Cumulative Effects Model for Prioritizing Recovery Actions (CEMPRA)

Guideline Documentation

March 31, 2023

Prepared for the BC Ministry of Environment

Cumulative Effects Model for Prioritizing Recovery Actions (CEMPRA)

Guideline Documentation

March 31, 2023

Authors:

Matthew Bayly, MJBA, ESSA Technologies Ltd.
Alexandra Tekatch, ESSA Technologies Ltd.
Dr. Jordan Rosenfeld, the University of British Columbia, BC Ministry of Environment
Dr. Eva Enders, INRS
Lauren Jarvis, Department of Fisheries and Oceans

Contributors:

Andrew Paul, Alberta Environment and Protected Areas
Kyle Wilson, Central Coast Indigenous Resource Alliance
Pedro Gonzalez, University of British Columbia
Laura MacPherson, Alberta Environment and Protected Areas
Isuru Dharmasena

Suggested Citation:

Bayly, M., J., Tekatch, A., M., Rosenfeld, J., Jarvis, L., Enders, E., 2023. Cumulative Effects Model for Prioritizing Recovery Actions (CEMPRA). Documentation prepared by ESSA Technologies Ltd. for the BC Ministry of Environment. 77 pp.

Funding provided by British Columbia Salmon Restoration and Innovation Fund (BCSRIF).

Cover Photo:
Chinook Salmon © Ryan Hagerty, USFWS

ESSA Technologies Ltd.
Vancouver, BC Canada V6H 3H4
www.essa.com

Table of Contents

Purpose

This document is the primary guidance document for the Cumulative Effects Model for Prioritizing Recovery Actions (CEMPRA). It includes a walkthrough of the model, its components, inputs and outputs, benefits and limitations, and instructions for using it in two available formats: an R package and an R Shiny Web Application. For further details on the model and examples of its implementation, please refer to MacPherson et al. (2018).

Acknowledgements

‘Joe Model’ Testimonial to Joe Nelson

The CEMPRA model contains the *Joe Model* as a subcomponent. The Joe Model nickname was given in honour of the University of Alberta Ichthyologist Dr. Joseph Nelson (MacPherson et al., 2018; Murray et al., 2012). We acknowledge Dr. Nelson’s profound impact on ichthyology and the original Alberta cumulative effects modelling framework acting as the foundation of the CEMPRA tool.

Executive Summary

The Cumulative Effects Model for Prioritizing Recovery Actions (CEMPRA) is a cumulative effects modelling framework. The CEMPRA tool uses a series of standardized stressor-response functions to link environmental attributes to the system capacity and productivity of a target species/system. This framework design is as generalizable, simple, and versatile as possible so that users can apply the model to various geographic regions, contexts, systems, and species. As the name suggests, the CEMPRA tool helps prioritize recovery actions for data-limited species and species-at-risk, with the flexibility to accommodate both data-rich and data-poor study systems. The CEMPRA tool is accessible as an open-source R package (https://github.com/essatech/JoeModelCE) and R Shiny interactive web application (https://github.com/essatech/JoeModelCEShiny).

Stressor-response functions form the foundation of the CEMPRA tool. A stressor variable is broadly characterized as an environmental driver resulting in an observable biological response in a target population (Rosenfeld et al., 2022). Within the CEMPRA tool, stressors represent and capture various metrics of cumulative effects (direct or proximal) and their associated impact pathways (e.g., stream temperature, sedimentation, habitat loss).

Stressor-response functions are developed for each metric in a standardized format and linked to population-level productivity (mean system capacity) or specific vital rates within a life cycle modelling framework. Users then populate another table of stressor-magnitude values joined to various locations (spatial units) of interest. Finally, the CEMPRA tool runs to generate stochastic simulations of the study system under different scenarios. Comparisons between scenarios are commonly made against a default reference (status quo) scenario. Scenarios generally consist of various “alternative futures” to characterize potential impacts from development activities and/or alternative restoration/recovery efforts. Comparisons between scenarios can be quantitative (e.g., looking at a weighted mean system capacity or relative productivity) or qualitative by simply looking at a heatmap of stressors across the landscape.

There are two endpoints within the CEMPRA framework: 1) Joe Model estimates of system capacity (a generalized response metric), and 2) life cycle model outputs, where stressor-response relationships are tied to a specific life stage and vital rate. The life cycle model complements the CEMPRA tool, where users can adjust vital rate parameters to estimate cumulative effects at the population level life cycle model.

Additional supporting resources are being developed to facilitate ease of use and collaboration between individuals studying cumulative effects. These resources include the development of an online stressor-response library (digital archive), example species profiles, case studies and tutorial resources.

1 Introduction

Many species are under significant pressure from human development and resource use. These effects are coupled with existing pressures from climate change and natural disturbances. Collectively, these pressures can compound and interact to have cumulative effects on many species and ecosystems. Attempts have been made to quantify, map, and model cumulative effects across a landscape or region of interest to understand current conditions, high vs low-risk areas, and potential future conditions under different (hypothetical) development scenarios. However, the applicability of many analytical tools and frameworks is linked to specific geographies, study systems or regulatory requirements. In addition, many tools or frameworks are too specific or over-generalized, causing researchers to start from scratch with duplicated efforts when undertaking assessments on cumulative effects.

The CEMPRA tool leverages (and automates) common data processing and analytical pathways reoccurring across numerous cumulative effects assessments. The development team behind the CEMPRA tool also recognizes that many cumulative effects assessments are collaborative and highly iterative. As such, an interactive web application was developed, alongside an analytical R-package, to facilitate the development of cumulative effects assessments in a workshop-like setting. The CEMPRA R-Shiny web application facilitates the inclusion and leadership of land stewards and decision-makers in the assessment and analytical process by working around a centralized interactive choropleth map. Key locations are represented by polygons on the map linked to interactive stressors and stressor-response relationships. Numerous interactive visualizations and summary tables are included within the tool to facilitate rapid “what if” assessments. The CEMPRA tool does not include all possible options for advanced analytics. However, it can be a useful starting point to engage different user groups and work towards a shared understanding of key drivers, processes and opportunities for a target study system and region of interest.

The inspiration for the CEMPRA tool originated from many case studies centred around aquatic species-at-risk in Canada. However, the underlying framework is not explicitly bound to aquatic or terrestrial ecosystems. Aquatic ecosystems have remained the primary focus of the CEMPRA tool’s development in response to numerous contemporary priorities, including a) practical guidance for watershed management, b) recent amendments to the Fisheries Act prompting the consideration of cumulative effects (past and current), c) empowering practitioners with tools and resources to initialize collaborative cumulative effects assessments.

Key components of the CEMPRA include a focal study system (e.g., a target population), a focal area of interest (i.e., a target region with defined sub-areas, locations, or spatial units), key stressors of interest (e.g., stream temperature, sedimentation, habitat loss), stressor-response functions linking the key stressors to the focal study system, estimates of stressor magnitude levels across the target region of interest (e.g., current conditions across the focal locations and spatial units), different assessment endpoints for data-rich and data-limited systems and scenarios (future “what if” possibilities) to represent in the model.

Core Components of the CEMPRA Tool:

Study System: In all applications of the CEMPRA tool, users should frame their assessment around a focal study system. A study system is commonly described as a valued (target) species within a defined geographic region. Previous (and ongoing) applications of the CEMPRA have included Athabasca Rainbow Trout (Sullivan et al. 2017), Chinook Salmon in the Nicola Basin (Pearsall et al. 2022) and the Plains Sucker in Southeastern Alberta (L. Jarvis, personal communication, February 23, 2023). A study system in the CEMPRA tool should be a population or valued ecosystem component. On occasion, the term “study system” is used in conjunction with “cumulative effects” to describe a project or development within the Environmental Impact Assessment (EIA) literature. However, the CEMPRA uses specific valued ecosystem components as the study system. Usually, these are target populations, but occasionally they can be interpreted as broader entities with basic applications of the CEMPRA where stressors are linked to “aquatic ecosystems”. If broader interpretations describe a study system, care should be taken to ensure that the stressor response functions are still interpretable and biologically relevant. For example, suppose the CEMPRA is used for a multi-species assessment. In that case, creating species-specific stressor response curves and running scenarios for each species will likely be necessary.

