Report of David Sampson's Activities as a Visiting Scientist with CAPAM (Center for the Advancement of Population Assessment Methodology) for the period 1 April 2014 through 30 June 2014

Dr. David Sampson, Professor of Fisheries from Oregon State University's (OSU) Coastal Oregon Marine Experiment Station and Department of Fisheries and Wildlife, worked with CAPAM collaborators from the Inter-American Tropical Tuna Commission (IATTC) and the National Marine Fisheries Service (NMFS) for three weeks during April and May 2014, with additional follow-up work conducted during June and beyond. The research project is focused on the subject of spatial structure in a fishery causing distortions to fishery selection (e.g., Sampson and Scott, 2011). During his stay at CAPAM Dr. Sampson set up a simulation experiment that he is using to explore the consequences of spatial structure for stock assessments conducted with Stock Synthesis. This report describes Dr. Sampson's activities during the contract period.

Project Activities while Visiting CAPAM in La Jolla

During his visit to CAPAM, which began on 21 April 2014, Dr. Sampson was provided an office at the NMFS Southwest Fisheries Science Center (SWFSC) in La Jolla, CA. He was also given user-privileges to remotely access a fast computer located in the nearby Inter-American Tropical Tuna Commission (IATTC) facilities. Dr. Sampson's activities while in residence focused on designing, constructing and testing procedures for conducting simulation tests of different approaches for representing spatial structure in age-structured stock assessments conducted using Stock Synthesis (Methot and Wetzel, 2010). His initial work involved modifying an existing operating model that simulates the population dynamics of a fish stock and generates data files that can be read directly by the Stock Synthesis software.

The operating model, constructed in Excel, simulates a fish stock that occupies three spatial regions whose age-structured sub-populations grow and die independently of one another and without interchange of fish among the regions. The sub-populations are linked by their contribution to a spawning biomass pool and shared annual recruitment, which is lognormally distributed with expectation given by a Beverton and Holt spawner-recruit relationship. Recruitment is apportioned randomly to the regions, with annual deviations averaging about 5% from the expected values. Although the fishing processes in all three regions use the same asymptotic fishery selection curve, the overall age-specific rates of fishing mortality can take on quite complicated forms, depending on the trends set for the regional rates of fishing mortality. The operating model generates lognormal survey estimates of biomass and multinomial random sample data for fishery and survey age-compositions by region. The operating model produces data files and accompanying control files that can be directly read by the Stock Synthesis software. The operating model produces sets of Synthesis files for a suite of model configurations that include: (a) a three-region model whose data and spatial structure are the same as the operating model; (b) one in which the data are treated as coming from a single region but fishery selection (one fleet only) is modeled using a flexible non-parametric form; and (c) another that takes the "areas-as-fleets" approach (Waterhouse et al., 2014), with three regional fleets that all operate on a single stock.

In addition to producing data sets for Synthesis that include random observation errors, the operating model also generates so-called "perfect" data sets that have no observation error. If these “perfect” data are fit by a correctly specified Stock Synthesis model, the resulting Synthesis model will produce a log-likelihood value of zero. The purpose of fitting these “perfect” data sets is to verify that the population-dynamics embedded in the operating model are the same as the dynamics embedded in Stock Synthesis.

The operating model, when it produces a Synthesis control file, determines whether Synthesis should treat parameters as estimated or fixed, and the operating model provides values for those parameters. Key parameters that have a very strong influence on a stock’s dynamics (e.g., the natural mortality coefficient and the parameter for steepness of the spawner-recruit relationship) were fixed at their true values for all Stock Synthesis model configurations considered in the experiment for this project. Other important aspects that were fixed at their true values include the annual catch weights, a time-invariant vector for weight-at-age, and the growth-in-length parameters. The parameters that were estimated for all the Stock Synthesis model configurations were the recruitment deviations and the fishing mortality coefficients for the initial year. When Synthesis was applied to data that were not perfect, the operating model configured the Synthesis control files to estimate the unfished recruitment and the fishery selection parameters and annual selection parameter deviations (if any).

One initial challenge for Dr. Sampson was to construct appropriate Stock Synthesis controls for the configuration that collapsed the Synthesis data as if they were from a single region and a single fishing fleet and a single survey. In this case fishery selection must be modeled using a very flexible non-parametric form that has age-specific coefficients that also vary each year. Dr. Sampson was able to resolve the issues that arose, as evidenced by the revised operating model’s ability to generate Synthesis control files that result in perfect fits (zero log-likelihood) to perfect data for this “collapsed” configuration.

After consultation with CAPAM collaborators Drs. Mark Maunder (IATTC) and Paul Crone (NMFS, SWFSC), and some initial testing of the modified operating model, Dr. Sampson decided to structure the main simulation experiment for the project to follow a full factorial design with three main effects and two levels for each factor. One factor represents different life-history and biological traits of the fish stock (shorter-lived and more productive versus longer-lived and less productive). A second factor controls the stock’s biomass trend and catch history (decreasing from lightly fished versus increasing from heavily fished). The third factor is the amount of variability in the population-selectivity (low versus high variability, measured in terms of the average of the age-specific standard deviations of log-scale changes in the population-selection coefficients). In concert with a large set of parameter values that regulate the operating model, these three factors control the properties of the data that the operating model produces for analysis by Stock Synthesis (fishery and survey age-composition data, survey biomass indices).

