Overview of Pallid Sturgeon Assessment Framework Evaluation

Missouri River Pallid Sturgeon Technical Team

27 January, 2018

# Evaluation context for the PSPAP V. 2.0

Redesign of the Pallid Sturgeon Population Assessment Program (PSPAP) to PSPAP v. 2.0 is intended to update population assessment to support adaptive management of the Missouri River Recovery Program (MRRP). PSPAP v. 2.0 is considered to be a necessary part of 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 (fig. 1).

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.
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 focus on the implementation monitoring (i.e., was the management action implemented as intended?) and process-based monitoring (i.e., did the management action achieve desired changes to ecological processes thought to lead to increased growth and survival?). They do not address population-level responses to the actions directly. Instead, they provide information complementary to PSPAP v2.0, the Collaborative Population Model, and focused research studies, which will be needed to link the management actions to the population-level effect. An overview of the effectiveness monitoring approach for each of the MRRP actions, with IRCs presented as a more detailed example, is provided in a companion white paper.
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 parameters 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 estimates (for example gender ratios, fecundity, age at first reproduction, and recrudescent intervals). Importantly, PSPAP v. 2.0 will provide the empirical data on population status and trends needed to validate population model results.

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, agencies contracted to collect data, consulting groups working on the adaptive management plan, and the USACE which 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 and 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: Complete 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 as fundamental objectives in the context of the monitoring program to quantify recruitment to age-1 and quantify pallid sturgeon population trend and abundance.

The first sub-objective (increase pallid sturgeon recruitment to age-1) is based on the understanding that bottlenecks for pallid sturgeon populations are likely in age-0 survival and that after the first year of life, survival increases substantially (DeLonay et al. 2016). Metrics to assess achievement of this objective are particularly challenging because of the rarity of the species and existing technical difficulties with capturing age-1 (and age-2, age-3) fish and in accurately determining ages. As direct assessment of recruitment to age-1 is unlikely to be successful, the PSPAP v. 2.0 process has proceeded with the assumption that recruitment will need to be assessed through estimates of abundance and trend of older fish, determination of hatchery or wild origin, and back-calculation of age-1 recruitment rates through a well-calibrated and validated population model. Success of the second sub-objective (maintain or increase numbers of pallid sturgeon as an interim measure) can be assessed through direct empirical abundance and trend data, but the assessment can be enhanced, and better linked to management actions, through incorporation of monitoring data into the collaborative population model.

# Redesign objectives

The objective of the redesign effort is to identify sampling designs and abundance and trend estimators that will give optimal estimates given budget constraints. Our approach is to compare metrics of estimator success and associated sampling costs across multiple estimators and sampling strategies, while accounting for uncertainties about 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 sufficient understanding of the process but not overwhelm the document with technical details. It should also be noted that evaluating alternative monitoring designs is not trivial, 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.

# PSPAP stakeholders and eliciting monitoring objectives

Many state and federal agencies contracted to collect pallid sturgeon population assessment data will be affected by changes to the current PSPAP design, as well as consulting groups and management agencies dependent on monitoring data to inform adaptive management. We convened a face-to-face stakeholder workshop during the MRNRC meeting in March 2017 to present an overview of the redesign process and to elicit stakeholder[[1]](#footnote-25) objectives for the program. Five fundamental information objectives of the PSPAP were identified at the workshop: 1) quantify recruitment to age-1, 2) quantify population trend and abundance, 3) provide collaborative population model inputs (improved parameter estimates), 4) maintain compatibility with legacy PSPAP data, and 5) remain within cost constraints. Stakeholders identified many means objectives that potentially contribute to the fundamental objectives; they also provided critical information on logistics of sampling. Means objectives included varying population metrics to monitor and approaches needed to quantify the metrics. Metrics identified during the objectives elicitation were organized to 8 categories. Specifically, stakeholders identified metrics relating to population structure (e.g., age and size structure, sex ratio), reproductive status (e.g., fecundity, reproductive cycling, size at sexual maturity), health status (e.g., stress, condition, diet, contaminants), population augmentation, movement (i.e., spawning, seasonal), demographic rates (e.g., recruitment, survival), fish community (e.g., competition, invasive species), and genetic status (e.g., effective population size, hybridization, local adaptation). Stakeholder objectives were organized in an influence diagram during the workshop. The influence diagram serves 2 purposes. First, it clarifies stakeholders' fundamental and means objectives, thereby increasing transparency. Second, the influence diagram can be developed into a Bayesian Decision Network (BDN) to evaluate alternative PSPAP designs (Marcot et al. 2001, Nyberg et al. 2006, Conroy and Peterson 2013). The development and parameterization of the BDN required 2 steps: 1) translate nodes in the influence diagram to quantifiable metrics that can be estimated from a monitoring program and 2) stochastically simulate the range of possible outcomes for a node, conditional on influencing nodes. This work is continuing and the methods below provide a limited overview of the process.

## 1) Translate nodes in influence diagram to quantifiable metrics

The influence diagram was modified to develop a BDN. Specifically, nodes in the influence diagram were classified as nature, decision, or utility nodes. Nature nodes represent quantifiable metrics and the range of possible outcomes quantified as probabilities. For example, it is uncertain what level of recruitment might occur in the system, which in turn may influence the optimal monitoring design. Therefore recruitment uncertainty can be included by including a range of plausible recruitment levels and frequencies. Additionally, other important variables were added to the BDN – although not identified during the objectives elicitation – if they were considered to influence whether or not the objectives could be achieved. For example, catch depends on gear-specific catchability which is uncertain but can be bounded to reasonable levels from estimates of capture probability.

The fundamental objectives were developed as utility functions that value the outcome of a monitoring design. For example, the fundamental objective to quantify population abundance and trend was quantified by calculating the bias, precision, and performance of alternative PSPAP designs. 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. Actual values can be treated as knowns since they are reported in or can be derived from a simulated reference population (described below). Precision is calculated as the expected ratio of the standard error to the absolute value of the estimate (more precise estimators have lower values of precision). Since some catch data (e.g., sparse data) will lead to errors in certain abundance estimator calculations (e.g., non-convergence), a measure of estimator performance also was calculated as the expected proportion of catch data that allows for reliable bend-level abundance estimates. Each estimator-sampling design combination is therefore 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 (i.e., proportional scoring Conroy and Peterson (2013)). The overall utility of the estimator-sampling design pair is evaluated as the weighted sum 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.) for 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, U, of a particular sampling-estimation design would be calculated as U=0.5*U\_abund+0.5*U\_trnd, where U\_abund and U\_trnd are the utilities of the chosen abundance estimator and the chosen trend estimator under the particular sampling design, respectively. These are standard approaches for valuing outcomes and relating to stakeholder objectives (Clemen et al. 2001, Conroy and Peterson 2013).

Additional valuation will be provided for the remaining fundamental objectives. Specifically, compatibility with legacy data will be quantified as the similarity in bend randomization among designs, provided collaborative population model inputs will be quantified as the proportion of model inputs estimated by design. Lastly, the cost for each sampling design will be evaluated taking several factors into account. While overhead costs are similar for all sampling designs, costs will vary among sampling designs with the number of sampling occasions within a year, the number of deployments per sampling occasion, sampling effort of each deployment, which gears are used, how much training is required for a particular gear, etc. Costs may also vary with estimators or what population characteristics are estimated (gender, age, fecundity, etc), especially if there is a need to purchase advanced equipment or hire staff with a particular set of advanced skills. For example, if an estimate requires samples to be sent out for something like blood-hormone analysis there would be an additional cost to the collection of the samples. 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 would not be monetarily feasible to implement such a design.

Potential sampling actions were collected into decision nodes. The decision nodes represent 4 classes of decisions: system-level sampling design, estimator, gear combinations, and measurements taken on individual fish. Alternative system-level sampling designs identified during the workshop and with follow-up stakeholder input are: 1) randomized selections of bends within the segment, 2) randomly selected bends within a segment initially and followed over time, and 3) a spatially balanced design (Stevens and Olsen 2004). Several estimators were identified that could potentially be used to achieve PSPAP fundamental objectives including catch-effort and capture-recapture estimators (e.g., robust design, Cormack-Jolly-Seber, Pradel, closed population estimators). Gears and gear combinations to evaluate were elicited from stakeholders and experts in the system. Lastly, the actions required to measure metrics of interest (e.g., stress, condition) were identified.

## 2) Simulate the likely outcomes for a node, conditional on influencing nodes

Simulation has been used (and is continuing) to generate the outcomes of the alternative monitoring designs. In short, this approach has four steps: simulate a known reference pallid sturgeon population, simulate sampling catch data from the reference population, estimate metrics from the catch data, and compare the estimates to the “true” values used to generate the reference population.

### 1. Simulate a reference PS population

The pallid sturgeon reference population is initialized using data from the PSPAP database and the pallid sturgeon literature. Each river bend is populated with pallid sturgeon based on expected segment[[2]](#footnote-29)-level densities, while each fish is assigned an initial length and von Bertalanffy growth parameters L\_inf and k from segment-level and basin -level distributions, respectively. Individual fish are then tracked for 10 years, recording individual survival status, bend location, and length on a yearly basis. Survival is binomially distributed with fixed parameter f, and growth is projected by individualized von Bertalanffy growth curves. Within-basin movement is based on a pallid sturgeon’s current bend location with the probability of being in a particular bend the following year increasing as distance to that bend decreases. Immigration and emigration (for example, to/from the Mississippi River) are critical unknowns that can also be evaluated within the reference population. Recruitment occurs randomly with a fixed expected frequency (e.g., every year, every 3 years), and the number of recruits, given there is recruitment, is drawn from a basin-dependent Poisson distribution. Each new recruit is tracked after being randomly assigned an age-0 location within basin, an age-0 length of 200 mm, and von Bertalanffy growth parameters. The population simulation here is generalized from the collaborative pallid sturgeon population model (Jacobson et al, 2016).

