**Overview of Pallid Sturgeon Assessment Framework Evaluation**

Redesign of the Pallid Sturgeon Population Assessment Program (PSPAP) to PSPAP v. 2.0 is intended to update population assessment to support decision making under the adaptive management of the Missouri River Recovery Program (MRRP). Notably, PSPAP v. 2.0 is considered to be *necessary* for the pallid sturgeon monitoring strategy, but it is not considered to be *sufficient* for monitoring needs. Instead, it is designed to be complementary to three other components: effectiveness monitoring, focused research studies, and the collaborative population model. These four components define the information framework needed to understand and predict population responses to management action.

1. Pallid Sturgeon Population Assessment Program v. 2.0. The fundamental objective of PSPAP v. 2.0 is to provide essential population-level information needed for the MRRP to make decisions about its fundamental objectives, including, but not limited to:
   1. Discern status and trends of the pallid sturgeon population; serve as validation of model predictions.
   2. Complement and enhance understanding of linkages from management actions to population responses. PSPAP v. 2.0 cannot provide direct linkages to all management actions but will be especially important in evaluating population responses to augmentation and stocking decisions as it may estimate abundance, survival, and effective population size.
2. Effectiveness monitoring. Each MRRP action has an associated monitoring plan that is designed to provide insights into whether the action has the intended ecological effect. These are presently being defined for flow cues, passage around Intake Dam, spawning habitat, interception-rearing complexes (IRCs), and rehabilitation of shallow-water habitat projects. These effectiveness monitoring plans will provide the links from the actions to habitat changes to biological responses. In some cases, the biological responses will be sufficient to quantify population-level effects. For example, the spawning site assessment may document x% increase in recruitment to the drift in constructed spawning sites, and the x% increase may be used to estimate changes in corresponding parameters in the population model, thereby allowing for evaluation of population-level effects. In other cases, measured biological responses may be only indirect indicators of population-level effects. For example, IRC monitoring may demonstrate percentage increase in interception, but fail to link interception to growth and survival parameters. In such cases, focused research studies and/or population assessment will be needed to link to the population-level effect.
3. Focused research studies. The level 1 and level 2 focused research studies described in appendix C of the Missouri River Science and Adaptive Management Plan (MRSAMP) are designed to provide fundamental understanding of pallid sturgeon ecology in the Missouri River and to develop quantitative response models. For example, level 1 mesocosm studies on foraging bioenergetics of age-0 sturgeon are meant to translate changes in habitat characteristics in IRCs (depths, velocities, and bottom conditions) into changes in growth and survival. Other level 1 studies are intended to develop technologies that can be used to measure responses, such as improvements in telemetry systems and direct measurements of habitat conditions and egg survival.
4. Collaborative Population Model. The Collaborative Population Model documented in Jacobson and others (2016) serves as the framework to integrate understanding from effectiveness monitoring, focused research, and PSPAP v. 2.0. Changes in model parameter values associated with actions (for example, increases in age-0 survival associated with IRCs or increases in viable gametes associated with flow cues or spawning habitats) will be incorporated into the model to provide predictive understanding (and uncertainties) of population-level responses associated with management action. PSPAP v. 2.0 may provide additional parameter values (for example gender ratios, fecundity, age at first reproduction, and recrudescent intervals) and, importantly, the empirical data on population status and trends needed to validate population model results conditional on cost constraints.