Stressors: Stressors and stressor-response functions are a core part of the CEMPRA tool. For this tool, “stressors” can be defined broadly as “any environmental variable (e.g., temperature, sediment, predation, competition) that can induce a biological response (positive or negative)” (Rosenfeld et al. 2022). For example, suppose that the assessment intends to understand cumulative effects across a study region. In that case, stressors will be described as environmental variables (or drivers) that prevent the study system from reaching a fully realized system capacity that would presumably be possible without any harmful impacts. Stressors and stressor-response functions are described further in the next section.

Stressor-Response Function: Stressor-response functions link stressors to the study system. The key defining components of stressor-response functions within the CEMPRA tools are that they represent quantitative linkages between raw stressor values and the predicted biological response. Stressor-response functions are equivalent to dose-response curves. They can also be interpreted as habitat suitability curves; however, the default assumption in most high-level cumulative effects assessments is that target locations under evaluation are already characterized as having the capacity to support the study system. Stressors and stressor-response functions are described further in the next section.

Locations: The terms “locations”, “study areas”, “spatial units” and/or “assessment units” are used interchangeably to describe discrete locations across the target study region. Locations are represented in the CEMPRA tool as spatial polygons. Locations should be partitioned on environmental heterogeneity across the study area and based on an area's presence (or capacity) to sustain the target species. Ideally, location breaks are chosen such that stressor values are largely consistent within a location relative to between locations. Location breaks can also represent known (or suspected) subpopulations. In most cumulative effects assessments of aquatic ecosystems, location breaks will almost always be generated based on the boundaries of watersheds or subbasins since these natural geographic breaks often drive differences in key stressor values across the landscape.

Assessment Endpoints: The ability to effectively link environmental stressors to a focal study system in the form of stressor-response functions is the core underlying theme of the CEMPRA tool. In related assessments, data availability and timelines often dictate how researchers and practitioners choose to link stressors to their target study system. Simplified roll-up-like summaries are convenient in data-limited systems, while more integrated modelling can often take place in data-rich systems. The CEMPRA tool provides two major workflows and assessment endpoints. Stressor-response functions can be summarized as the product of the scaled response of all stressors across locations (the classical 'Joe Model'), or stressors can be run through an integrated life cycle model where stressor-response functions are linked to vital rates (e.g., survivorship, capacity). The simplified roll-up ('Joe Model') is useful for data-limited species, but it also serves as a convenient framework for a rapid first-pass assessment to generate stressor heatmaps. The integrated life cycle model has special utility for more data-rich systems where it's possible to weight and understand stressors through the lens of a demographic framework. This can be especially important where some stressors will have disproportionate impacts on a specific bottleneck.

Scenarios: Scenarios represent unique sets of conditions of constraints to represent in the tool for structured comparisons. Scenarios (or scenario profiles) can include a combination of changes in one or more stressors at one or more locations to represent a hypothetical management action (or inaction). Scenarios are developed as "what if" assessments (e.g., "what if restoration activity x occurred at location(s) y" "what if development x continued across location y" etc.). Scenarios can be implemented as changes to the stressor values (across locations) and/or changes to the stressor-response relationships, changing the underlying assumption of impact pathways.

The purpose of this user guide is to introduce these concepts in further detail with illustrative examples and demonstrate how they can be brought together to implement cumulative effects assessments in the CEMPRA tool. The following sections of the guide include setup instructions, a quick start guide, an overview of data inputs and sample use cases. The purpose of the CEMPRA tool is to create a generalizable and easy-to-use cumulative effects modelling framework that is adaptable to many different systems and/or species.

2 Stressor-Response Functions

Stressor-response functions describe the relationship between a specific stressor (such as a pollutant or a change in temperature) and the response of a target species (such as growth rate, reproduction, or mortality). Stressor-response functions are used to predict how a population (or study system) will respond to changes in the environment and to identify thresholds or "critical levels" at which a stressor becomes harmful. Stressor-response functions are often used to inform environmental policy and management decisions, for example, by identifying risk levels of pollution or temperature change for a particular species or ecosystem[1]. Stressor-response functions are generally developed through primary research (i.e., mechanistic, empirical, experimental etc.) and expert opinion.

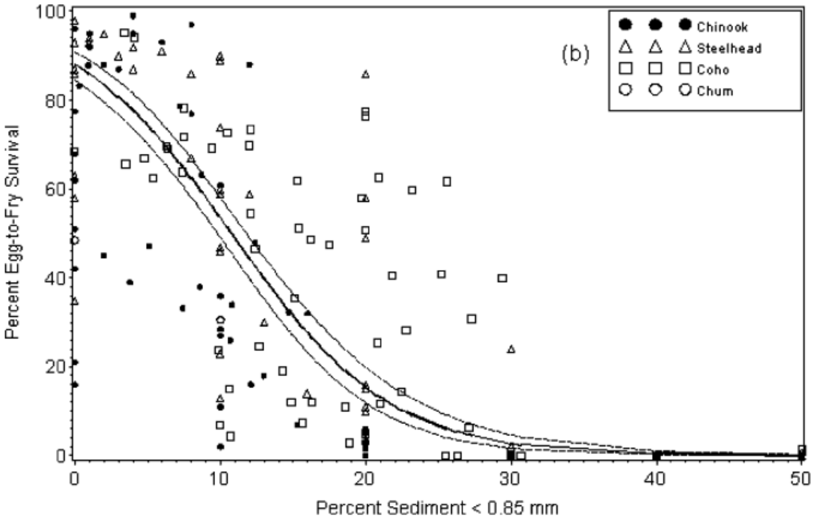


Figure 1. Example of a stressor-response function for Pacific salmon from Jensen et al. (2009) showing the relationship between a stressor (sedimentation percentage) on the x-axis and the biological response (percent egg-to-fry survivorship) on the y-axis.

There are several types of stressor-response functions, including linear, threshold, and non-linear. Linear functions describe a simple, linear relationship between the stressor and the response, with the response increasing or decreasing at a constant rate as the stressor increases. Threshold functions describe a breakpoint at which a stressor becomes harmful, beyond which the response increases rapidly. Non-linear functions describe more complex relationships, with the response changing at different rates as the stressor increases. The example provided in Figure 1 shows an example of a customized non-linear stressor-response function fit to empirical data (reference points). It is also possible to represent stressor-response functions as interactions between variables, such as the risk of exposure to a harmful pathogen being temperature dependent.

For a more in-depth discussion on the foundations of stressor-response functions, refer to the following resources:

- **Rosenfeld et al. 2022**: Conceptual overview of stressor-response functions as a generalizable model for context dependence. This paper provides a valuable overview to conceptualize stressors as a mechanism to characterize the state of a system and

---

[1] https://www2.gov.bc.ca/assets/gov/environment/air-land-water/water/waterquality/water-quality-guidelines/approved-wqgs/wqg_summary_aquaticlife_wildlife_agri.pdf

ecological process. This also outlines common forms of stressor-response functions and key considerations for the creation of a stressor-response function from empirical data.

- **Piet et al. 2021**: *A roadmap towards quantitative cumulative impact assessments: Every step of the way*. Provides an important roadmap for working groups to consider linkages between land-use activities, resulting ecosystem pressures, functional linkages in space and time (exposure) and the consideration of endpoints to target study systems.
- **Incorporating Indigenous Knowledge**: In many instances, stressor-response functions may not be developed exclusively from empirical data but instead developed through expert opinion from local communities based on value systems. Where appropriate, working groups may include a customized stressor-response function to represent potential risks and values based on traditional knowledge systems and expert opinion. Refer to Hounde (2007) and Alexander et al. (2019) for further discussion. Examples of many other values-based Indigenous-led cumulative effects management programs exist across Canada[2,3].

