APPENDIX T

S. Heppell, *Review of Chaloupka model*

1. Executive Summary

This report is a review of the stochastic simulation model of Western Pacific leatherback sea turtle metapopulation dynamics developed by Dr. Milani Chaloupka for the U.S. National Marine Fisheries Services, Honolulu Laboratory, Honolulu, Hawaii. The files examined were the Berkeley Madonna (Copyright 1997-2001, Robert I. Macey and George F. Oster) file, *leatherback(User).mmd*, and the associated user manual, *Stochastic simulation model of Western Pacific leatherback sea turtle metapopulation dynamics: User's Guide (January 2002)*. Berkeley Madonna version 8.0.1 was used to run the simulation model.

The report examines the Chaloupka simulation model within the context of modeling in general. Parameter uncertainty, suggested protocol to assess the impacts of fisheries, and sensitivity analysis of model parameters on critical endpoints are discussed. The report concludes with a summary and statement of research recommendations

The Chaloupka simulation model has the potential to serve as a key component in a suite of analytical tools to assess the viability of sea turtle populations. We recommend that such research address the following tasks:

- Perform external validation to search for points that validate and define the boundaries and validity of the Chaloupka model.
- Identify sentinel variables (observable indicators of trends that are sensitive to parameter settings) to determine when the model can be applied and produce results consistent with real-world conditions.
- Perform prediction risk to determine consequences of incorrect model results.
- Construct a meta-model of model accuracy (below).

The complexity of the Chaloupka simulation model justifies making the tool itself and the results it produces objects of study. Sensitivity analyses will inform analysts and decision makers about inherent variability in model results. In particular, a "meta-model" that focuses on parameter values deemed speculative by expert analysts is warranted to gauge confidence in specific instantiations of the model. By varying all of those speculative parameters simultaneously, this meta-model would serve as a complement to the simulation model to alert analysts to possible interactions among certain parameter values that might potentially produce non-additive effects. These synergistic effects are potentially dangerous because they can precipitate rapid model population declines.

While this model provides a good framework for heuristic evaluation of "what if" scenarios, we strongly recommend that the research tasks listed above be completed before the model is used to make quantitative predictions for management. This is particularly true for leatherbacks, which are a poorly understood species, at best. It is critical to understand how the model results are affected by parameter uncertainty and assumptions that are incorporated into the structure of the model (which are not necessarily obvious to users), and to anticipate the potential for conflicting, yet equally valid, results that may be generated by various stakeholders.

2. About This Report

This short report represents the opinions of Dr. Selina Heppell, a fisheries ecologist at Oregon State University and author of several papers on sea turtle demographic models, and Dr. Jane Jorgensen, a private consultant with expertise in ecological models and Bayesian belief networks. We were unable to do more than a precursory review of the model and its potential as a management tool due to the short time frame allotted for review (1 month). Our goal was to determine the general structure of the leatherback model and to offer suggestions for how the model should be used by NMFS. We did not review the other simulation models (green and loggerhead turtles); although the general structure of those models is quite similar, far more empirical data exists on those species. Nevertheless, our concerns about the complexity of the simulation, given the available data and the precision of those data, apply to all of the models.

3. Overview: How Models are Created, Parameterized, Analyzed, and Applied to Management Problems

The Chaloupka simulation model is a powerful tool for detecting and projecting trends in sea turtle populations. Careful use of this model can help us heuristically evaluate the relative merits of alternative management programs under a range of plausible biological scenarios. It creates a common forum for discussion by creating a framework within which to relate multiple observations and hypotheses.

However, as with any fine tool, the quality of the resulting product will depend in large part on the quality of the raw materials (input data) and the skill of the user. Levins (1966) described models as tradeoffs among generality, realism and precision. He listed three model-building strategies:

1. Sacrifice generality for realism and precision. Models produced using this strategy yield precise predictions for tightly constrained situations. This approach has been adopted by natural resource managers to formulate precise, testable predictions based on the short-term behavior of organisms.
2. Sacrifice realism for generality and precision. This approach yields very general models that generate very precise predictions. However, the equations may be unrealistic given the conditions in the natural world. Small departures from initial assumptions often have large effects upon predicted outcome.
3. Sacrifice precision for generality and realism. This approach produces models that focus on change in terms of relative change among variables in complex dynamic systems. The results of the models are very generalizable in terms of the qualitative nature of behaviors, but do not contain numerical precision.