The remainder of this paper focuses on the design process elements for PSPAP v. 2.0. It should be noted that this redesign process and possible changes affect many stakeholders. Stakeholders include state and federal agencies contracted to collect data, consulting groups working on the adaptive management plan, and the USACE who provides the resources to conduct the assessment. Therefore we are using a structured decision making process to provide a transparent and rigorous approach to evaluate alternative monitoring designs in the context of stakeholder objectives while accounting for uncertainty (Conroy and Peterson 2013). This process began in March 2017 and is expected to extend through early spring 2019, according to this schedule:

* March 21, 2017: Workshop developing objective hierarchy and elicit logistical sampling information from stakeholders
* Spring 2017 – Fall 2017: Model sampling scenarios – explore approaches to meeting objectives, mix of methods, benefit:cost of objectives. Communicate progress through PSPAP v. 2.0 blog.
* August 31, 2017: Webinar to document progress, elicit additional information on objectives hierarchy and sampling logistics.
* October 2017: White paper for ISAP review.
* Late Fall 2017: Meeting to present design results to agencies, stakeholders.
* December 2017: Deliver draft PSPAP v. 2.0 sample-design report.
* Winter 2017-2018: Agency review, ISAP review, and revisions
* Spring 2018: implement the design on a pilot basis.
* Spring – Fall 2018: revise, refine models and protocol. Finalize design report.
* Spring 2019: Implement PSPAP v. 2.0.

**Background of PSPAP v. 2.0 Design**

The need to redesign the PSPAP was triggered by the recognition that the current PSPAP may not allow evaluation of whether pallid sturgeon fundamental objectives identified in the AM plan were achieved on an annual basis or estimated with any level of certainty. Specifically, sub-objectives listed in section 4.1.1. of the AM plan specify 1) increase pallid sturgeon recruitment to age-1, and 2) maintain or increase numbers of pallid sturgeon as an interim measure until sufficient and sustained natural recruitment occurs, both of which are needed to achieve the fundamental objective set by the USFWS to preclude species jeopardy. These two sub-objectives are redefined in the context of the monitoring program as fundamental objectives to 1) quantify recruitment to age-1 and 2) quantify pallid sturgeon population trend and abundance. To achieve the fundamental objective(s) identified for a population assessment program, it is important to identify sampling strategies (and associated abundance and trend estimators) that will give optimal estimates given budget constraints. Hence, it is pertinent to compare metrics of estimator success and associated sampling costs across multiple estimators and sampling strategies while accounting for uncertainties and whether monitoring program objectives were achieved. The following sections outline the approach used for this comparison. The methods described below are overviews intended to provide a summary of the process but not overwhelm the document with technical details. It should also be noted that evaluating alternative monitoring designs is not a trivial task, and at times requires days of computing time to run estimators and various simulations with sufficient numbers of replications to fully characterize the potential outcomes and uncertainties.

Many state and federal agencies contracted to collect pallid sturgeon population assessment data will be affected by changes to the current PSPAP design. We held an in face stakeholder workshop during the MRNRC in March 2017 to overview the process and elicit stakeholder objectives of a population assessment program. Five fundamental objectives of the PSPAP were identified: 1) quantify recruitment to age 1, 2) quantify population trend and abundance, 3) provide collaborative population model inputs, 4) maintain compatibility with legacy PSPAP data, and 5) remain within cost constraints. Objectives were then converted to quantitative indices. For example, bias, precision, and performance metrics are used to evaluate stakeholder objectives to quantify recruitment to age 1, population trend, and abundance. Compatibility was quantified as similarity to existing PSPAP design, and providing quantitative population model inputs was measured as the proportion of population model inputs that could be estimated from the monitoring program.

Afour step process was used to generate the outcomes from monitoring designs. First simulate a reference PS population, then simulate sampling catch data for varying sampling designs and gears, then estimate abundance and trend from the catch data, and lastly calculate metrics of estimator success by comparing the estimates to the true values used to generate the reference population. Simulations were spatially explicit using river bend and segment as the spatial grain and extent, respectively. When possible the PS reference population is initialized using data from the PSPAP database; otherwise plausible values were used based off of expert opinion and the PS literature. A combination of fixed and random sampling designs can be implemented, and a spatially balanced design is being considered (Stevens and Olsen 2004). Potential sampling strategies also differ in the number of sampling occasions per year, as well as the number of gear deployments within a particular occasion. Additionally, each sampling design can be implemented with varying gear combinations: gill nets, trammel nets, otter trawls, and trotlines.

