>th</sup> percent confidence intervals for the suitability criteria.<sup>2</sup>
5. For PHABSIM studies using 1-dimensional hydraulic models, a plot showing bootstrap confidence intervals around the curve of WUA over discharge, as described in Williams (1996).
6. Comparisons of the PHABSIM results with results of alternative models, e.g., that redd density depends on position along the stream.

## **Recommendations**

1. For most purposes, models should be tailored to the problem at hand, and simpler models should be favored over more complex models. Unnecessary complexity should be avoided.
2. Where adequate data exist, the potential effects of important restoration measures should be evaluated with estimation models; simulations should be used otherwise.
3. Monitoring programs should be assessed with simulations.
4. More people with good training in modeling and statistics should be employed or retained by agencies concerned with Central Valley salmon issues, and training opportunities should be provided for existing staff.

---

<sup>2</sup> With confidence intervals with three habitat suitability criteria, there are 720 possible comparisons to be made, so clearly some judgment must be exercised in selecting examples to plot for this recommendation to be practical.