3 Assessment Endpoints

The CEMPRA framework offers different assessment endpoints to choose from. The specific endpoint to use will depend on the application, data availability, and overall objectives of the assessment. In the following section, we will describe stressor-response functions and the simplified "Joe Model" that combines the effects of multiple stressors. We will also discuss the integrated life cycle model, which links stressors to vital rates to project productivity and capacity of the target system. Although assessment endpoints can be mutually exclusive, they all rely on stressor-response functions as the central theme.

---

[2] https://www.icce-caec.ca/

[3] http://metlakatlacem.ca/ability-to-steward/

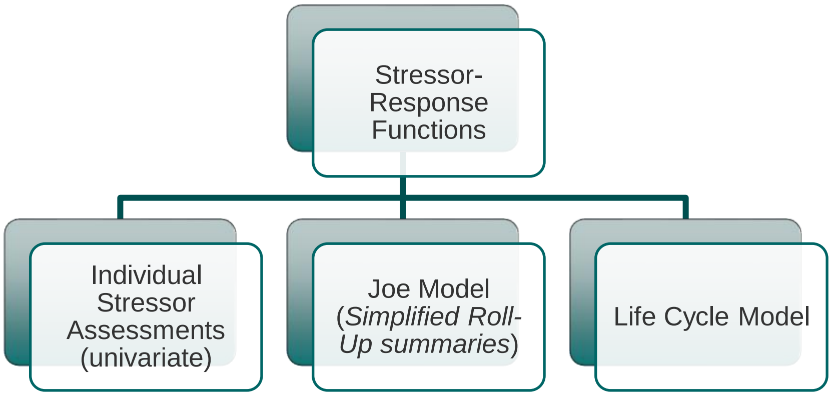


Figure 2. Overview of alternative assessment endpoints within the CEMPRA framework.

3.1 Individual Stressor Assessments (univariate)

The CEMPRA tool is based on a centralized map interface (see Section 6). In many instances, it is convenient to flip through stressors individually to produce univariate heatmaps (choropleth maps) of each stressor (see Figure 3). These summaries are simple yet useful to provide a general overview of the study area, stressors, and stressor-response functions. The interpretation of univariate stressor summaries is largely qualitative. The intent is to identify hotspots and/or determine which stressors are high everywhere or low everywhere based on the input stressor magnitude data for each location and the corresponding stressor response function for each stressor.

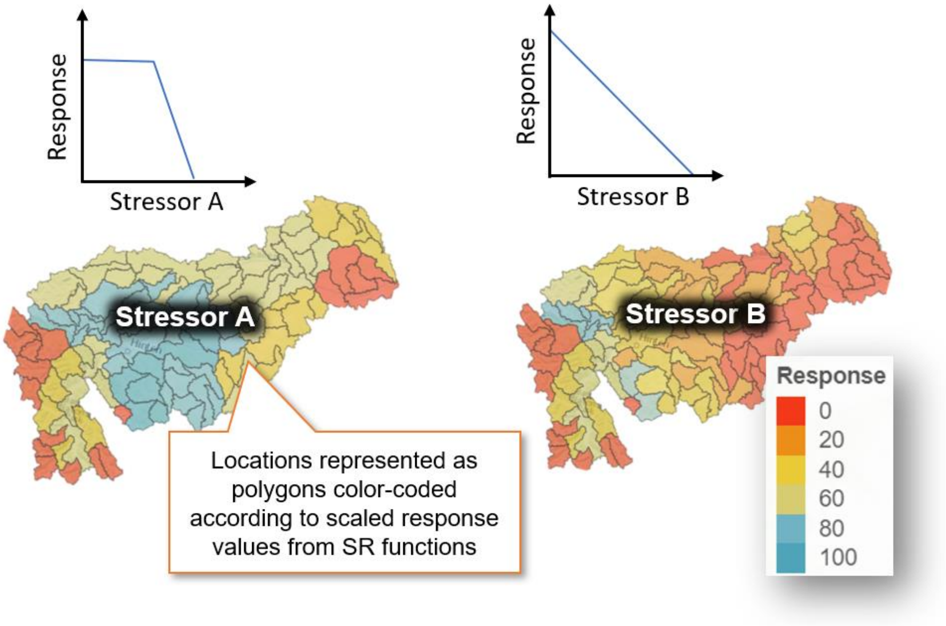


Figure 3. Example of single (univariate) assessment of stressors and stressor-response relationship for hypothetical stressors A & B with univariate heatmaps generated for the study area.

3.2 Joe Model: Simplified Stressor Roll-up Summaries

The Joe Model component of the CEMPRA tool leverages the library of stressor-response functions, loaded by the user, to generate a simplified roll-up summary across stressors. These summaries are simply the product of the scaled response values (0 to 1) across stressors for each location (**Figure 4**). Different stressors can be selected (or omitted) from this summary to characterize different impact pathways.

In the Joe Model, the response component of each stressor-response function (y-axis) is characterized broadly as the ‘Mean System Capacity’ of the target study system. Mean System Capacity’ is a generalized term used to allow for simplified multi-stressor comparisons. For convenience, we assume that each stressor is linked to the adult system capacity, but in reality, individual stressors will likely have more complex relationships.

In the Joe Model summaries, the cumulative effect score across stressors is calculated additively by multiplying the mean system capacity metrics together. The Joe Model does not require the weighting of individual stressors since each stressor is used to predict the same dimensionless response metric (mean system capacity, %). By doing this, the Joe Model avoids long-standing difficulties associated with weighting impacts (Walters, 1997).