### 2. Simulate sampling catch data

For each of the simulated reference populations, various sampling designs can be implemented to obtain simulated catch data. All sampling designs include segments 2-4, 7-10, 13, and 14, and at a minimum, the number of bends sampled within a segment matches those listed in Table A1 of Welker et al. (2016). The way bends are chosen, however, may vary. In a random sampling design bends within a segment are chosen each year randomly, while in a fixed sampling design they are chosen randomly once and then fixed to be sampled each of the following years. A combination of fixed and random designs can also be implemented, and a spatially balanced design is being considered. Sampling strategies can also differ in the number of sampling occasions (times within a year that each bend is sampled), as well as the number of gear deployments within a particular occasion. Additionally, each sampling design can be implemented with varying gears: gill nets, trammel nets, otter trawls, trotlines, or a combination of these. The catch is simulated from gear-specific catchability values which were bounded to produce plausible overall capture probabilities where catchability is defined as the probability of capture 1 fish per unit effort (Hubert and Fabrizio 2007).

### 3. Estimate metrics from the catch data

Once a sampling design is selected, 10 years of occasion-level catch data are simulated for each of the selected bends. Each fish within a selected bend has a probability, P, of being captured. This occasion-level capture probability varies from occasion to occasion, as it is calculated from the individualized deployment catchability and effort values. For each deployment, effort values f are drawn from a gear- and basin-specific gamma distribution, which was fit to PSPAP effort data. Deployment catchability q, or the probability of catching a single fish with one unit of effort, is drawn from a gear-specific distribution. Deployment-specific capture probabilities, p, are calculated as p=q\*f and then aggregated to the occasion level to obtain P. When a fish is successfully caught, fish id, location (bend), timing (occasion within year) of catch, gear used, and other important individual attributes (e.g., length, weight, sex) are recorded, simulating a complete capture history at the bend level for the given sampling design.

### 4. Compare the estimates to the “true” values

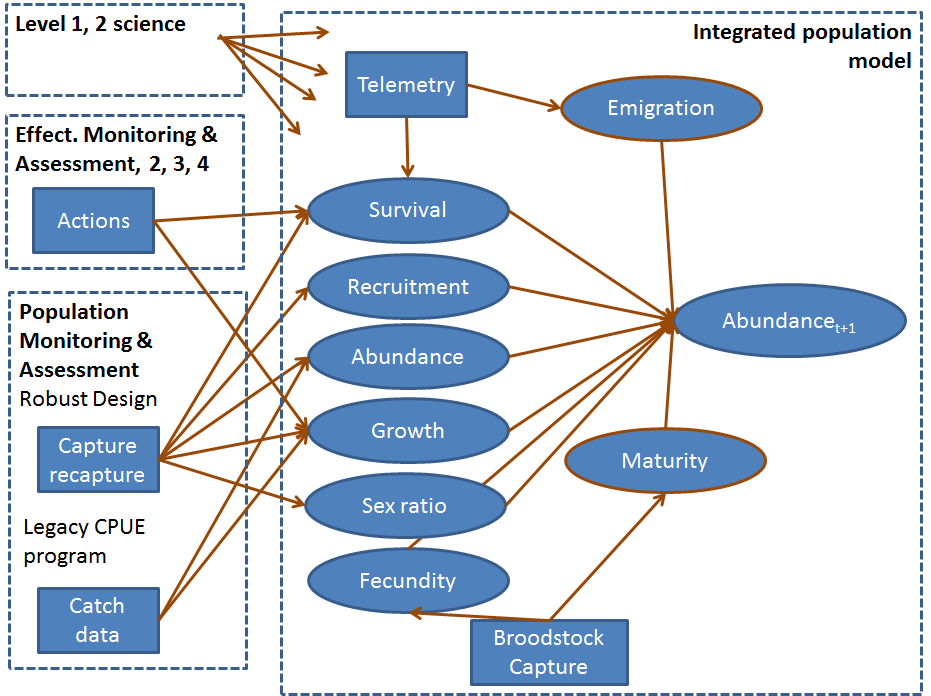
Several catch data simulations are being evaluated per reference population. All estimators are applied to each of the catch data simulations except where sampling design limits the use of an estimator, forcing the application of a smaller subset. Abundance estimates are first computed on the bend-level using various estimators: closed population M\_0 and M\_t (Otis et al. 1978), Cormack-Jolly-Seber, robust design (Pollock 1982), and catch for single-occasion estimates. Bend-level abundance estimates are then aggregated to the segment-level where estimate uncertainty is calculated using the delta method to index estimate precision (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). The precision, bias, and performance of estimates are calculated and used to parameterize the BDN.

# Additional considerations and summary

The simulation outcomes are being used to parameterize a Bayesian Decision Network (BDN). The BDN can evaluate alternative sampling designs accounting for uncertainty (i.e., the outcomes are not 100% certain). Specifically, in the stochastic simulations described 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, immigration/emigration, and gear catchability are uncertain and not reliably available but can be bound to biologically plausible values. 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 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. Results of sensitivity analyses will inform level 1 or 2 research that may be needed to reduce decision uncertainty. The process used to evaluate alternative monitoring programs is rigorous and allows transparency. Lastly, stakeholder input and objectives have been accommodated throughout the process.

##### page break

# Figures



# References

# Introduction

# Methods

## Overview

1. Trigger
2. Problem
3. Stakeholder identification
4. Objectives elicitation and structuring
5. Decision model development
6. Parameterization
7. Ancillary objectives
   1. Identification
   2. Valuation
   3. Optimization

## Eliciting stakeholder objectives

### Objectives hierarchy and attributes

The numbered objectives correspond to fundamental objectives identified during the workshop. Bulleted lists within each numbered objective in bold are measurable attributes that can be used to quantify each objective. For example there are 3 attributes under objective 1 that can be quantified for each monitoring alternative. Assuming these attributes are scaled to a common scale (e.g., 0 to 1, 0 to 100) then each bullet may receive a weight of 33% if each attribute is equally important to decision makers. Alternatively these values can be weighted to reflect perceived importance by decision makers.

1. **Quantify PS recruitment to age-1 (Natural origin)**
   * Power to detect age-1 natural origin recruits if recruitment occurs
   * Segment level age-1 abundance
     + bias
     + precision
   * Estimate age-1 recruitment rate (natural origin)
     + bias
     + precision
2. **Quantify PS population trend (natural and hatchery origin)**
   * Estimate RPMA level population growth rate
     + bias
     + precision
3. **Provide relevant PS model inputs**
   * Estimate segment-level abundance, origin and stage specific
     + bias
     + precision
     + spatial distribution
   * Survival (RPMA level)
     + bias
     + precision
   * Fecundity (RPMA level)
     + bias
     + precision
   * Growth (RPMA)
     + bias
     + precision
   * Movement
     + fidelity
     + among segment movement
   * Population structure and characteristics (segment level)
     + Size structure
       - bias
       - precision
   * Sex ratio (segment level)
     + bias
     + precision
4. **Maintain compatibility with legacy PSPAP data**
   * Proportion of randomly selected bends within segment
   * Gears similarity: proportion of standard gears used by design
   * Effort similarity: deviation from average effort
5. **Stay below cost constraints**
   * Minimize costs

# Overview

A set of reference populations of known pallid sturgeon abundance was needed to evaluate varying monitoring designs and their ability to achieve the fundamental objectives identified during the first population assessment program workshop (See Section ??). A reference population provides data on the annual survival and movement of each fish in the population (including new recruits) at the bend-level (Figure 1). Since fish length may influence sampling effectiveness (e.g. gear selectivity), data on individuals (e.g., growth) was also simulated.

### Objectives

The objectives of this analysis were to:

1. Create a set of spatially explicit pallid sturgeon population with known bend-level annual abundances,
2. Simulate reference population dynamics for 10 years with varying but known annual survival, individual growth, movement, and recruitment, and
3. Allow movement and recruitment to vary over a wide range of biologically reasonable conditions in order to better understand their relationship with monitoring design outcomes.

### Major Assumptions

1. Survival probability is homogeneous among all individuals and independent of space and time.
2. Individual fish growth follows a von Bertalanffy growth equation.
3. There is no movement within years.
4. Movement within a basin can occur between years.
5. Recruitment occurs at the basin-level and is stochastic.

Assumptions 1 and 2 are reasonable simplifications for the purposes of this evaluation and have been used in various fish studies (Cite Kirk, Holan & Wu, & Rotella). Assumptions 3 and 4 match the assumptions of many closed population estimators from mark-recapture estimation to robust design estimation. Making these assumptions allows for the evaluation of such estimators under the best possible movement conditions---if an estimator performs poorly under the best movement conditions, then its use in less favorable movement conditions is further questionable. However, since it is unclear how closely assumptions 3 and 4 are met in reality, we explore and discuss the consequences of breaking these movement assumptions on estimator outputs in Appendix ??. Lastly, since recruitment is expected to vary from year to year and fish are expected to migrate to spawning habitats within basin (and not necessarily stay in particular bends), assumption 5 is appropriate for studying pallid sturgeon.

# Methods

### Spatial extent and grain

This analysis was constrained to generate reference pallid sturgeon populations for segments and bends within the upper basin (segments 1-4) and lower basin (segments 7-10, 13 and 14) of the Missouri River. Bends of varying length were used to generate the spatial distribution of pallid sturgeon (Table 1).

### Population Initialization

The pallid sturgeon reference population was initialized using data from the PSPAP database and the pallid sturgeon literature. Population initialization required xx steps: 1. initializing bend-level abundance, 2. initializing lengths and growth parameters of individuals within each bend, 3. assigning demographic rates (i.e., sex, origin).