The Chaloupka simulation is realistic because it has a sound scientific basis, and precise because it can produce more or less exact results based on parameter estimates. Generalizability in the simulation is provided by the user-driven input of model parameter values using slider bars. However, as Levins noted, no modeling strategy can fully exploit all three strategies simultaneously. The Chaloupka simulation model most closely adheres to Levins's second strategy, maximization of realism and precision.

There are at least 151 user-defined parameters in the model, including survival rates, harvest levels, periods of time in which harvest occurs, density-dependent effects, nesting beach temperatures, and the level of stochastic environmental variability. This level of biological realism has not been available in other published sea turtle models (e.g., Crouse et al. 1987, Heppell et al. 1996), and is appealing to biologists (S. Eckert, *personal communication*). Input parameters can be entered into the simulation using a slider-bar interface. These input values may be driven by observational data or by expert opinion mediating these data. Because of the migratory habits of the sea turtles, the difficulties associated with observing them in their natural state, and the surreptitious manner of many anthropogenic hazards (e.g., poaching of adults and eggs), the observational data are inadequate for the leatherback. There are many facts about sea turtles that remain unknown, such as dwell-time in fishing areas, response of individuals and vital rates (growth, survival and reproduction) to population density, and quantified data on human-induced and natural mortality. Even age at maturity is somewhat speculative in leatherbacks.

While uncertainty within the model may be controlled and explored, the data from which parameter estimates are drawn may contain substantial uncertainty. Schmitt and Klein (1996) identified four sources of uncertainty, all of which are present in the data regarding sea turtles:

1. Missing information. Information is unavailable. It has not been received or has been received but cannot be located when needed; for example, information about dwell-times in fishing areas.
2. Unreliable information. The credibility of the source is low, or is perceived to be low even if the information is highly accurate; for example, by-catch by non-US longline fisheries and nesting beach data from several Pacific locales.
3. Ambiguous or conflicting information. It is difficult to integrate the different facets of the data; for example, radio-transmitter data for migratory patterns of individuals and migratory routes by genetic stocks.
4. Complex information. It is difficult to integrate different facets of data; for example, the estimated abundance of nesting females may be dependent upon a host of associated factors, including sampling error.

Sea turtles, particularly leatherbacks, suffer from all four of these sources of uncertainty. It is crucial to accept and incorporate that uncertainty into any assessment, regardless of how precise the model output from individual simulations may be. Unfortunately, uncertainty in the covariance of these parameters may be impossible to assess.

Analyses using this model will produce a wealth of "data" that may be used to formulate and prioritize management actions as alternative hypotheses. The flexible modeling framework invites the creation of alternative models for discussion and analysis. Managers should be aware, however, that the end result of this increased modeling and analysis activity may not be increased clarity. Shenk (1997) refers to this phenomenon as the "Sixth Law of Data Smog – Too many experts spoil the clarity". The New York

Times calls the plethora of models generated by opposing experts "volleys of data." Managers should prepare for extended argumentation that although scientifically sound, is not definitive. Communication of the reasons for management activities to stakeholders may likewise become more difficult as the complexity of the arguments driving these activities increases.

The User's Manual is well written, at least for ecologists moderately familiar with modeling terminology, and provides an excellent introduction to the simulation model. The graphical interface is intuitive enough to allow naïve users access to the model. As an accessible model that may be used to evaluate management alternatives, the Chaloupka simulation may encourage relatively inexperienced decision makers to approach the formulation of management plans by adjusting parameters until a desirable output is produced. Glover et al. (1997) found that structured aids influenced these novice analysts to approach aided tasks "mechanistically, without becoming actively involved in the task or judgment." Paradoxically, while this tool has been designed to enhance understanding of sea turtle population dynamics, as a decision-making aid it may have the opposite effect on some subset of user analysts. This phenomenon applies to all structured decision making aids and in not unique to natural resource management models. However, in a supervised learning environment, the simulation model may prove immensely valuable as a didactic tool to demonstrate the way that complex effects are coherently related through an expert's eyes.