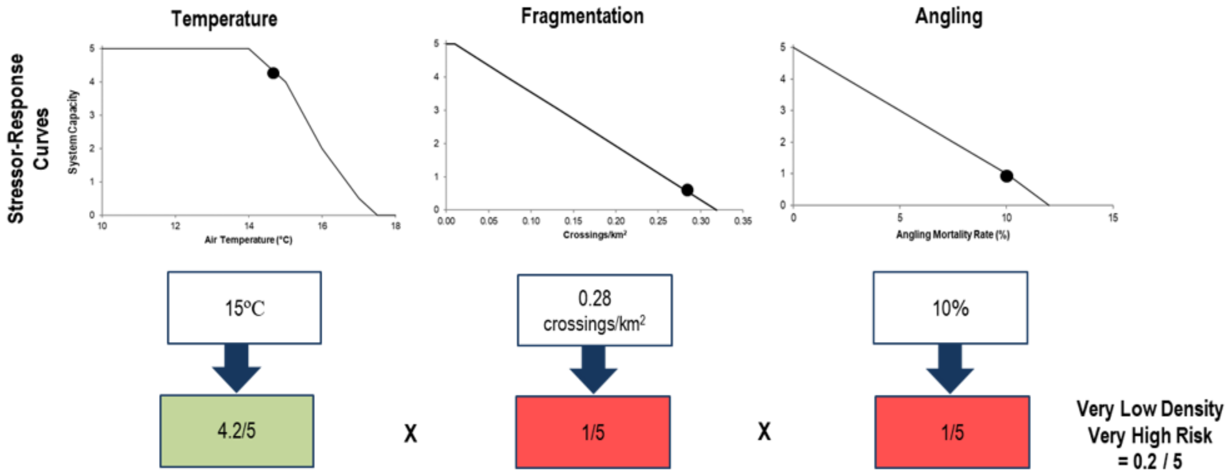


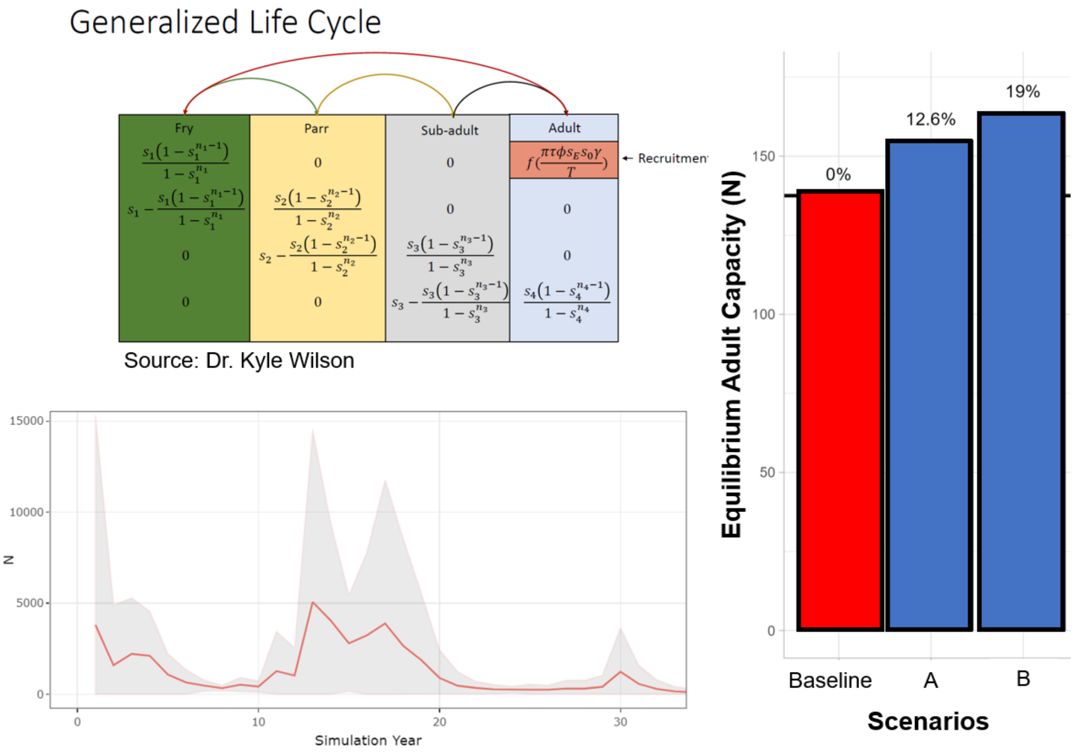
Figure 4: Additive effect of multiple stressors in the Joe Model. Figure from MacPherson et al., 2020.

The full implementation “Joe Model” extends this basic summary to include stochastic simulations with uncertainty represented in both the raw stressor values for each location and the response function (described further in Section 5). These summaries are useful to represent the risk of cumulative effects or current conditions as a distribution rather than a fixed value.

3.3 Life Cycle Model

An integrated life cycle modelling framework is also included within the CEMPRA tool to evaluate the effects of stressors on population-level productivity and capacity (Figure 5). The embedded life cycle modelling framework consists of a stage-structured matrix model that allows users to link stressor-response functions to individual vital rates (e.g., egg survivorship, fry capacity etc.). The life cycle model estimates relative changes to population-level productivity and capacity through simulations. A flexible species profile input dataset is available such that users can change values in the species profile to represent different study systems (e.g., Athabasca Rainbow Trout, Chinook Salmon etc.). See sections 6.4 and 7 for a detailed summary of the life cycle modelling framework within the CEMPRA tool.

Developing a life cycle model is not a requirement for all applications of the CEMPRA tool; however, it can be a powerful extension to identify stressors, impact pathways, and scenarios to be tightly coupled to critical demographic bottlenecks.



Source: Dr. Kyle Wilson

Figure 5. Sample outputs from the Life Cycle Modelling component of the CEMPRA tool.

4 Initial Setup

There are two ways to interact with the CEMPRA (Joe Model): through the R Shiny web application or directly using the JoeModelCE R package. Individuals unfamiliar with R can access the web application currently available here:

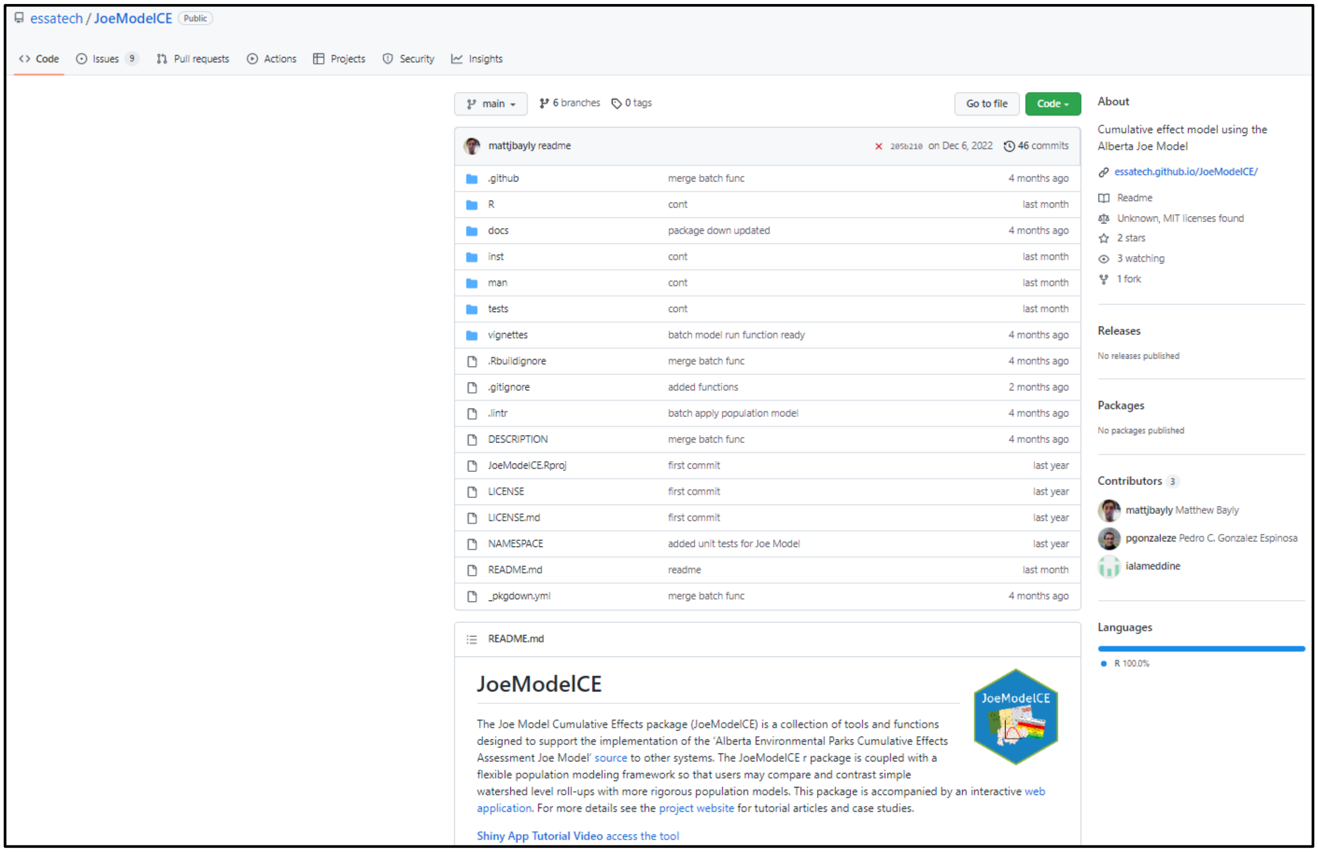
The web version of the CEMPRA Tool: https://essa.shinyapps.io/JoeModelCEShiny/

However, we strongly recommend that individuals familiar with R download a local copy of the CEMPRA (Joe Model) Shiny application and run it on their own computers through RStudio. Running the application from your own computer reduces latency and other issues.