#### Bend Abundance

We populated each river bend with pallid sturgeon based on expected segment-level densities by origin (Table 2). The number of pallid sturgeon assigned to a bend was stochastically generated from a Poisson distribution given the segment density and bend length as:

where

* is the number of pallid sturgeon within each bend,
* is the segment and origin specific density in fish per RKM,
* is bend length in RKM,
* indexes segment within universe,
* indexes bend within segment, and
* indexes whether pallid sturgeon were hatchery or natural origin.

Generating bend abundances from a Poisson distribution allows for variation in bend densities among bends within segment while guaranteeing that the bend is populated with a whole number of fish. Additionally, one could derive a spatial Poisson distribution with the following assumptions:

1. The probability of 1 pallid being in a column of water with length and the width of the river at RKM is .
2. The probability of 2 or more pallid sturgeon in a column of water with length is negligible .
3. The number of pallid sturgeon in disjoint (non-overlapping) water columns and are independent. In other words the presence or absence of a fish in water column 1 does not effect the probability of a fish being present in water column 2, even if these two water columns are adjacent to each other.

#### Assigning length and growth parameters to individuals

We randomly assigned an initial length to each individual fish. Initial lengths were generated from a segment specific distribution constructed from recent (2015-??) PSPAP database length data. Only pallid sturgeon data where both length and weight were reported as non-negative values were used. Additionally, data points that were residual outliers to the log-log regression of weight and length were excluded from the analysis. Lastly, only lengths greater than 200mm were included in the analysis. Lengths 200mm or less were considered to be from age-0 pallid sturgeon and these fish would instead be added to the reference population through an analysis of next year's recruitment (described in a later section below).

The cleaned up data was then analyzed by segment. For each segment, a numerical inverse cummulative distribution function for length was generated by interpolating between the percentiles of the length data for the given segment. This was achieved using "approxfun" in R's stats package. Inverse cummulative distribution functions take a probability value between 0 and 1 as input and output a length value. Therefore, we were able to generate an individual's initial length by using a uniformly sampled probability value as input to the inverse cummulative distribution function associated with the individual's segment.

Additionally, we assigned individual von Bertalanffy growth parameters and , or the asymptotic length and Brody growth coefficient of a fish's growth trajectory, respectively, to each fish. Growth parameters were generated from a basin specific bivariate normal distribution fitted to the length data in the PSPAP database. (FURTHER DESCRIPTION HERE; NEED MIKE TO DO SINCE NOT SURE HOW DONE)

To avoid unrealistic growth parameter values we truncated each bivariate normal distribution (one for each basin) to the middle 80%. All points of the truncated distribution are contained on or within an ellipse centered at , the basin specific mean values for and . Since the R package truncates multivariate normal distributions to rectangular regions, as opposed to elliptical regions, we used a different approach to randomly draw points from the desired elliptical region.

In short, we first drew points from the middle 80% of a standard bivariate normal distribution and then transformed these points to the bivariate normal distribution of interest, giving us randomly drawn points that fall on or within the 80% ellipse. The truncation process was achieved using twice in R, once for the normal distribution associated with and then again for the normal distribution associated with conditional on . For more details see Appendix ??.

### Simulating Annual Transitions

After initializing the reference population, we track all individual fish for 10 years, recording individual survival status, bend location, and length on a yearly basis. Additionally, new recruits may be added to the population and tracked.

#### Survival & Growth

Survival is stochastic but simulated with a survival probability that is homogeneous among individuals with all fish in the river having an equal probability, , of surviving the year. Survival probability is independent of individual age, length, sex, origin, location, and year. In other words, each year individual fish survival is a Bernoulli trial with fixed probability of success, .

Annual growth is projected by individual von Bertalanffy growth curves as

where is the individual's length during year and each fish has its own randomly generated asymptotic length, , and Brody growth rate, , as described in the previous section on initializing the population.

#### Movement

Within basin movement occurs between years and is considered at the bend-level. Pallid sturgeon may move from one bend to another bend within the same segment, from one bend to another bend within a different segment (but same basin), or stay within the same bend. Movement probabilities are based on current bend locations with the probability of being in bend the following year increasing as the distance between the fish's current bend and bend decreases. In particular,

where , , and are particular bends, is the individual's bend location in year , is the distance in river kilometers (RKM) from the center of bend to the center of bend , and is a basin specific parameter[[3]](#footnote-54).

Currently, we have simulated reference populations that do not move (individuals can always be found in the same bend from one sampling season to the next), as well as reference populations with very little between year movement (; high site fidelity). This allows for an evaluation of whether there are any stark discontinuities in optimal monitoring decisions between the simplified case of no movement and the case of very little between year movement. Additionally, to account for the uncertain nature of movement probabilities, we plan to vary from simulation to simulation, allowing the analysis of populations that range in having high site fidelity to low site fidelity (Figure 2). Since (3) is discontinuous in terms of river kilometer, a similar bend-level conditional probability distribution that is derived from movement probabilities that are continuous in space is also being considered as a movement model.

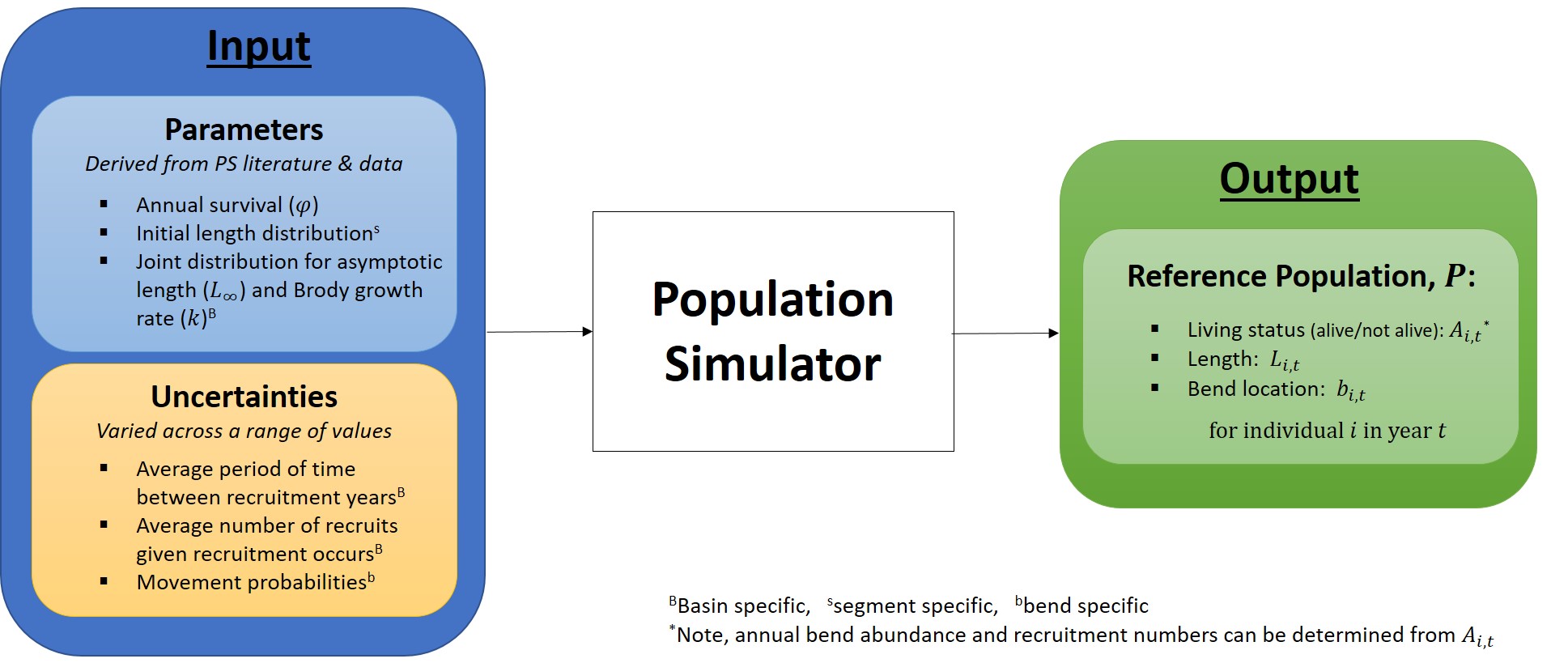
#### Recruitment

Each year recruitment was determined by two factors: (1) whether or not spawning that led to successful recruitment occurred and (2) given recruitment occurred, how many age-0 fish survived to age-1 (number of recruits). Recruitment years occurred with a fixed expected frequency (e.g. once every 3 years) and were determined each year as the result of a Bernoulli trial (e.g. probability of success 1/3). During years when recruitment occurred, the number of recruits is drawn from a basin specific Poisson distribution. Assuming that spawning results in a large number of free embryos and that survival to age-1 is rare, then a Poisson distribution mathematically provides a good approximation for the number of recruits. Furthermore, the use of a Poisson distribution aligns well with how actual recruit data may be analyzed (?), as well as with the set-up of the collaborative population model.