Dr. Sampson constructed the operating model to produce data and control files for running Stock Synthesis models for six distinct assessment model configurations for each random replicate of the population dynamics. Differences between replicates within a given treatment are driven by randomness each year in recruitment, in the spatial distribution of the recruits, and in the regional fishing mortality coefficients. The six assessment model configurations are as follows.

1. Perfect data (no observation error) for a Synthesis model that has the same structure as the operating model (three-regions, each with its own fleet).
2. Random data (with observation error) for a Synthesis model that has the same structure as the operating model (three-regions, each with its own fleet).

3. Perfect data (no observation error) for a Synthesis model that has collapsed the spatial structure to a single region and fleet.
4. Random data (with observation error) for a Synthesis model that has collapsed the spatial structure to a single region and fleet.
5. Random data (with observation error) for a Synthesis model that has a single region but three fleets (i.e., an areas-as-fleets configuration). The fishery catch-at-age data matrices for this configuration are identical to the fishery catch-at-age data matrices used in configuration (2). Selection curves for the three fleets in the Synthesis model are set up as time-invariant double-normal curves.
6. Identical to configuration (5) except that the selection curves for the three fleets are set up as time-varying double-normal curves.

Subsequent Project Activities

Although Dr. Sampson was able to develop the operating model during his three week visit to La Jolla and use it to conduct a series of trial runs (for the purpose of debugging the software and evaluating the time required for a series of Synthesis runs), he was unable to begin the full experiment. The primary stumbling block was related to assessment model configuration (6), the areas-as-fleets configuration with time-varying double-normal fishery selection curves. At issue was how to specify values for Synthesis to use for controlling random-walk variation in the fishery selection parameters, a set of two or three parameters for each fleet's double-normal selection curve. These co-called *dev_SD* parameters, which must be provided as inputs to Synthesis, control the magnitude of the log-likelihood component associated with parameter deviations. They potentially could have a strong influence on the accuracy of Stock Synthesis estimates. Consequently, it is important that appropriate values be used for the sake of the scientific objectivity of the experiment. A similar problem arose for assessment model configurations (3) and (4), which had collapsed the spatial structure, but for those configurations Dr. Sampson was able to derive the correct values for the *dev_SD* parameters from sets of population-selection coefficients calculated within the operating model. Deriving areas-as-fleets selection coefficients within the operating model, however, is a significantly more complicated problem. During June (and after the conclusion of the contract period) Dr. Sampson continued to work on possible techniques for deriving values for the set of *dev_SD* parameters needed as inputs for assessment model configuration (6), but the problem had not been resolved at the time this report was written.

After several sets of email correspondence and a telephone conversation between Dr. Sampson and Drs. Maunder and Crone, it was agreed that the experiment could begin but using only assessment model configurations (1) through (5). While Stock Synthesis was running through those portions of the experiment, Dr. Sampson would conduct further research into approaches for deriving the *dev_SD* values needed for configuration (6). The Synthesis runs for configuration (6) would be conducted at a later time.

On 21 July Dr. Sampson began running a set of batch files on the IATTC computer. Each batch file conducts a set of Stock Synthesis runs using 20 replicate data sets with assessment model configurations (1) to (5). There are eight batch files, one for each experimental treatment, and all eight files can be run simultaneously on the IATTC computer, each in its own operating system window. To complete a set of batch runs requires more than 48 hours.

When the results from the Stock Synthesis runs are finally available, Dr. Sampson will analyze them using standard analysis of variance techniques to measure the relative importance of the three experimental factors and their interactions. The analysis will focus on the relative error of the Stock Synthesis estimates of unfished and final-year spawning biomass and their ratio (depletion), unfished recruitment, the maximum sustainable yield and its associated spawning biomass. A manuscript, providing a full description of the experiment and its results, will be written with assistance from Drs. Maunder and Crone and submitted for publication in *Fisheries Research* or another suitable fisheries science journal.

References

Methot, R.D., Wetzel, C.R., 2012. Stock synthesis: a biological and statistical frame-work for fish stock assessment and fishery management. *Fisheries Research* 142: 86–99.

Sampson, D.B. and Scott,R.D., 2011. A spatial model for fishery age-selection at the population level. *Canadian Journal of Fisheries and Aquatic Sciences* 68: 1077-1086.

Waterhouse, L. Sampson, D.B., Maunder, M. and Semmens, B.X., 2014. Using areas-as-fleets selectivity to model spatial fishing: Asymptotic curves are unlikely under equilibrium conditions. *Fisheries Research* 158: 15-25.

Prepared by David Sampson, 30 July 2014