The CEMPRA (Joe Model) framework is available as an R package and an R Shiny application. All components of the project are freely available and open source. Both the R package and the R Shiny application are freely available for download from GitHub:

R package: https://github.com/essatech/JoeModelCE/
R Shiny application: https://github.com/essatech/JoeModelCEShiny

4.1 R-Package



mattjbayly readme ✕ 205b210 on Dec 6, 2022 46 commits

.github	merge batch func	4 months ago
R	cont	last month
docs	package down updated	4 months ago
inst	cont	last month
man	cont	last month
tests	cont	last month
vignettes	batch model run function ready	4 months ago
.Rbuildignore	merge batch func	4 months ago
.gitignore	added functions	2 months ago
.lintr	batch apply population model	4 months ago
DESCRIPTION	merge batch func	4 months ago
JoeModelCE.Rproj	first commit	last year
LICENSE	first commit	last year
LICENSE.md	first commit	last year
NAMESPACE	added unit tests for Joe Model	last year
README.md	readme	last month
_pkgdown.yml	merge batch func	4 months ago

About

Cumulative effect model using the Alberta Joe Model

essatech.github.io/JoeModelCE/

Readme

Unknown, MIT licenses found

2 stars

3 watching

1 fork

Releases

No releases published

Packages

No packages published

Contributors 3

mattjbayly Matthew Bayly

pgonzaleze Pedro C. Gonzalez Espinosa

ialameddine

Languages

R 100.0%

README.md

JoeModelCE

The Joe Model Cumulative Effects package (JoeModelCE) is a collection of tools and functions designed to support the implementation of the 'Alberta Environmental Parks Cumulative Effects Assessment Joe Model' source to other systems. The JoeModelCE r package is coupled with a flexible population modeling framework so that users may compare and contrast simple watershed level roll-ups with more rigorous population models. This package is accompanied by an interactive web application. For more details see the project website for tutorial articles and case studies.

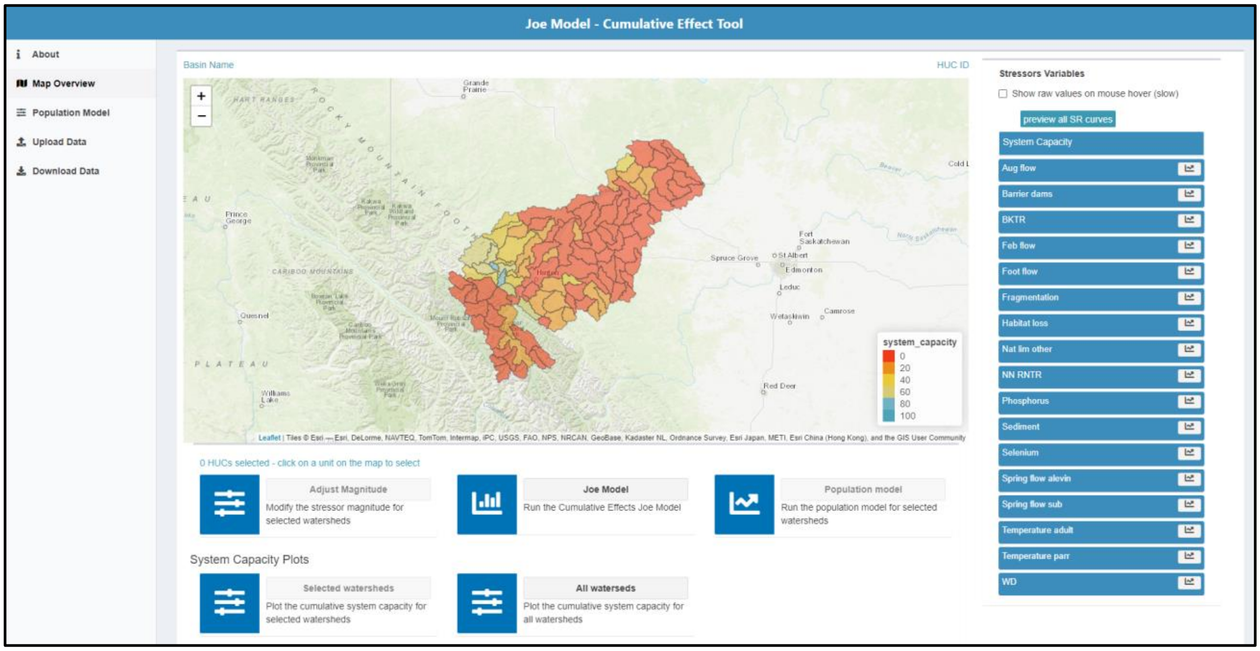
Shiny App Tutorial Video access the tool

Figure 6: JoeModelCE package repository available on GitHub (https://github.com/essatech/JoeModelCE/).

The *JoeModelCE* package is part of a larger initiative to develop a framework for modelling cumulative impacts to Species-at-Risk (SAR) to guide recovery planning and adaptive management based on stressor-response functions related to taxa-specific threats. This framework allows users to generate models across a range of complexity and data quality, treating stressor-response functions as modular entities.

A quick start guide is provided below, but see Appendix D for tutorials and function documentation.

4.2 R Shiny Application



Modify the stressor magnitude for selected watersheds

Run the Cumulative Effects Joe Model

Run the population model for selected watersheds

Plot the cumulative system capacity for selected watersheds

Plot the cumulative system capacity for all watersheds

Figure 7: Map Overview of results in the JoeModelCE Shiny web application.

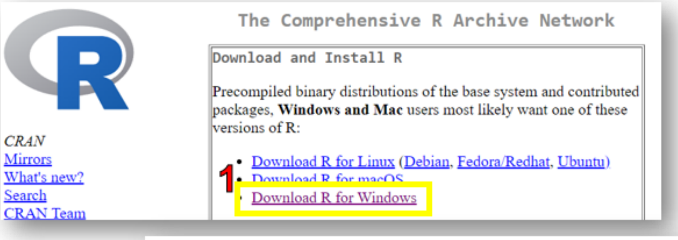
The *JoeModelCE* R package is accompanied by an interactive R Shiny web application (https://essa.shinyapps.io/JoeModelCEShiny/). This application acts as a flexible, user-friendly interface to interact with key components of the CEMPRA tool. It accepts user inputs for stressor-response functions, stressor magnitude, spatial units/polygons, and vital rates (for the life cycle model portion). Results are easily mapped, summarized, and plotted within the **Map Overview** page of the web application. Scenario results generated from the tool (and the associated life cycle model) can be exported directly from the application as an Excel spreadsheet (.xlsx). See Section 5 for details on data inputs and Section 6 for a complete walkthrough of the application.

4.3 Initial Setup and Installation

Inexperienced R users can run the Joe Model directly through the web application (https://essa.shinyapps.io/JoeModelCEShiny/). However, downloading and running the model directly in RStudio is recommended for improved performance. Running the application locally (outside of shinyapps.io) also guarantees data privacy. To run the application locally, users must first have R and RStudio downloaded and installed on their computers. The *JoeModelCE* R package (https://github.com/essatech/JoeModelCE/) and the local version of the R Shiny web application (https://github.com/essatech/JoeModelCEShiny) can be downloaded from GitHub.

To install R and RStudio on your computer:

1. Go to the R website (https://cran.r-project.org/) and follow the instructions to download the latest version of R for your operating system (Windows, Mac, or Linux).



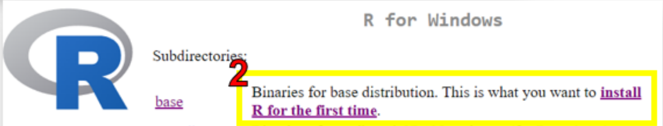
The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- Download R for Linux (Debian, Fedora/Redhat, Ubuntu)
- Download R for macOS
- Download R for Windows

CRAN
Mirrors
What's new?
Search
CRAN Team



R for Windows

Subdirectories:

base	Binaries for base distribution. This is what you want to **install R for the first time**.
contrib	Binaries of contributed CRAN packages (for R >= 3.4.x).

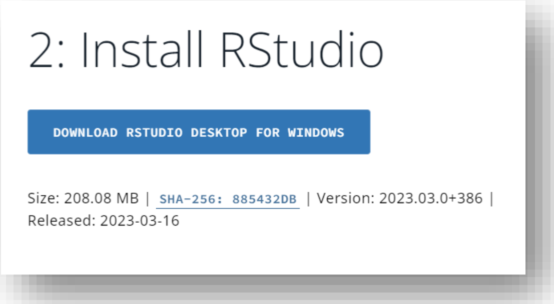
CRAN

R-4.2.3 for Windows

Download R-4.2.3 for Windows (77 megabytes, 64 bit)
README on the Windows binary distribution
New features in this version

CRAN

2. Once the download is complete, double-click (open) the installer file and follow the prompts to install R on your computer.
3. To open and run R scripts (files ending in .R), you can use RStudio, a popular Integrated Development Environment (IDE) for R. You can download the latest version of RStudio from: https://rstudio.com/products/rstudio/download/.