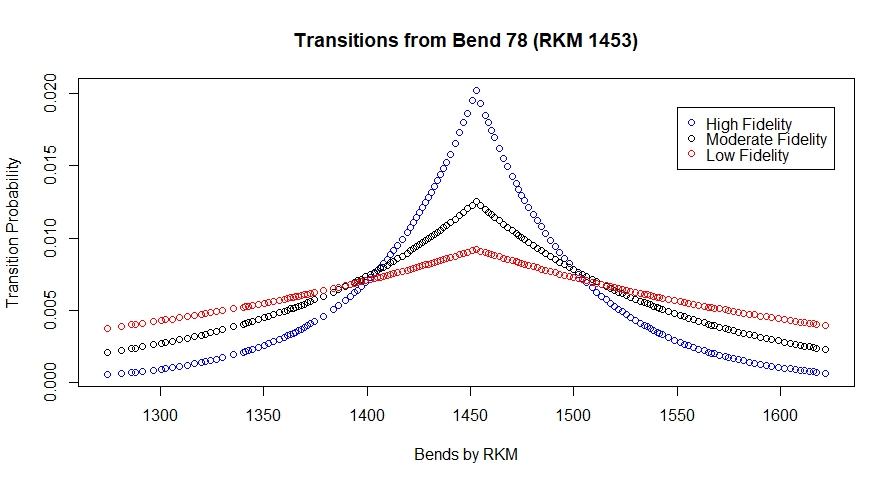
Each new recruit is tracked after being randomly assigned an age-0 location within basin, an age-0 length of 200mm, and von Bertalanffy growth parameters. Age-0 bend locations are generated from a discrete uniform distribution that includes all bends within basin, while growth parameters are generated from the same basin specific truncated bivariate normal distribution used during the initialization of the population. Initial length was chosen as 200mm, because (REASON HERE). Despite fixed initial lengths, the growth trajectories of recruits will differ due to variation in their individual growth parameters.

##### Page Break

# Figures



##### Page Break



##### Page Break

# Tables

Table 1. Summary of bends within PSPAP Missouri river segments.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Basin | Segment | Number of bends | Minimum length (km) | Mean length (km) | Maximum length (km) |
| UB | 1 | 1 | 7.89 | 7.9 | 7.9 |
|  | 2 | 40 | 0.64 | 2.2 | 3.9 |
|  | 3 | 91 | 0.64 | 2.0 | 7.1 |
|  | 4 | 24 | 0.97 | 3.3 | 8.0 |
| LB | 7 | 34 | 0.16 | 2.6 | 7.9 |
|  | 8 | 61 | 1.13 | 4.0 | 9.2 |
|  | 9 | 80 | 1.13 | 4.4 | 11.7 |
|  | 10 | 39 | 1.61 | 4.7 | 10.3 |
|  | 13 | 45 | 1.77 | 4.1 | 10.6 |
|  | 14 | 56 | 1.45 | 3.9 | 19.0 |

##### Page Break

Table 2. Segment and river level densities by origin: hatchery (H) and natural/wild (W). Minimum and maximum densities were taken from data across a few recent years (when available), while mean densities are those reported in the literature from the most recent year's data.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Basin | Segments | Fish Type | Minimum Density (fish/rkm) | Maximum Density (fish/rkm) | Mean Density (fish/rkm) | Most Recent Year | References |
| UB | 1-4 | H | 91.57 | 91.57 | 91.57 | 2013 | [1] |
|  |  | W | 0.28 | 0.37 | 0.28 | 2008\* | [2]-[4] |
| LB | 7-9 | H | 28.62 | 32.30 | 32.30 | 2010 | [5] |
|  |  | W | 5.43 | 8.88 | 5.70 | 2010 | [5] |
|  | 10, 13, 14 | H | 5.53 | 10.17 | 5.53 | 2013 | [6] |
|  |  | W | 0.56 | 0.93 | 0.93 | 2013 | [6] |

\*Estimated year of data collection based on reference date.

# Overview

Simulated catch data was essential to calculating estimates of population abundance and trend using alternative estimators. It is because we could compare the estimates from simulated catch data to the underlying reference population, that we were able to evaluate estimator accuracy and precision. Therefore, simulating catch data was a key component of evaluating varying monitoring designs and their ability to achieve the fundamental objectives identified during the first population assessment program workshop (See Section ??).

The catchability and capture probability of particular gears are important elements of simulating catch data. Since these values are highly uncertain and in some case unknown, we simulate catch data over a range of plausible gear catchabilities.

### Objectives

The objectives of this analysis were to:

1. Simulate 10 years of bend sampling data (i.e. which bends are sampled during which years) for various monitoring designs,
2. Simulate 10 years of catch data for various sampling strategies using bends as the spatial scale,
3. Include information in the catch data that could be reasonably assessed by sampling crews in the field and is relevant to meeting fundamental or sub-objectives,
4. Allow gear catchability to vary over a wide range of values in order to evaluate its effect on monitoring design outcomes, and
5. Simulate several catch data tables per combination of reference population and monitoring design.

### Major Assumptions

1. All fish can be uniquely identified.
2. Fish ids are not lost or overlooked.
3. The sampling protocol is followed without exception.
4. Expected catch, can be approximated as , a function of gear catchability, , effort described as the time in minutes a gear is deployed, , and the population size, .
5. Gear deployments are independent of one another and do not interact.

Assumptions 1 and 2, while unlikely to be perfectly met, are reasonable assumptions for working with pallid sturgeon (ref) and are typical assumptions made when working with capture-recapture data (ref). For a well designed sampling protocol, assumption 3 is also a reasonable, although we recognized that it may not be perfectly met due to safety or other unforeseen issues.

Describing expected catch as in assumption 4: , where is the capture probability and , is also a commonly made assumption (ref). This equation is likely a good approximation for small effort values (); however, it is clear that as effort increases without bound, capture probability () will become greater than 1 for fixed . In other words, the model breaks down for large effort values, and therefore, we must be cautious drawing conclusions from outcomes that were generated from large effort values. To indirectly account for this we have flagged cases in our simulations where the capture probability is abnormally high.

To understand assumption 5, it is important to first describe the sampling relationship among years, occasions, deployments, and bends. In short, during each year a number of bends will be selected to be sampled. During a particular year, each selected bend will be sampled on several occasions. Occasions are short periods of time, say 24 hours, during which multiple deployments of a gear are used to sample the population throughout the bend. Using this terminology assumption 4 means both that there is a zero probability that the same fish will be caught by two separate deployments (within the same occasion), and that any particular deployment does not affect the probability that a fish is caught by any other deployment within the same sampling occasion. For some gear types these assumptions are very likely met, while with others there are some caveats that we examine in THE DISCUSSION/APPENDIX ??.

# Methods

### Monitoring Designs

For each of the simulated reference populations, various sampling decisions can be implemented to obtain simulated catch data. The choice of monitoring design determines the particular combination of decisions used in this process---system-level sampling design, within bend and year sampling design, gear selection, and measurements taken on individual fish are all prescribed by the monitoring design (Figure 1). Alternative sampling decisions of interest were identified during the PSPAP workshop at the 2017 MRNRC meeting, as well as during follow-up input from stakeholders and experts in the system. Simulated catch data have been generated from the range of the identified alternative sampling decisions, with the exception that some individual measurements have yet to be implemented and the spatially balanced system-level sampling design is still a work in progress.

For each combination of reference population and monitoring design, we simulated catch data (Figure 2) using two major steps:

1. For each year, select which bends to sample.
2. For each occasion, select which individual fish (of those located in the sampled bends) are caught.

### Bend Selection

All monitoring designs considered included sampling bends located within segments 2-4, 7-10, 13, and 14. The number of bends sampled within each segment was chosen to be consistent with past sampling procedures. In particular, we chose the number of sampled bends within segement to match those given in USACE's 2017 PSPAP Guiding Document Table A1 (CITATION) and reproduced in Table 1 below. The way selected bends were generated, however, varied and was determined by the choice of system-level sampling design laid out in the monitoring design. In particular, three alternative system-level sampling designs were considered: a random design, a fixed design, and a spatially balanced design (Stevens and Olsen 2004). Under the random sampling design bends within a segment were chosen each year uniformly at random, while under the fixed sampling design they were chosen uniformly at random for the first year and then fixed to be sampled each of the following 9 years. A combination of the fixed and random designs can also be implemented,and the spatially balanced design is in the works.

While monitoring designs can also differ in decisions made about within bend and year sampling design, gear selection, and measurement choices, none of these decisions affected which bends were selected for sampling each year. They did, however, affect the individual capture histories of the catch data.

### Individual Captures

Once the bends to be sampled were selected for each of the 10 years, 10 years of catch data can be simulated. Each year, catch was simulated spatially at the bend level and temporally at the sampling occasion level. We defined a sampling occasion as a short time period (e.g. 24 hours) within a year during which the given bend was sampled. The fixed number of sampling occasions per year per sampled bend was given as part of the monitoring design of interest. In particular, we simulated catch data for monitoring designs that included 1-4 sampling occasions per year for each sampled bend. Additionally, in order to thoroughly sample a bend during a sampling occasion, several gear deployments are used to capture fish. Currently simulations all utilize 8 gear deployments per sampling occasion; however, future analysis will include catch data simulated for a range of deployment numbers.

For a particular sampled bend during a particular year, each fish located in the given bend had a probability, , of being captured during occasion . This occasion level capture probability varied from bend to bend and occasion to occasion, as it was calculated from individualized deployment catchability and effort values. For each of the 8 deployments within occasion , an effort value was generated from a gear and basin specific gamma distribution, and a catchability value was generated from a gear specific distribution. Deployment specific capture probabilities, , were calculated as , and the 8 within occasion deployment capture probabilities were summed (and bounded at 1) to obtain occasion level . Therefore, if a fish was located in the given sampled bend during the given year, its capture was a Bernoulli trial with probability of success , resulting in total occasion level catch numbers for the particular bend being binomially distributed.

#### Effort Distributions

Distributions for deployment level effort () were generated using the PSPAP database. We defined effort as the time (in minutes) that a gear was set, dragged, or pulled for during a single deployment[[4]](#footnote-72). Since effort was calculated using the start and stop times, all PSPAP data without a start or stop time was excluded from this analysis. The relevant data entries were then merged with the gear data and a stop date column was added to take into account gears that were set over night. Effort was calculated in minutes as the difference between the stop time and the start time with any discrepancies in the set date and stop date taken into account by using the strptime function in R to link the start and stop times to their corresponding set and stop dates, respectively. Before generating the effort distributions, all entries resulting in negative effort values were removed.

