**How to generate effort and catch estimates using the “CreelEstimates” code repository**

*Background*

This document outlines the steps for estimation of freshwater fishery angler effort and catch from survey data collected during Washington Department of Fish and Wildlife (WDFW) freshwater sportfishing monitoring (creel) projects, using a scripted analysis in the CreelEstimates code repository. It is applicable to creel projects which adhere to study design guidelines outlined in the [Statewide Protocol for Roving Creel Surveys](https://stateofwa.sharepoint.com/:w:/r/sites/DFW-TeamFPFWCreelMonitoringProgram163/_layouts/15/Doc.aspx?sourcedoc=%7B83F2B81F-39FC-49E0-955B-B6E9145AB8B7%7D&file=Protocol%20-%20Statewide%20roving%20creel%20surveys%20(Part%201%20of%202)%20-%20STUDY%20DESIGN%20(in%20progress).docx&action=default&mobileredirect=true).

CreelEstimates is a code repository with a template script that provides a workflow to fetch raw observed data, generate intermediate summaries, produce expanded point estimates and associated uncertainty measures, and output tables and figures. Data used in the scripted analysis is fetched from [data.wa.gov](https://data.wa.gov/browse?q=creel) and filtered to a user-specified fishery using the field “fishery\_name”, which describes the geographic boundaries, dates, and one or more fish species of interest.

The process of generating creel estimates relies on three primary data components: index effort counts, census effort counts, and interviews. These observed data are used to generate estimates of angler effort and catch, stratified by angler type, day type, section, and time-period (table 1), for open legal fishing days within the specified monitoring period. Index count data are collected at a discrete set of locations in the fishery area, either with direct counts of anglers or indirect counts of objects assumed to represent anglers (e.g., vehicles, boat trailers). The script currently provides two analysis options to generate creel estimates, discussed below in more detail.

Table 1. Fields used to stratify estimates of angler effort and catch in creel surveys using the WDFW Statewide Protocol for Roving Creel Surveys.

|  |  |  |
| --- | --- | --- |
| field | description | options |
| angler type | Categorical assignment of the angler effort and catch based on a combination of observed fishing location and reported means to access the fishery area | bank, boat |
| day type | Specification of the “day type” used to aggregate *n* days within a week | weekday, weekend |
| section | Specification of the spatial domain used to aggregate observations from individual survey locations. A section contains one or more sites where effort counts and interviews occur. | each fishery project specifies the sites and corresponding sections which define the fishery area |
| time-period | Specification of the time step used to aggregate *n* days of observations from creel surveys | varies by estimation method; daily, week, month, "duration" |
| angler type | Categorical assignment of the angler effort and catch based on a combination of observed fishing location and reported means to access the fishery area | bank, boat |

The point estimate (PE) method uses modified “direct expansion estimators” (Pollock et al. 1994; Thompson 2002; Jones and Pollock 2012), where mean daily catch is the product of mean daily effort and mean daily catch rate. The following equations document the process to generate angler effort and catch estimates using indirect index effort counts:

Counts of vehicles and boat trailers are converted to estimated angler counts using a ratio of the season-long sum of persons per angler group to the season-long sum of the *g*th index count type (vehicle or trailer) observed from interviews:

Where is the person count per angler group *a* for the *q*th index count type, is the vehicle or trailer count (*q*th index count type) per angler group *a*. The mean daily index count is calculated using the formula:

as the mean of the sum of *h* number of site-level index effort counts for each index count type *q*, day *d*, period *j*, and section *s*

Mean daily index count estimates of boat anglers are calculated using

Where the mean daily index count of trailers is multiplied by the angler per trailer ratio

Mean daily index count estimates of total anglers are calculated using

Where the mean daily index count of vehicles is multiplied by the angler per trailer ratio

Mean daily index count estimates of bank anglers are then calculated using

Mean daily index effort (angler hours) is calculated using

d = day

j = period

g = angler type

q = index count type

s = section

a = angler group

h = effort count event

I = index effort count

R = vehicle or trailer count from interview

B = person count from interview

C = CPUE variable

E = Effort variable

L = total fishing hours

Where is the instantaneous index count of *i*th sampling unit, Mean daily catch is estimated by multiplying mean daily effort by a daily catch rate , mean daily effort is the product of the instantaneous index effort count and day length *T*.

,and total catch is calculated as the product of mean daily catch and the total number of days in the monitoring period *N* using the following modified equations:

Indirect counts are converted to angler count estimates by multiplying mean daily counts by an interview-derived rate of anglers per vehicle or anglers per trailer. Mean daily effort is then multiplied by a “bias term ratio”, calculated by dividing census counts of anglers by the estimated counts of anglers from index effort counts.

Where is the instantaneous count of anglers the *i*th sampling unit

The Bayesian state space (BSS) method is - insert developed description of BSS from Thomas and Kale

The functions used to summarize, aggregate, and model effort and catch in CreelEstimates require formatting consistent with WDFW’s [creel database](https://dfw-fp-r5.s3.us-west-2.amazonaws.com/data-dictionaries/creel_data_dictionary.html#Database_Overview) and “proofed” data (i.e., reviewed for data collection / entry errors). Before proceeding to the instructions portion of this document, review and complete the checklist of analysis preparation steps (table 2).