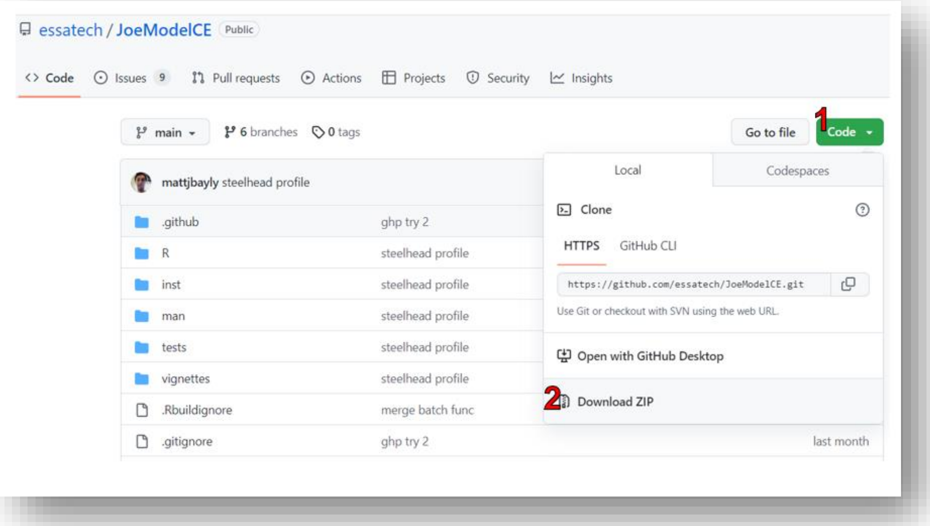


2: Install RStudio

Size: 208.08 MB | SHA-256: 885432DB | Version: 2023.03.0+386 | Released: 2023-03-16

To install the *JoeModelCE* R package on your computer:

1. Download the *JoeModelCE* R package from GitHub (https://github.com/essatech/JoeModelCE/) onto your computer by clicking the green "Code" button and selecting "Download ZIP" from the dropdown. Unzip the folder once the download is complete.



mattjbayly steelhead profile

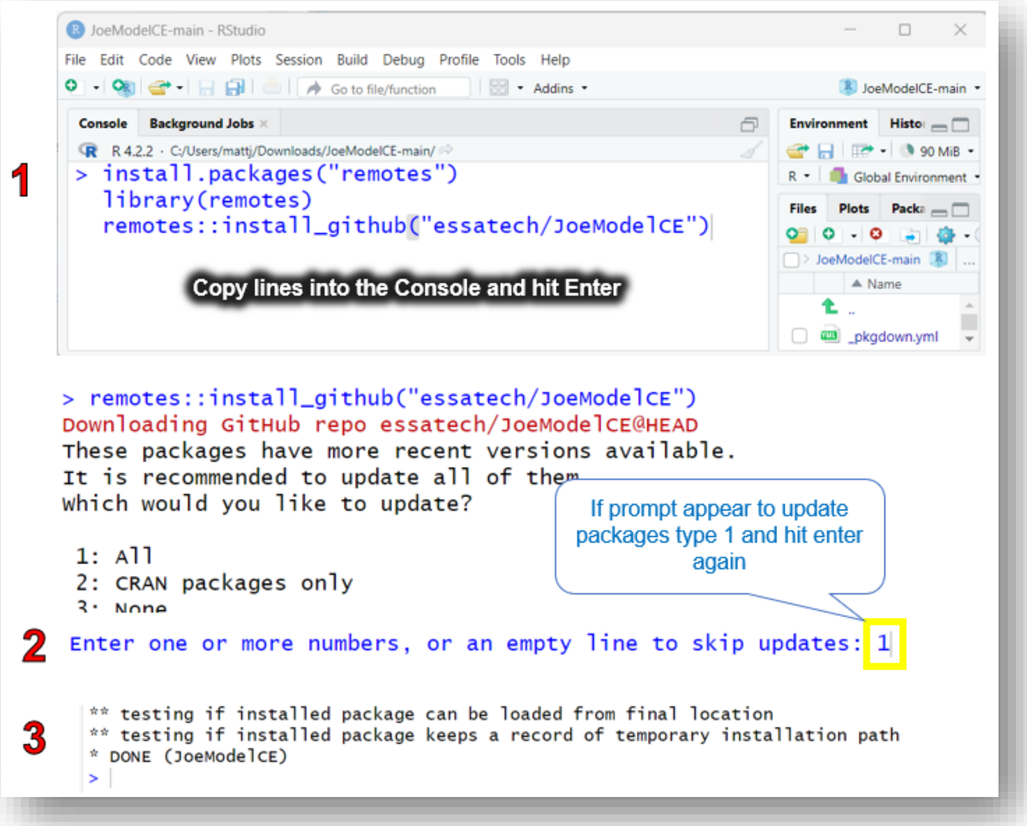
.github	ghp try 2	
R	steelhead profile	
inst	steelhead profile	
man	steelhead profile	
tests	steelhead profile	
vignettes	steelhead profile	
.Rbuildignore	merge batch func	
.gitignore	ghp try 2	last month

`https://github.com/essatech/JoeModelCE.git`

Use Git or checkout with SVN using the web URL.

2. Install the *JoeModelCE* R Package. The easiest way to install the *JoeModelCE* package is from within RStudio using `remotes::install_github()`. At this time, the package has not been published to CRAN, so the default `install.packages()` will not work for installing the *JoeModelCE* package. Instead, use the *remotes* package (or *devtools*). Open RStudio and install the *remotes* package using the `install.packages("remotes")` command in the Console. Next, install the *JoeModelCE* package from GitHub using the following commands in the Console:

   ```
   install.packages("remotes")
   library(remotes)
   remotes::install_github("essatech/JoeModelCE")
   ```



```
> install.packages("remotes")
  library(remotes)
  remotes::install_github("essatech/JoeModelCE")
```

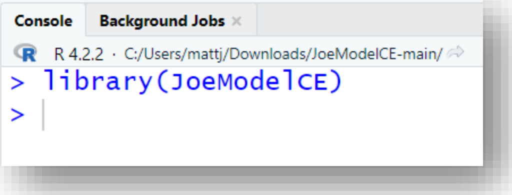
Copy lines into the Console and hit Enter

```
> remotes::install_github("essatech/JoeModelCE")
Downloading GitHub repo essatech/JoeModelCE@HEAD
These packages have more recent versions available.
It is recommended to update all of them.
Which would you like to update?

 1: All
 2: CRAN packages only
 3: None

Enter one or more numbers, or an empty line to skip updates: 1
```

If prompt appear to update packages type 1 and hit enter again

```
** testing if installed package can be loaded from final location
** testing if installed package keeps a record of temporary installation path
* DONE (JoeModelCE)
>
```