Gamma distributions were fit to the cleaned up PSPAP effort data for standard gear types[[5]](#footnote-73) by basin. The mean effort, standard deviation of the effort data, minimum effort, maximum effort, and median effort were also calculated for each gear by basin (Table 2). In general, gamma distributions provided a good fit to the effort data (e.g. Figure 3), but see Appendix ?? for more details.

#### Catchability Distributions

Gear specific catchability distributions were set up such that the log-odds of catchability () was normally distributed:

where and is the expected catchability of the gear. Because deployment catchability values, or the probability of catching a single fish with one unit of effort, are a source of uncertainty we allowed the choices of and to vary for among catch data simulations (CITE A TABLE OR SOMETHING THAT LISTS RANGES OF PARAMS).

#### Gear Selection and Size Selectivity

The choice of monitoring design also indicates which gears should be used in sampling. This is important as different gears will have different catch rates and associated costs of deployment. Currently, catch simulations include catch data from gill nets, trammel nets, otter trawls, trotlines, or a combination of these.

Additionally, gears may be size selective. In this case different gears will interact with the size distribution of the reference population to produce different catch data outcomes. Because gear selectivity is not precisely known, we are currently accounting for size selectivity by comparing the catch data results simulated under various selction curves. See Appendix ?? for details.

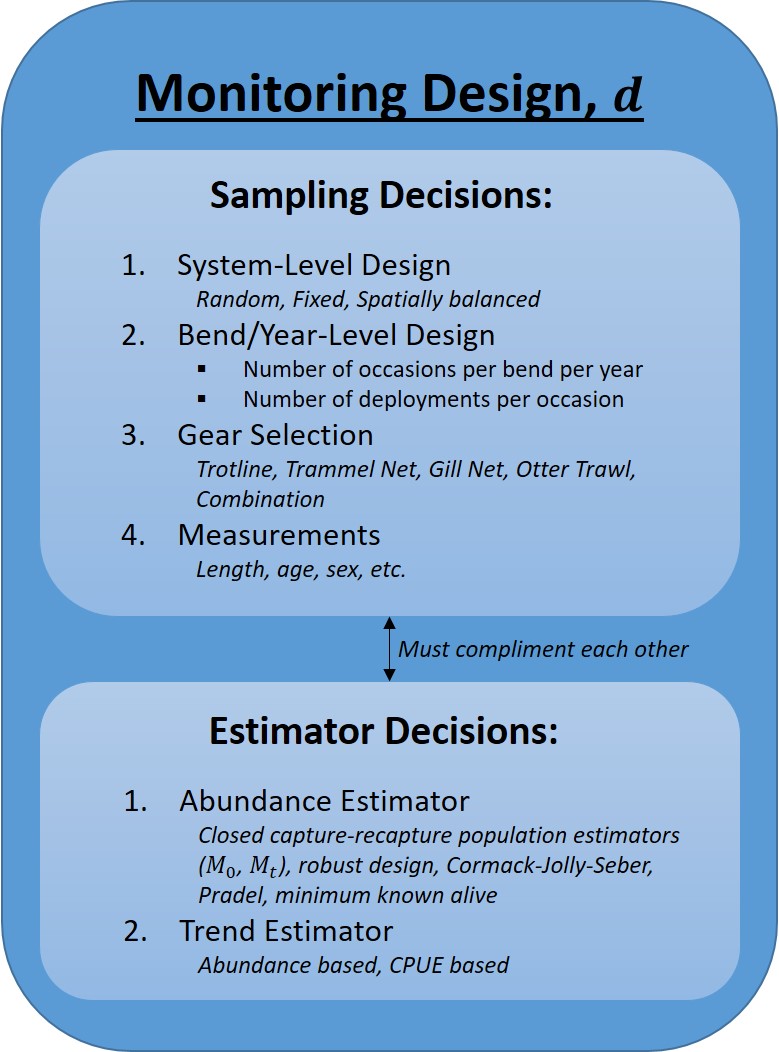
#### Individual Measurements

For a particular bend and year, a Bernoulli trial simulated whether or not an individual fish was captured during each occassion, producing a within year capture history for each fish. Fish that were never captured had their capture histories removed from the data. Fish that were captured had their individual attributes for the given year recorded from the reference population data. Specifically, whenever a fish is successfuly caught its fish id, location (bend), timing (occasion within year) of catch, gear used, and length are recorded.

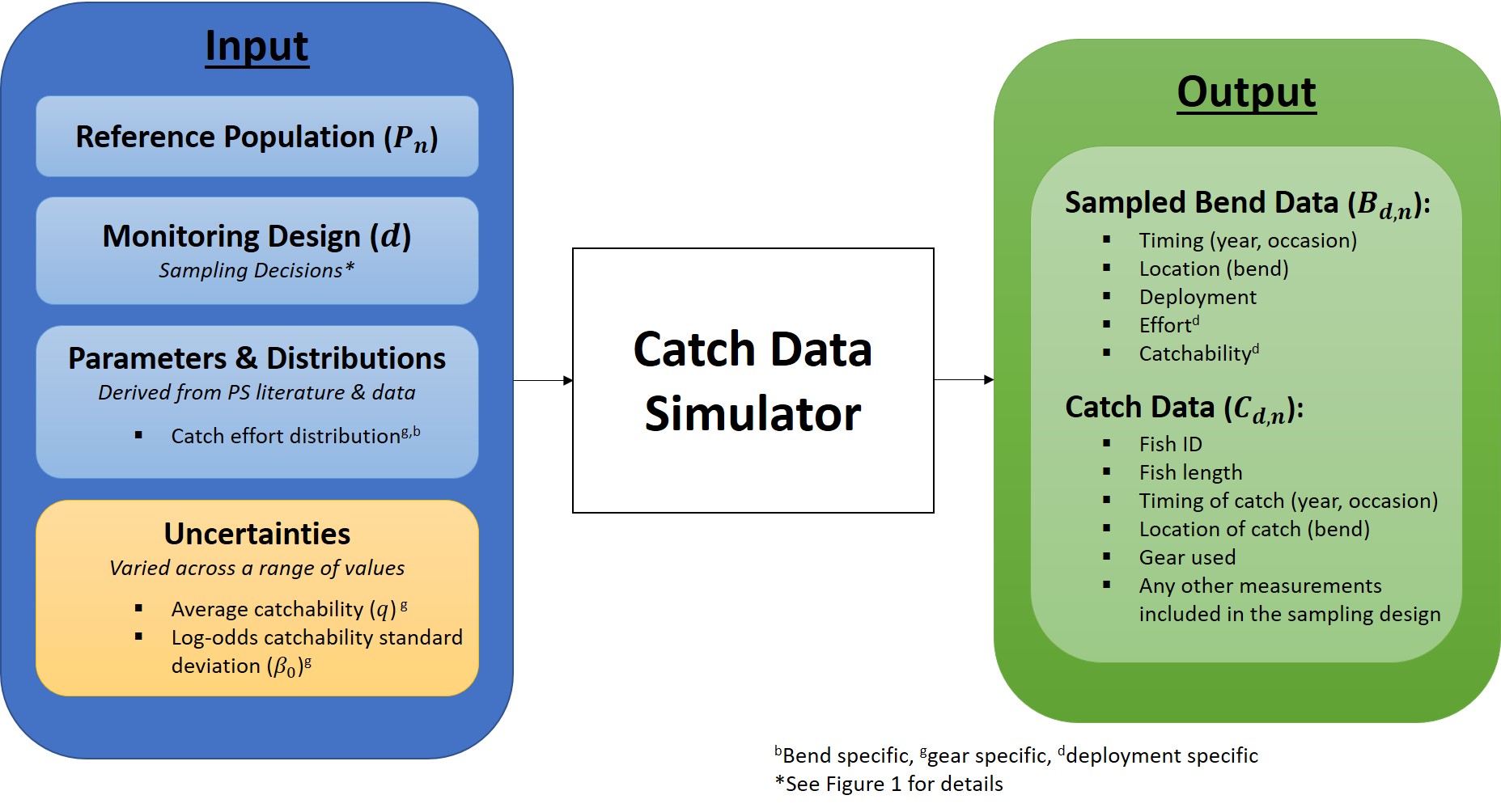
We have primarily focused on measuring length, as it is vital to the metrics for sub-objective 2[[6]](#footnote-77) as outlined in Section 4.1.1 of the 2016 AM plan. Additionally, we are currently working on including fish sex, origin (natural or hatchery), and age due to their importance to sub-objective metrics and to understanding effective populations size. Several other individual measurements were voiced to be of interest to stakeholders at the 2017 MRNRC PSPAP Workshop. These measurements may be considered as part of future analyses.