4. Appropriate Consideration of Parameter Uncertainty for Assessment of Management Alternatives

The complexity of the Chaloupka simulation model justifies making the tool itself and the results it produces objects of study. The tool, as described in the User's Guide allows for the incorporation of 100s of parameters. There are over 150 easily modified (user-defined) parameters alone, and many more within the program itself. The model incorporates environmental and demographic stochasticity by adding temporal variability within stocks and probabilistic variability in fecundity and survival probabilities. The choice of parameter values and the perceived precision of the result encourage a very mechanistic approach to management of these species. When results about the numbers of individuals that might be impacted are derived from very small changes in parameter values, the tendency to trust these numbers as reflective of more than a general trend may arise. An analogy familiar to biologists would be the tendency of novice data analysts to run through a list of available models in a Windows-based statistical program until the "p-value" generated for the data was <0.05.

For example, as a tool for monitoring and detecting current change to support management decisions that seek to optimize human activity while maintaining sea turtle populations, the simulation model may produce results from the model that are perceived to be precise and encourage generation of targets that are perceived to be equally as precise. In reality, given the amount of information that remains unknown about this species, such narrow targets may be inappropriate. At minimum, users of this model and its results must be aware that any statement such as "Given these risk criteria, it is apparent that the Western Pacific leatherback sea turtle stock would most likely be well on the way to extinction given harvesting of 15%of adults each year for 100 years or more" (Users Guide, p.34) *must* be prefaced with, "The model results, given user-defined parameter set A, are..." because *any* result will depend on that parameter set (as well as those parameters set by Dr. Chaloupka in the program code). Parameter sets should be included explicitly wherever quantitative results are presented, and effort should be made to assure that results are consistent over a wide range of plausible parameter sets.

As a tool for guiding restoration efforts, such as habitat restoration, where both management activities and their concomitant effects on sea turtle populations will become manifest at some future date, the Chaloupka simulation model provides a valuable structure for prioritizing and evaluating the relative impacts of alternative management strategies. However, the recovery of sea turtle populations will also be affected by influences that cannot be incorporated into the model, and the response of the population to a management strategy may be entirely unexpected. This is a problem with any model, and is one of the reasons why continuous updates and modifications are needed as new data are acquired. Even if the population behaves exactly as theorized in the model, the environment may vary in such a way to make predictions inaccurate. It is important to monitor the environment as the context for the simulation model, and to incorporate them expeditiously, as forecasts may be affected. Assessment of prediction risks should be integrated into any management decision taken.

The development of new and sophisticated analytical tools has been anticipated for many years. Walters and Holling (1990) cautioned that even with the advent of these tools, "we must not pretend that process research and diligent data analysis alone will provide answers that resource managers can trust." Likewise, in his seminal paper on the future of conservation biology, Caughley (1994) stressed that over-parameterized models for poorly-known species give is a "false sense of precision" that may overstep our understanding of how populations respond to perturbations.

5. Suggested Protocol for Use of the Chaloupka Model to Assess the Impacts of Fisheries

Levins (1995) admonishes us to prepare for surprise. Holling (1995) cautions: "... knowledge of the system we deal with is always incomplete. Surprise is inevitable. Not only is the science incomplete, but the system itself is a moving target, evolving because of the impact of management and the progressive expansion of the scale of human influences on the planet."

We recommend that:

- management decisions be supported by several modeling approaches whenever possible, including approaches that reflect coarser granularity (i.e., Levins's third modeling strategy, maximize generality and realism at the expense of precision), such as life history perturbation analysis models (Heppell et al. 2000), qualitative community models (Puccia and Levins 1985) and coarser-grained temporal stochastic models.

- prediction risks (risks of incorrect predictions) should be assessed when comparing management alternatives.
- multiple reference points should be evaluated for each set of alternative scenarios, including short-and long-term reference points that may be used in decision-making.
- sensitivity of the model to reasonable changes in parameter inputs should be carefully examined. These are discussed in greater detail in (6.) following.
- clusters of model inputs should be evaluated as variables themselves. Conflicting models may be reducible to a small set of scenarios more amenable to other types of models.