Table 2. Analysis preparation steps that are require completion before using the CreelEstimates tool.

|  |  |
| --- | --- |
| step | description |
| 1 | The proposed creel project has been reviewed and approved by the Study Design Lead using the Creel Questionnaire |
| 2 | Data is stored in the WDFW creel database and accessible via [data.wa.gov](https://data.wa.gov/browse?q=creel) |
| 3 | The schedule that defined data collection events used a multi-state, stratified, probabilistic approach (see [CreelScheduleGenerator](https://github.com/wdfw-fp/CreelScheduleGenerator) Github repository) |
| 4 | The data of interest is associated with a specific fishery\_name, defined in the database with a fishery\_name, start\_date, and end\_date |
| 5 | Index sites and census sections for the fishery are specified in the fishery\_location\_lut |
| 6 | Closure dates during the monitoring period are specified in the closures table (TBD, proposed table/view name “fishery\_closures\_lut”) |
| 7 | Data has been undergone a QA/QC process and is approved in "Data Proofed" field in Fish Apps |

*Instructions*

The following steps provide an overview of how to generate estimates of effort and catch from creel surveys using the CreelEstimates tool. If additional guidance is needed, contact your Study Design Lead.

1. **Clone the “CreelEstimates” repository**
   * The most up-to-date version of CreelEstimates is located on GitHub **here** and is stored as a code [repository](https://docs.github.com/en/repositories/creating-and-managing-repositories/about-repositories)
   * To use CreelEstimates, the repository must be cloned to your local computer (for further details see “[Clone the Repository with RStudio](https://resources.github.com/github-and-rstudio/)” sub-section). The repository can be located anywhere on your computer since you’ll interact with an RMarkdown script file (.Rmd) that runs within a RStudio project file (.Rproj) that uses relative filepaths.
   * Before a repository can be downloaded, a user must have a GitHub account (see “[Prerequisites](https://resources.github.com/github-and-rstudio/)” subsection) and have Git and RStudio installed on their computer (see “[Install Git and RStudio](https://resources.github.com/github-and-rstudio/)”)
2. **Create a new analysis project** 
   * Open the “CreelEstimates” folder
   * Open the “CreelEstimates.Rproj” file

**This section here pending final plan for file structure / user outputs**

Draft option used here:

* open the projects folder
* open “establish\_analysis\_project.R” script. Enter the fishery\_name, start date, and end date for the analysis of interest. Run the script, which will create a project folder and renamed copy of the fw\_creel.Rmd analysis template script
* open newly created analysis script .Rmd

1. **Specify the metadata parameters of the script**
   * In the [YAML](https://yaml.org/spec/) metadata section at the top of the script, enter parameter values which specify the data and analysis control options for the .Rmd script (table 3)

Table 3. Parameters in the YAML metadata section which determine the data fetched from data.wa.gov and analysis controls linked to the study design and methods for a specific creel project.

|  |  |
| --- | --- |
| parameter | description |
| fishery\_name | character string with waterbody area, focal species, year / year-group to describe distinct creel project |
| est\_date\_start | minimum date of observations fetched from data.wa.gov |
| est\_date\_end | maximum date of observations fetched from data.wa.gov |
| est\_catch\_groups | specification of one or more groups of focal species, life stage, mark status, and encounter types of interest |
| person\_count\_type | specification designating method used to count anglers during census surveys, either counting the total number of people in angler groups (“group”) or counting only people observed actively fishing (“angler”) |
| period\_pe | time period for aggregation of observations; for period\_pe options are: “week”, “month", and "duration" |
| period\_bss | time period for aggregation of observations; for period\_bss options are: "day" and "week" |
| days\_wkend | Days assigned to "weekend" day type group (weekend or weekday) |
| index\_count\_types | specification for objects that are counted during index effort counts, either "Vehicle/Trailers Only" or "Bank Angler / Boat Angler" |
| census\_expansion | Option to use either observed census data ("Direct") or assumed values ("Indirect") to calculate the effort bias term ratio |
| min\_fishing\_time | The minimum fishing time per angler group interview that is retained to calculate their catch group specific catch rate |
| dir\_output | directory name in file path where output from script is saved |

1. **Interactively run the project-specific fw\_creel.Rmd script**
   * The CreelEstimates script is broken up into code chunks.
     + Within each chunk, there are rows of code and most of which simply needs to be run (and thus not edited). However, see next bullet.
     + There is also code located within in *R\_functions* and is run behind the scenes. Feel free to open these source files to understand how CreelEstimates is working. This code but should not be edited unless you establish a formal plan to do this using a distributed version control system (i.e., create a branch to CreelEstimates, make edits to the code, commit and push the changes to your branch of the repo, and submit a pull request to tell collaborators about the changes in your branch)
2. **Evaluate output** 
   * **Table with fields and descriptions of output (table \_\_**