##### page break

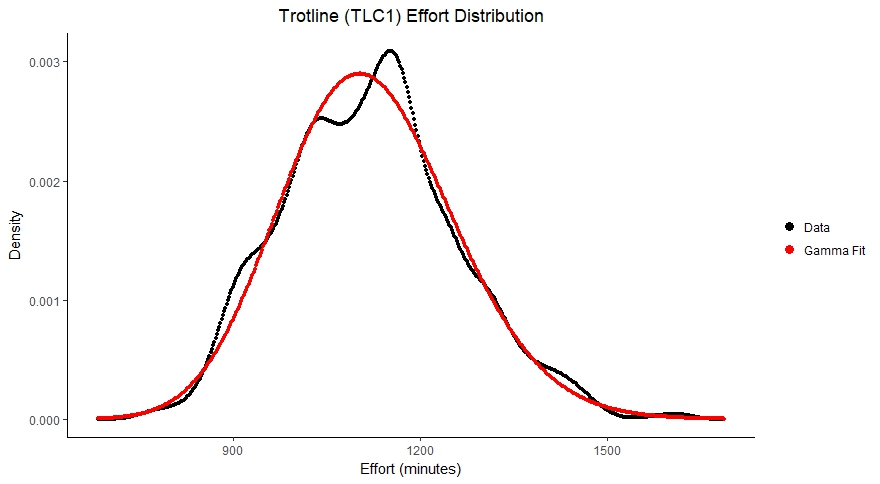
# Figures



##### page break



##### page break



##### page break

# Tables

Table 1. Number of Bends Sampled per Segment

|  |  |
| --- | --- |
| Segment | No. of Sampled Bends |
| 2 | 12 |
| 3 | 21 |
| 4 | 12 |
| 7 | 12 |
| 8 | 15 |
| 9 | 20 |
| 10 | 10 |
| 13 | 11 |
| 14 | 14 |

##### page break

Table 2. Summary of effort data by gear and basin, where effort is measured in minutes. The shape and rate columns are the results of fitting a gamma distribution to the data.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Basin | Gear | Gear ID | No. of Observations | Mean Effort | SD of Effort | Minimum Effort | Maximum Effort | Median Effort | Shape | Rate |
| LB | GN14 | 18 | 1523 | 1326 | 89 | 1002 | 1633 | 1333 | 224.493086 | 0.1692404 |
| LB | GN18 | 21 | 4698 | 1221 | 123 | 750 | 1695 | 1221 | 98.540664 | 0.0807099 |
| LB | GN41 | 18 | 1118 | 1327 | 91 | 969 | 1563 | 1340 | 210.667976 | 0.1587651 |
| LB | GN81 | 21 | 4094 | 1219 | 117 | 750 | 1678 | 1222 | 109.460899 | 0.0897863 |
| LB | MF | 41 | 7411 | 1278 | 126 | 729 | 2106 | 1288 | 103.412037 | 0.0808979 |
| LB | OT16 | 52 | 606 | 3 | 7 | 1 | 178 | 2 | 1.986635 | 0.7149057 |
| LB | TLC1 | 87 | 2164 | 1336 | 60 | 1067 | 1542 | 1341 | 500.663839 | 0.3746263 |
| LB | TLC2 | 87 | 5219 | 1206 | 109 | 859 | 1591 | 1210 | 122.773069 | 0.1017996 |
| LB | TN | 65 | 1072 | 4 | 3 | 1 | 69 | 3 | 2.568699 | 0.6765714 |
| UB | GN14 | 18 | 700 | 1130 | 183 | 310 | 1751 | 1151 | 29.735758 | 0.0263068 |
| UB | GN41 | 18 | 716 | 1134 | 183 | 295 | 1762 | 1138 | 30.634266 | 0.0270084 |
| UB | MF | 41 | 4281 | 1182 | 142 | 698 | 1708 | 1180 | 69.168422 | 0.0584990 |
| UB | OT16 | 52 | 9081 | 4 | 7 | 1 | 511 | 4 | 6.787978 | 1.6706862 |
| UB | TLC1 | 87 | 3616 | 1120 | 139 | 755 | 1615 | 1119 | 65.327401 | 0.0583306 |
| UB | TLC2 | 87 | 80 | 1025 | 150 | 694 | 1416 | 994 | 49.933263 | 0.0486958 |
| UB | TN | 65 | 10915 | 7 | 7 | 1 | 610 | 6 | 5.128287 | 0.7734619 |

# Valuing fundamental objectives

## Objectives hierarchy and attributes

The objectives identified by stakeholders can be valued in varying ways which in turn are used to calculate the value of a monitoring design. Some objectives

The numbered objectives correspond to fundamental objectives identified during the workshop. Bulleted lists within each numbered objective in bold are measurable attributes that can be used to quantify each objective. For example there are 3 attributes under objective 1 that can be quantified for each monitoring alternative. Assuming these attributes are scaled to a common scale (e.g., 0 to 1, 0 to 100) then each bullet may receive a weight of 33% if each attribute is equally important to decision makers. Alternatively these values can be weighted to reflect perceived importance by decision makers.

### 1. Quantify PS recruitment to age-1 (Natural origin)

1. Power to detect age-1 natural origin recruits if recruitment occurs
2. Segment level age-1 abundance
   1. bias
   2. precision
3. Estimate age-1 recruitment rate (natural origin)
   1. bias
   2. precision

### 2. Quantify PS population trend (natural and hatchery origin)

In simulating population monitoring designs, we are using 3 metrics to quantify how a monitoring design meets the objective of *quantifying population trend*. Specifically, the estimates from a monitoring program, as it relates to trend are evaluated by estimating basin level population growth rate .

1. **Bias: How does a trend estimate compare to the true trend.** Bias is calculated as the true value minus the estimated value. We then divide the bias by the true value to be able to combine estimates of varying magnitude (i.e., survival, abundance), recall the previous post that used proportional bias.
2. **Precision: How precise are estimated trend values.** Precision is specified as the coefficient of variation (CV) calculated as the standard error of the estimate divided by the parameter estimate. There is no real threshold for what is optimal for estimator precision regarding decision making. Generally speaking, the more precise an estimate, the better. There are other alternatives to CV. However, CV is commonly used in fisheries and therefore likely to be familiar.
3. **Performance: How likely is an estimator successful.** In some cases, an estimator like the Robust Design may not have enough information for the estimator to provide an actual estimate. This measure is quantified as the proportion of stochastic simulations where the estimator did not converge, or convergence was problematic. Let's step through an example to clarify exactly what we are talking about. Suppose we have randomly generated 200 Pallid Sturgeon Populations. Then we simulate 2 alternative monitoring programs, a catch effort program and a capture recapture program. Then we estimate trend from the estimates from the 2 designs. In the case of a catch effort based monitoring program, the performance is 100% because there are no instances where trend cannot be estimated from CPUE data, albeit zeros can be an issue at low abundances or capture probability, but that does not preclude us from calculating CPUE. However, if capture probability is low, then there may be instances where a capture recapture estimator just does not work, and estimates cannot be made because the capture recapture histories are just too sparse.

### 3. Maintain compatibility with legacy PSPAP data

1. Proportion of randomly selected bends within segment
2. Gears similarity: proportion of standard gears used by design
3. Effort similarity: deviation from average effort

### 4. Provide relevant PS model inputs

### 5. ## Remain in cost constraints

Cost is the ultimate PSPAP constraint. Designs exceeding cost containments will not be considered. Significant uncertainty exists in costs because of several factors. First, field crews tend to become more efficient over time and therefore recent PSPAP costs may not represent the actual costs. However, if the next iteration of the PSPAP uses similar sampling units (i.e., bends within segments) to the current PSPAP, then costs for field sampling will likely be similar.

Minimize costs and stay within cost constraints. The expected cost for each design will be calculated. The expected cost can be used 2 ways. First it can screen for designs that exceed cost constrains. Second, it can be used in an absolute sense to quantify design cost. Lastly,it can be used relatively by dividing the value by the expected cost of the current PSPAP and therefore values less than 1 are cheaper and values greater than 1 are more expensive. 1. Designs exceeding cost constraints not considered.

#### Estimating cost per gear deployment

Estimating cost for a day of sampling is difficult. It is the function of several factors, including but not limited to the number of personnel, pay rates, other personnel expenses, travel, time sampling, and gear maintenance. Additionally, the cost of sampling varies among years and costs may be nonlinearly related if there is economy of scale (i.e., it may cost less per sampling unit if many are done).

There are 7 field offices that have conducted PSPAP sampling since 2003. Annually the USACE provides funding to these field offices to perform PSPAP sampling. Annual values vary from

Table X.1. Field office affiliations and Missouri River segments assigned for PSPAP sampling.

|  |  |
| --- | --- |
| Field office | Segment |
| Montana Fish Wildlife and Parks (MT) | 1, 2, 3 |
| Missouri River FWCO (MR) | 4 |
| Great Plains FWCO (GP) | 5, 6 |
| South Dakota Game Fish and Parks (SD) | 7 |
| Nebraska Game and Parks Commission (NE) | 8, 1/2 of 9 |
| Missouri Department of Conservation (MO) | 1/2 of 9, 10, 11 |
| Columbia FWCO (CF) | 13, 14 |

1. Annual funding provided to each field crew
2. Calculate the number of bends sampled in a year
3. Divide the cost by the total to get a field crew specific cost per bend.
4. Model or empirical distribution

* The season will begin when water temperatures decline to 12.8C or less (in the fall) and will continue through June 30.
* The Fish Community Season will be July 1 through October 30 throughout the geographic range of the PSPAP.