Furthermore, we suggest that two general questions be considered each time the model is used in decision-making:
Question 1. How do two reference points (one short-term, one long-term) change with varying input parameters?
Question 2. Under what range of input conditions can we achieve the same endpoint?

We strongly advise that each user group carefully review the entire list of parameters. Only a subset of user-defined parameters is included in the default list for the slider toolbox; at minimum, the list of user-defined parameters should be understood and assessed using available data.

In a summary review of the model, Donald Kobayashi of the NMFS Honolulu Lab describes a simulation scenario for determining the effects of the Hawaiian longline fishery on loggerhead populations. The number of adult turtles in 2085 was used as a reference point, and the effects of all possible anthropogenic hazards available in the model were evaluated by removing each hazard individually and recording changes in the model endpoint. This analysis suggested that the Hawaiian fishery has little or no effect on the number of mature loggerheads in 2085, regardless of the intensity of its effect, when the other hazards were included – "a minor component of the aggregate anthropogenic hazard". While the loggerhead model is quite likely more reliable than the leatherback model, given that the range of inputs for each parameter is grounded in empirical data, this analysis is insufficient for policy guidance. The analysis was only performed with one set of possible input parameters, and should be repeated for a range of potential input sets. Without such an analysis, far too much faith is put on the user's available input data and the preset "defaults" of the model, and there is no guarantee that a different group of model users would come to the same conclusions.

6. Sensitivity Analysis of Model Parameters on Critical Endpoints

Sensitivity analyses will inform analysts and decision makers about inherent variability in model results. The goal of such an analysis can be to locate:

1. A single point that produces an inconsistent or implausible result;
2. Proximal points that produce very different results.

The systematic study of the Chaloupka model by changing one parameter at a time is physically impractical due to large number of parameters that may be included in any one model. A factorial analysis of the changes in model output following systematic changes in multiple parameters is out of the question. Nevertheless, we encourage the use of sensitivity analyses to assure that results are robust to a wide range of plausible inputs.

Sensitivity of clusters of parameter values could be accomplished by construction of a meta-model of model accuracy, a knowledge-engineered model that describes parameter regions that produce valid results and that can be used to identify plausible and implausible clusters of parameter values. Pearl (1988) has referred to this approach of representing confidence in terms of higher-order probabilities as 'probabilities of probabilities.' In this approach, the event $P(A) = p$ is a random variable that depends on the occurrence or nonoccurrence of some other event in the model. For example, the parameter settings for the scenario shown in Figure 3 of the User's Manual (at risk group adults; constant annual harvest rate = 0.15; harvest duration = 100 years starting from 1975) might fall in one cluster of input parameters. Annual harvest rates of 0.2 or greater might be deemed speculative by expert analysts, reducing confidence in that instantiation of the model. Probabilities regarding confidence can be assembled into a Bayesian belief network (BBN) that, like the simulation model, serves as a framework within which to relate multiple observations and hypotheses. As new evidence about the biology, ecology, environment and anthropogenic hazards to sea turtle species is discovered, it can be applied incrementally to the BBN of model confidence.

Such a model would be an attempt to circumscribe the state of knowledge in the field. It defines regions of parameter space where experts feel comfortable and can express their level of comfort with their belief that what is known accurately represents natural processes. It incorporates ranges of parameter values with which we have experience, ranges of values with which we have little or no experience, but are confident that overall behaviors within this range will not change, and ranges of parameter values where there is significant uncertainty. The contribution of this model over the parameter processing provided by the Chaloupka simulation is that it can alert analysts to possible interactions among certain parameter values that can produce non-additive effects. These synergistic effects are potentially dangerous because they can precipitate rapid population declines. Construction of the belief model is a nontrivial task, but it may be a potentially valuable complement to the Chaloupka simulation model.

The Chaloupka model simulation is a sensitive tool for predicting trends in the present and future. It is important to remember, however, that any result produced by the model will be mediated by current environmental conditions and anthropogenic hazards. The granularity of proposed management actions must acknowledge the uncertainty contained in the model and in the surrounding context of the model. One strategy to monitor context would be to identify sentinel variables (observable indicators of trends that are sensitive to parameter settings) to determine when ranges of parameter values can be applied to produce results consistent with real-world conditions.