All estimators are applied to each of the catch data simulations except where sampling design limits the use of an estimator. Abundance estimates are first computed on the bend-level using various estimators: closed population and (Otis et al. 1978), Cormack-Jolly-Seber, robust design (Pollock 1982), and catch for single occasion estimates (i.e., relative abundance). Bend-level abundance estimates are then aggregated to the segment-level where estimate uncertainty is calculated using the delta method (Hilborn and Mangel 1997, Powell 2007). Trend estimates are computed as the slope of the linear model of either annual abundance estimates or annual catch per unit effort for each segment. Alternatively, in absence of recruitment, survival is the same as population trend. Additionally, estimates of other population attributes identified by stakeholders are calculated (e.g., mean length, size structure, condition).

Bias, precision, and performance are computed for each estimator-sampling design combination. The bias of an estimator used with a particular sampling design is computed as the expected value of the difference between the estimated value and the actual value, where the expectation is taken over all estimates made by the given estimator on all catch data simulated under the given sampling design. Note, actual values are known since they are reported in or can be derived from the reference population data. Precision is calculated as the ratio of the standard error to the absolute value of the estimate. Since sparse catch data will lead to errors in certain abundance estimator calculations, estimator performance was calculated as the proportion of the catch data that allowed for reliable bend abundance estimates.

Each estimator-sampling design combination is associated with a measure of bias, precision, and performance. For each of the three metrics we computed utilities as values in a common range: 0-1, with values scoring the closest to 1 giving the highest utility. The overall utility of the estimator-sampling design pair is evaluated as the weighted mean of these three utilities, where the weights are determined by the importance of each metric as established by stakeholders. Similarly, the overall utility of a suite of estimators (one abundance estimator, one trend estimator, etc.) used under the same sampling design is the weighted mean of the utilities for each estimator. For example, if stakeholders decide that abundance and trend estimates are equally important, then the utility, , of a particular sampling-estimation design would be calculated as , where and are the utilities of the chosen abundance estimator and the chosen trend estimator under the particular sampling design, respectively. Additionally, if the cost of any sampling-estimation design is greater than that allotted for in the budget, then its utility becomes zero since it is not monetarily feasible to implement such a design.

The outcomes of simulations will be used to parameterize a Bayesian Decision Network (BDN). The BDN can evaluate alternative sampling designs accounting for uncertainty. Specifically, in the stochastic simulations describe above, many parameter distributions were derived from the PSPAP database or the pallid sturgeon literature. However, mean values and standard deviations for movement probabilities, recruitment frequency, number of recruits, and gear catchability are highly uncertain and not reliably available. To account for this, reference populations and catch data are simulated for a wide range of mean values for these parameters. A sensitivity analysis will be then be used to evaluate the effect of these uncertainties on the results (Conroy and Peterson 2013). For example, it is possible that for all ranges of catchability the robust design has the most utility, and hence management decisions can be made confidently knowing that precise values of catchability will not affect the decision. However, if Strategy A is better given high catchability and Strategy B is better given low catchability, then we’ve discovered that it is important to learn more about catchability to make a smart decision.

References

Conroy, M. J., and J. T. Peterson. 2013. Decision making in natural resource management: a structured, adaptive approach. Wiley.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey.

Jacobson, R. B., M. L. Annis, M. E. Colvin, D. James, T. L. Welker, and M. J. Parsley. 2016. Missouri River *Scaphirhynchus albus* (Pallid Sturgeon) Effects Analysis—Integrative Report 2016. Scientific Investigations Report 2016-5064, U.S. Geological Survey.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical-inference from capture data on closed animal populations. Wildlife Monographs:7-135.

Pollock, K. H. 1982. A Capture-Recapture Design Robust to Unequal Probability of Capture. Page 752. The Wildlife Society.

Powell, L. A. 2007. Approximating variance of demographic parameters using the delta method: A reference for avian biologists. Condor **109**:949-954.

Stevens, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association **99**:262-278