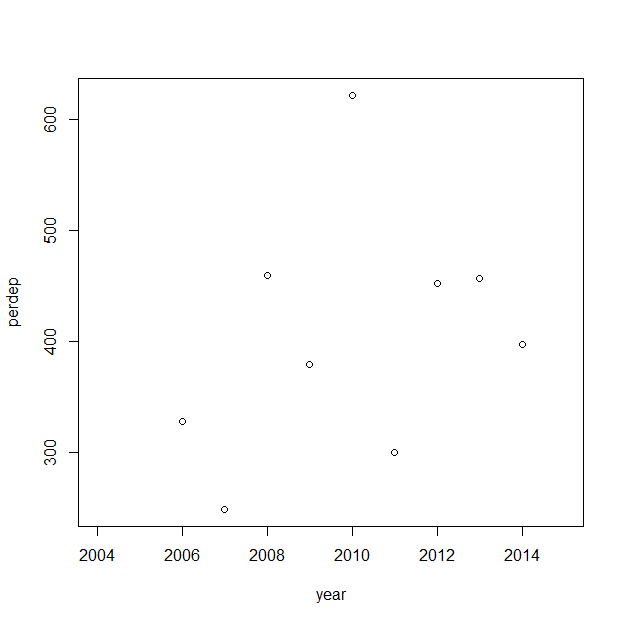
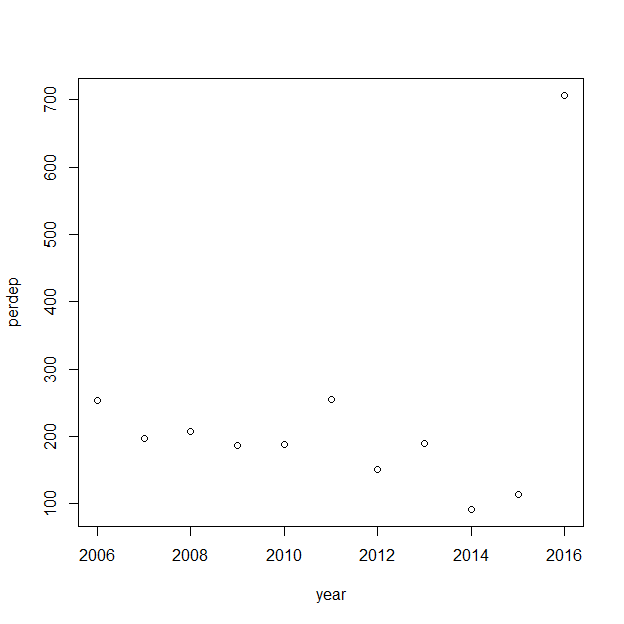
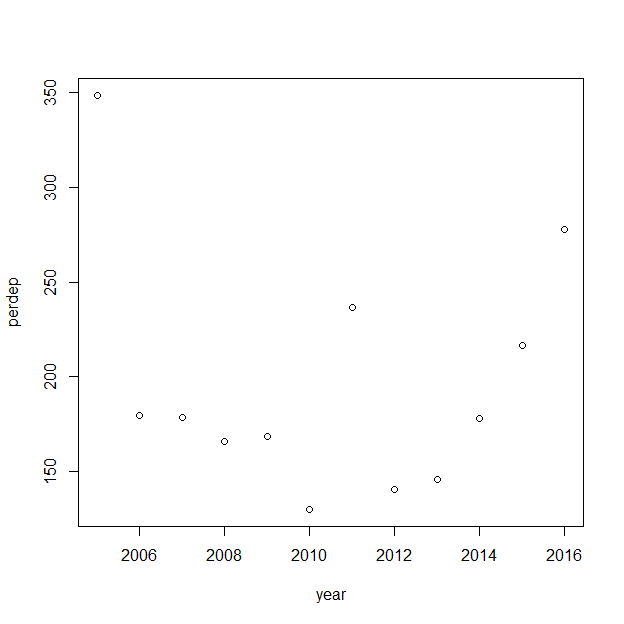
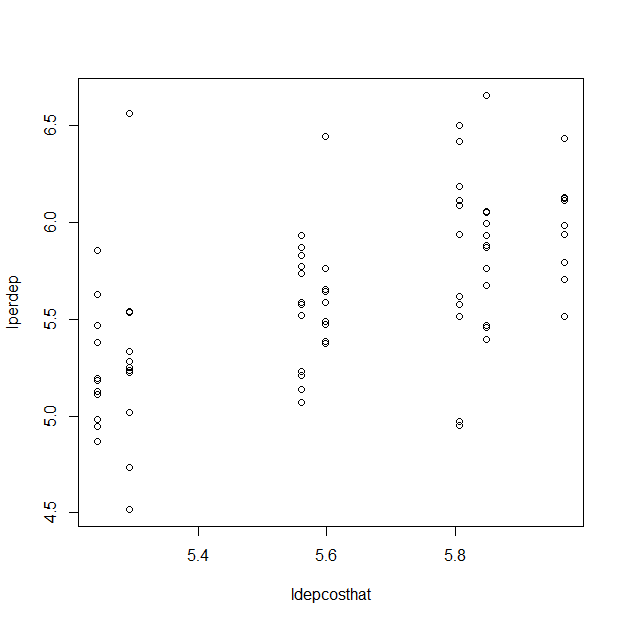
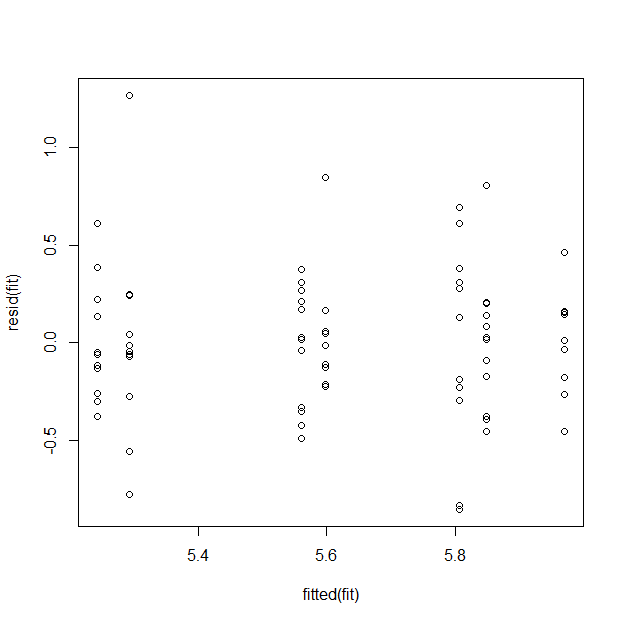
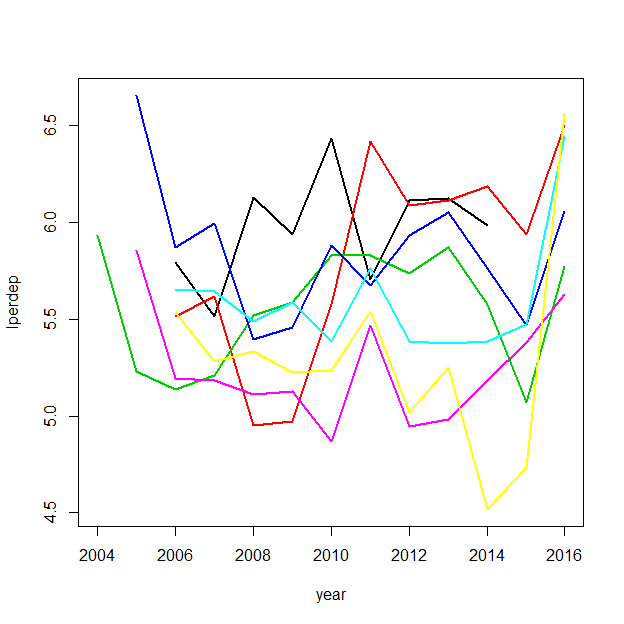
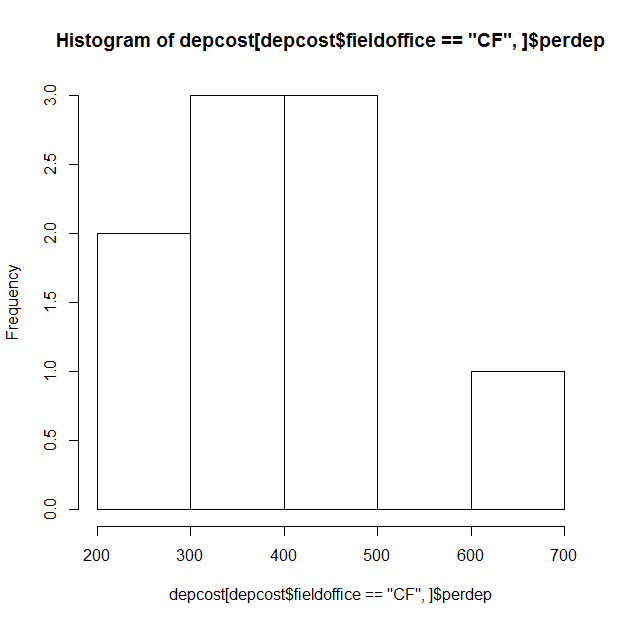
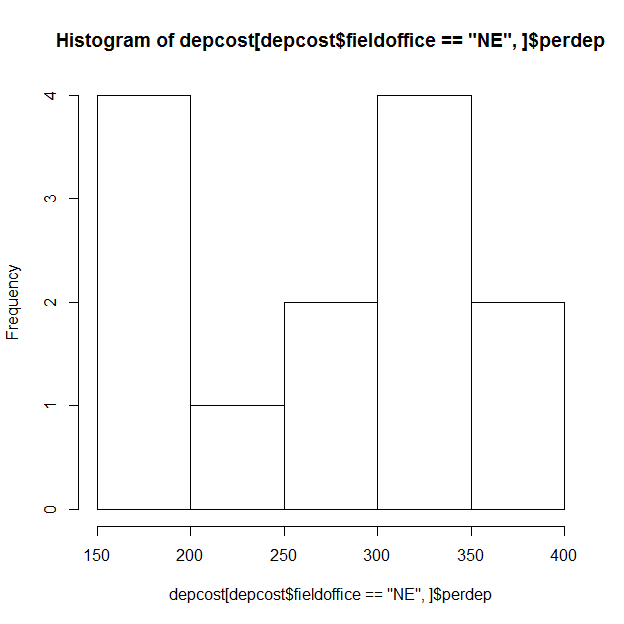
Table X.2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Field crew | Mean cost | Median cost | Std. deviation | Minimum cost | Maximum cost |
| CF | 404.75 | 397.26 | 110.14 | 248.51 | 621.71 |
| GP | 373.49 | 377.99 | 175.72 | 141.73 | 666.15 |
| MO | 368.09 | 355.54 | 147.79 | 220.28 | 776.91 |
| MR | 284.65 | 241.93 | 118.90 | 216.40 | 628.08 |
| MT | 230.92 | 190.41 | 165.62 | 91.65 | 706.87 |
| NE | 270.87 | 267.22 | 76.35 | 159.17 | 377.69 |
| SD | 197.29 | 178.42 | 63.54 | 130.14 | 348.40 |

#### Important assumptions

Currently the analysis assumes that the period of performance for field crews is the same as the fiscal year. Discussions with USACE contracting and PSPAP personnel suggested this was a reasonable assumption.