7. Conclusions and Research Recommendations

The Chaloupka simulation model has the potential to serve as a key component in a suite of analytical tools to assess the viability of sea turtle populations. The interactive method by which human analysts can adapt the model to specific scenarios increases its potential for extended use. While the simulation models the behavior of sea turtle populations, the complexity of the model is sufficient motivation for the study of the behavior of the model itself. We recommend that such research address the following tasks:

- Perform external validation to search for points that validate and define the boundaries and validity of the Chaloupka model.
- Identify sentinel variables (observable indicators of trends that are sensitive to parameter settings) to determine when the model can be applied and produce results consistent with real-world conditions.
- Perform prediction risk to determine consequences of incorrect model results.
- Construct a meta-model of model accuracy: Construct a knowledge-engineered model (a Bayesian network) that describes parameter regions that produce valid results. Use to identify plausible and implausible regions of parameter values, singly and in combination.

Dr. Chaloupka himself has stressed that the primary use of this model is to learn how sea turtle populations might behave under various scenarios – heuristic, rather than predictive, analyses. The User Guide for the Dermochelys Model runs through a variety of examples in its tutorial that emphasize the heuristic nature of the simulation. The primary task for users of this model is to determine how robust the various model endpoints are to a wide range of user inputs and parameter uncertainty. It is perhaps unsurprising that, relative to the many potential sources of mortality experienced by sea turtles, the small fleet of Hawaiian longliners are not to blame for population declines. At a meeting in Honolulu in 1995, many different models, both simple and complex, came to the same conclusion (Bolten et al. 1996). Thus, this result IS likely to be robust to many model formulations, but others will not be.

8. References

Berke RL. Volleys of data replace blatant attacks of 1988. *The New York Times*, October 29, 1992.

Bolten, A. B., J. A. Wetherall, G. H. Balazs and S. G. Pooley. 1996. Status of Marine Turtles in the Pacific Ocean Relevant to Incidental Take in the Hawaii-based Longline Fishery. U.S. Department of Commerce, NOAA Tech. Memo NOAA-TM-NMFS-SWFSC-230, 167 pp.

Caughley, G. 1994. Directions in conservation biology. Journal of Animal Ecology 63:215-244.

Crouse, D. T., L. B. Crowder, and H. Caswell. 1987. A stage-based population model for loggerhead sea turtles and implications for conservation. Ecology 68:1412-1423.

Heppell, S. S., H. Caswell and L. B. Crowder. 2000. Life histories and elasticity patterns: perturbation analysis for species with minimal demographic data. *Ecology* 81(3):654-665.

Heppell, S. S., D. T. Crouse, and L. B. Crowder. 1996. A model evaluation of headstarting as a management tool for long-lived turtles. Ecological Applications 6:556-565.

Holling CS. 1995. What barriers? What bridges? In: Gunderson LH, Holling CS, Light SS, eds. Barriers and Bridges to the Renewal of Ecosystems and Institutions. Columbia University Press, New York.

Levins R. 1966. The strategy of model building in population biology. *American Scientist* 54: 421-431.

Levins R. 1995. Preparing for uncertainty. *Ecosystem Health* 1:47-57.

Glover SM, Prawitt DF, Spiker BC. The influence of decision aids on user behavior: implications for knowledge acquisition and inappropriate reliance. *Organizational Behavior and Human Decision Processes* 1997; 72 (2): 232-255.

Pearl J. 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, San Mateo CA.

Primack RB. 1993. Essentials of Conservation Biology. Sinauer Associates, Sunderland MA.

Puccia C and Levins R. 1985. Qualitative Modeling of Complex Systems: an Introduction to Loop Analysis and Time Averaging. Harvard University Press, Cambridge MA.

Schmitt JF and Klein G. 1996. Fighting in the fog: Dealing with battlefield uncertainty. *Marine Corps Gazette* 80:62-69.

Shenk D. Data smog: surviving the information glut. 1997. HarperEdge, San Francisco, CA.

Walters CJ and Holling CS. 1990. Large-scale management experiments and learning by doing. *Ecology* 71:2060-2068.