This was an example outlining how alternative monitoring programs can be objectively evaluated and compared in the context of meeting agency objectives. Many uncertainties remain in Pallid Sturgeon population dynamics and capture that will need to be accounted for. Additionally, the weighting of utility values can be a treacherous territory and how metrics are weighted can drive outcomes. However, sensitivity analyses can be conducted to evaluate the influence of weighting on outcomes. This process of evaluating alternative monitoring programs is designed purposely to be as objective as possible and therefore formally linking the outcomes of alternative monitoring designs to agency objectives with quantifiable metrics is necessary.



1. Survival (RPMA level)
   1. bias
   2. precision
2. Fecundity (RPMA level)
   1. bias
   2. precision
3. Growth (RPMA)
   1. bias
   2. precision
4. Movement
   1. fidelity
   2. among segment movement
5. Population structure and characteristics (segment level)
   1. Size structure
      1. bias
      2. precision
   2. Sex ratio (segment level)
      1. bias
      2. precision

### 3. Estimate segment-level abundance, origin and stage specific

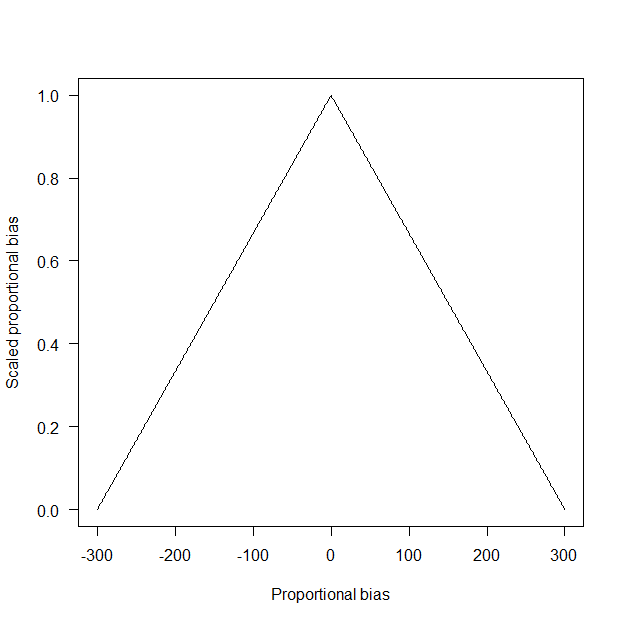
1. bias
2. precision
3. spatial distribution

### Valuing the fundamental objective

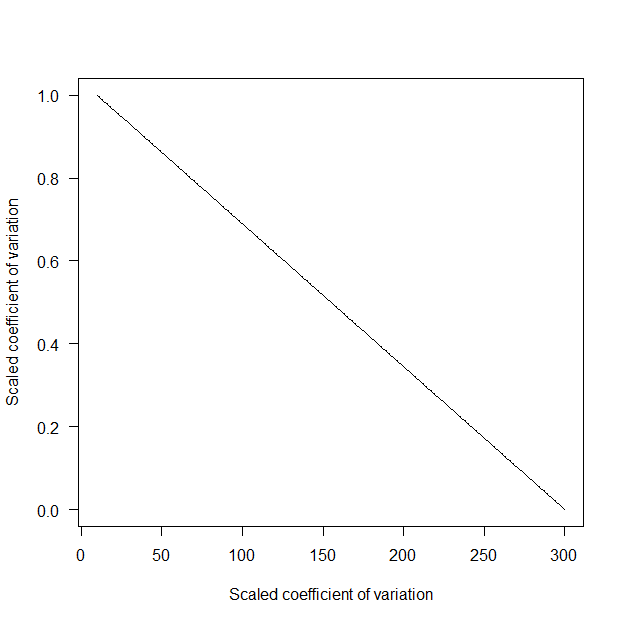
The 3 metrics described above can be combined into a single metric---commonly referred to as a utility---representing the objective to *quantify population trend* [@RN4402]. The utility is then used to evaluate alternative monitoring programs. However, one problem we run into with the metrics above is that they are on different scales. Bias can be negative or positive with values approaching 0 being best, precision is a positive number varying from 0 (best) to potentially large numbers (worst), and conformance is constrained between 0 (worst) and 100 (best).

To convert the 3 metrics to a common scale we can use methods like proportional scaling which normalized values to a specified minimum and maximum. For example, we can scale the bias to values varying from 0 to 1 as:

where is the absolute value of bias. We use absolute value here because we are assuming negative and positive bias are equally bad regarding satisfying the objective to *quantify population trend*. In the plot below, values with lower proportional bias are given higher values, and increasing values approach 0.

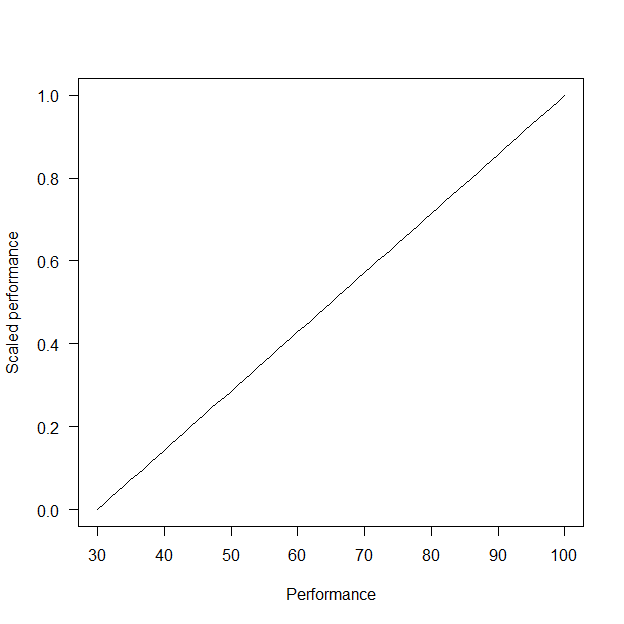


Now let's look at the precision metric. Suppose it varies from 10 to 300 (not very precise). The equation to calculate scaled precision is the same as before. However, we do not need to take absolute values since all values are positive. In the plot below, values with lower CV values have higher values, and increasing values approach 0.



Lastly, let's look at performance. Suppose performance values from simulating multiple replicates of each monitoring design vary from 35% to 100%. The difference between the performance metric with bias and precision is that higher values are more desirable and therefore we need to rearrange the proportional scaling equation to reflect this as,

In the plot below, values with lower performance values have lower values and increasing values approach 1.



### Linking monitoring designs and objectives

We can combine the utility values now such that monitoring designs resulting in a trend estimate with low bias, good precision, and good performance have higher values (i.e., approaching 1) and estimates of the trend that is biased, imprecise with poor performance approach 0. Values approach 1 because each of the scaled metrics is weighted. For example, if each metric is valued equally, then the weights would be 1/3. Alternatively, if really precise estimates of trend were desired the weight for scaled trend could be 0.5 and the remaining metrics weighted at 0.25. Suppose this last weighting scheme is the case and the output from 2 monitoring programs, a catch effort based and capture recapture program, are the values below.

* Catch effort
  + Proportional bias = -60, scaled = 0.8
  + Precision = 112, scaled = 0.648
  + Performance = 100, scaled = 1
* Capture recapture
  + Proportional bias = 5, scaled = 0.983
  + Precision = 115, scaled = 0.638
  + Performance = 90, scaled = 0.857

The scaled utility for the catch effort program is:

and the scaled utility for the capture recapture program is

The combined utility values indicate that the capture recapture program has slightly more value to achieve the objective of *quantifing population trend*.

# Results

1. This is a general measurement of effort that can be calculated for all gear types. However, we recognize that gear set time is more relavent to measuring effort for passive gears and we plan to further incorporate effort values based on distance or area metrics when considering active gears. [↑](#footnote-ref-25)
2. We use the term "standard gear" as defined in the USACE's 2017 Missouri River SOP for Fish Sampling and Data Collection (See green boxes of Appendix K) (CITATION). [↑](#footnote-ref-29)
3. 2016 AM Sub-objective 2 metric: "Population estimates for pallid sturgeon for all size and age classes, particularly for ages 2 to 3 to assess recent trends in recruitment; catch rates of all pallid sturgeon by size class (to maintain legacy data)." (CITATION) [↑](#footnote-ref-54)
4. This is a general measurement of effort that can be calculated for all gear types. However, we recognize that gear set time is more relavent to measuring effort for passive gears and we plan to further incorporate effort values based on distance or area metrics when considering active gears. [↑](#footnote-ref-72)
5. We use the term "standard gear" as defined in the USACE's 2017 Missouri River SOP for Fish Sampling and Data Collection (See green boxes of Appendix K) (CITATION). [↑](#footnote-ref-73)
6. 2016 AM Sub-objective 2 metric: "Population estimates for pallid sturgeon for all size and age classes, particularly for ages 2 to 3 to assess recent trends in recruitment; catch rates of all pallid sturgeon by size class (to maintain legacy data)." (CITATION) [↑](#footnote-ref-77)