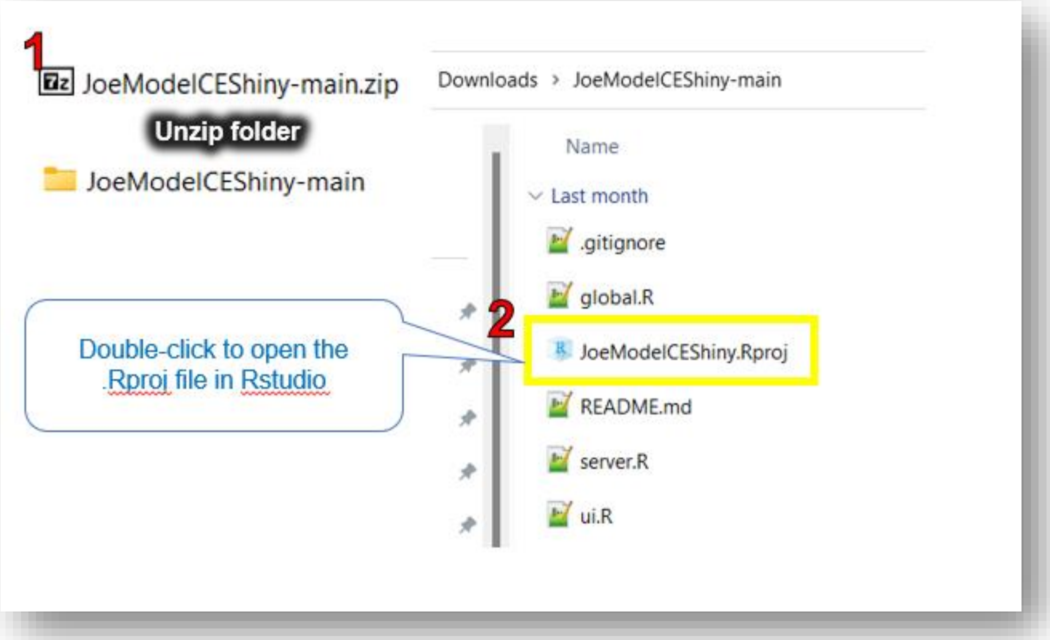
3. Once installed, use the `library(JoeModelCE)` command to call the *JoeModelCE* package into RStudio. You should only have to do the above steps once on your computer.



```
> library(JoeModelCE)
>
```

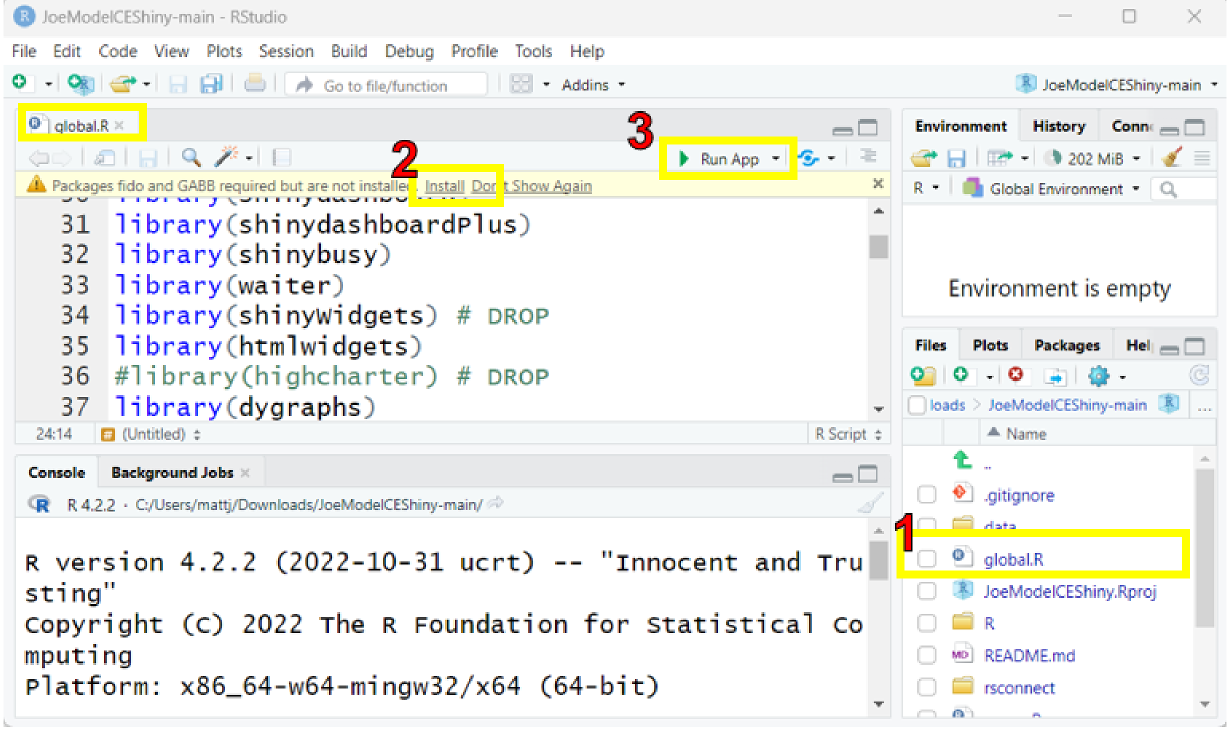
To access the raw code and example data for the R Shiny application:

1. Download the *JoeModelCEShiny* repository from GitHub (https://github.com/essatech/JoeModelCEShiny) onto your computer by clicking the green “Code” button and selecting “Download ZIP” from the dropdown. Unzip the folder once the download is complete.
2. Open the .Rproj file in R-Studio by double-clicking on it.



Double-click to open the Rproj file in Rstudio

3. Open a script called global.R by double-clicking on it in the bottom right corner. Next, click the ‘Install’ link on the yellow banner to install additional dependencies. Once complete, click the green arrow labelled ‘Run App’ to launch the application.



Packages fido and GABB required but are not installe

```
library(shinydashboardPlus)
library(shinybusy)
library(waiter)
library(shinyWidgets) # DROP
library(htmlwidgets)
#library(highcharter) # DROP
library(dygraphs)
```

```
R version 4.2.2 (2022-10-31 ucrt) -- "Innocent and Tru
sting"
Copyright (C) 2022 The R Foundation for Statistical Co
mputing
Platform: x86_64-w64-mingw32/x64 (64-bit)
```

4. Within the *JoeModelCEShiny* repository, example datasets are stored in the “demo” subfolder within the “data” folder: https://github.com/essatech/JoeModelCEShiny/tree/main/data/demo

You only need to do the above steps once on your computer. Next time you need to launch the application, simply click on the JoeModelCEShiny.Rproj file to open it in R Studio, then click on the green ‘Run App’ button or simply type ‘shiny::runApp()’ in the console to launch the application.

(Advanced) Using Windows command line to clone a GitHub repository:

For users comfortable using the command line who wish to contribute to the project, GitHub repositories can be quickly cloned into a local directory using this alternative method:

*Note: Prior to using Git commands in the command line, you must download and install Git on your computer (link).

1. Navigate to the desired GitHub repository (*JoeModelCE* R package: https://github.com/essatech/JoeModelCE/; *JoeModelCEShiny* application: https://github.com/essatech/JoeModelCEShiny).
2. Click the green “Code” button and copy the URL from the HTTPS tab.
3. Open the Windows Command Prompt window on your computer. Set your working directory using the following command:

   ```
   cd “<file path>”
   ```

   For example:

   ```
   cd “C:\Users\username\Documents”
   ```

4. Next, use the git clone command to clone the GitHub repository into a folder in your working directory. For example:

   ```
   git clone https://github.com/essatech/JoeModelCE.git
   ```

5 Data Inputs

A working example of all data inputs can be downloaded from a ‘demo’ directory with the JoeModelCEShiny project repository:
https://github.com/essatech/JoeModelCEShiny/tree/main/data/demo
Download the repository locally and navigate to the demo folder: .\JoeModelCEShiny-main\data\demo:

- stressor_response_demo.xlsx
- stressor_magnitude_demo.xlsx

- watersheds.gpkg
- life cycles.csv

5.1 Stressor-Response Workbook

5.1.1 Purpose

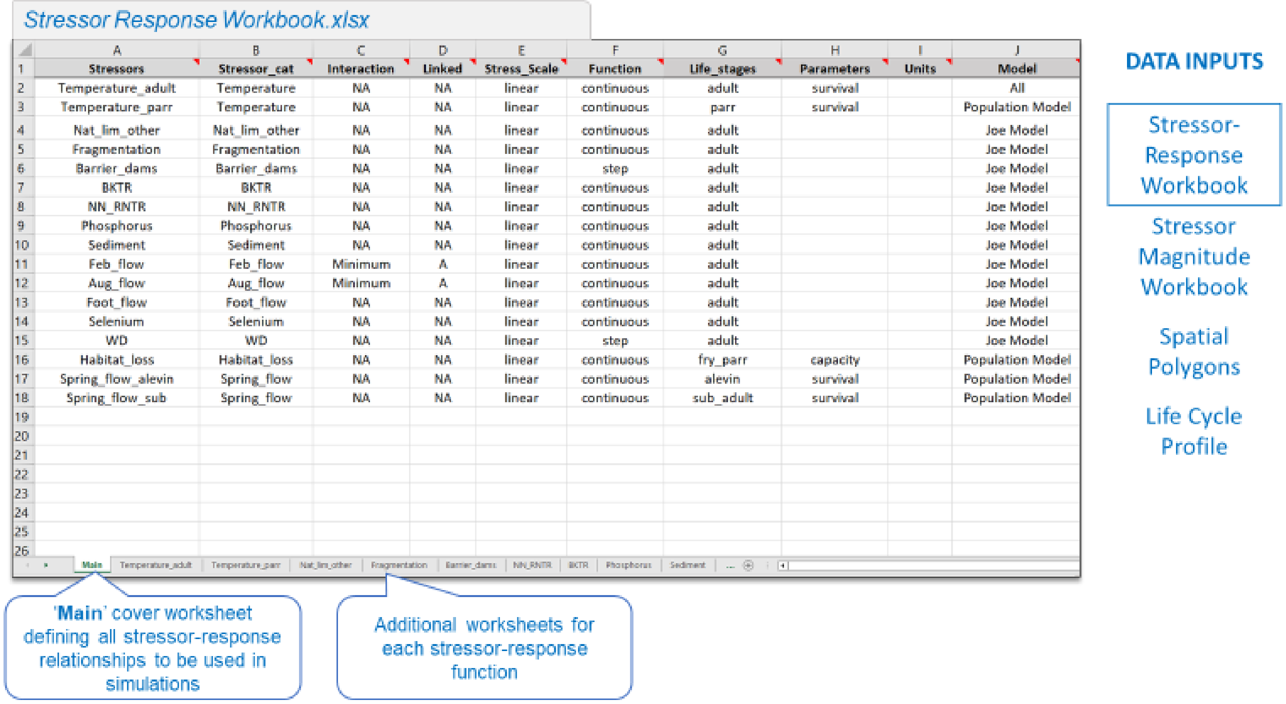
The stressor-response workbook contains all the stressor-response curves applicable to a target study system (e.g., Athabasca Rainbow Trout). These stressor-response curves are used within the CEMPRA (Joe Model) tool to predict cumulative effects additively given stressor magnitude values (discussed in the next section).

5.1.2 Layout

The stressor-response workbook is an Excel workbook containing all stressor-response functions to be used in the CEMPRA (Joe Model). The first worksheet contained within this Excel workbook must be titled "Main". This worksheet is used as a coversheet to describe and organize each of the stressor-response curves. Subsequent worksheets describe each of the stressor-response functions relevant to a particular species, where each stressor-response function has its own worksheet. Note that the spelling of the stressor name must be identical between the "Stressors" column in the "Main" worksheet and the worksheet title (on the bottom tab) for each stressor.

Work is currently underway to develop an online stressor-response web database (a digital repository of stressor-response functions across species, systems and geographic areas compiled across reference literature. As this database grows, functionality will be expanded to automatically generate a stressor-response workbook of selected stressors using the R-package (JoeModelCE). See details in Appendix A.

Main Worksheet



Stressors	Stressor_cat	Interaction	Linked	Stress_Scale	Function	Life_stages	Parameters	Units	Model
Temperature_adult	Temperature	NA	NA	linear	continuous	adult	survival		All
Temperature_parr	Temperature	NA	NA	linear	continuous	parr	survival		Population Model
Nat_lim_other	Nat_lim_other	NA	NA	linear	continuous	adult			Joe Model
Fragmentation	Fragmentation	NA	NA	linear	continuous	adult			Joe Model
Barrier_dams	Barrier_dams	NA	NA	linear	step	adult			Joe Model
BKTR	BKTR	NA	NA	linear	continuous	adult			Joe Model
NN_RNTR	NN_RNTR	NA	NA	linear	continuous	adult			Joe Model
Phosphorus	Phosphorus	NA	NA	linear	continuous	adult			Joe Model
Sediment	Sediment	NA	NA	linear	continuous	adult			Joe Model
Feb_flow	Feb_flow	Minimum	A	linear	continuous	adult			Joe Model
Aug_flow	Aug_flow	Minimum	A	linear	continuous	adult			Joe Model
Foot_flow	Foot_flow	NA	NA	linear	continuous	adult			Joe Model
Selenium	Selenium	NA	NA	linear	continuous	adult			Joe Model
WD	WD	NA	NA	linear	step	adult			Joe Model
Habitat_loss	Habitat_loss	NA	NA	linear	continuous	fry_parr	capacity		Population Model
Spring_flow_alevin	Spring_flow	NA	NA	linear	continuous	alevin	survival		Population Model
Spring_flow_sub	Spring_flow	NA	NA	linear	continuous	sub_adult	survival		Population Model

'**Main**' cover worksheet defining all stressor-response relationships to be used in simulations

Additional worksheets for each stressor-response function

Figure 8: Main worksheet in the Stressor-Response workbook.

The purpose of the "Main" worksheet is to organize and summarize stressor-response functions within the workbook. The columns within this worksheet are set up as follows. Inputs must be diligently followed in order for the code to work:

Column Name	Description
Stressors	Name of the stressor. This must match the stressor-response worksheet title. It must also match the "Stressor" column in the stressor magnitude workbook. Avoid the use of spaces in stressor names.
Stressor_cat	Category of the stressor. Only relevant if multiple stressors are linked with a defined interaction. If there is no interaction, simply copy the stressor name here.
Interaction	Interaction between stressors. Set to NA or leave blank if there are no interactions. Possible interactions available include "Minimum". If multiple variables are linked together with "Minimum", the variable with the lowest mean system capacity in the Joe Model will be used, and the other variables will be omitted from the CEMPRA (Joe Model) calculation. This also helps determine which terms should be linked together by defining groups in the "Linked" column. Use letters A, B, C etc., to define

	distinct groups. If no special interaction is defined between variables, then set these cells to NA.
Linked	For Variable Linkages: Use NA if no interaction is defined for the target variable; otherwise, choose letters A, B or C to specify variable groups. For example, if there were four stressor-response curves for December, January, February, and March flows, and they were encoded as three separate variables, users could apply the “Minimum” function to link the lowest system capacity across all four terms but only have one winter flow function in the CEMPRA (Joe Model). In this case, the letter A would be used to group all these terms together.
Stress_Scale	Stressor Scale can be either “linear” or “log” to specify a linear or logarithmic function, respectively, for interpolation.
Function	Either “continuous” or “step”. “Continuous” will apply linear interpolation between values, and “step” will adjust mean system capacity values in discrete steps.
Life_stages	Which life stage should each stressor be linked to? **Set all values to “adult” if unspecified**. The default assumption in the Joe Model is that stressor-response curves are linked to “adult” system capacity. If you are making use of the integrated life cycle model, then possible life stages include **“SE”** for eggs and **“S0”** for hatchling/fry. Both SE and S0 are Age-0 individuals. For subsequent stages, use “**stage_1”**, “**stage_2”**, “**stage_3”** etc. for stage-specific linkages (see Section 6.4 and Section 7 for clarification). If the same stressor is linked to multiple life stages, special terms such as “**sub_adult**” or “**adult**” can be used to define linkages to all immature (Age-1+) stage classes and all mature (Age-1+) stage classes respectively. The term “**all**” will mean that the stressor response curve is linked to all life stages. ***If the same environmental variable is linked to different life stages with different stressor-response curves, then it is recommended that the user duplicate stressor values and treat them as distinct environmental variables to avoid confusion (e.g., “Temperature_adult”, “Temperature_parr”).
Parameters	***Only relevant for the integrated life cycle model: leave blank for scenarios that only make use of the Joe Model.*** If the stressor-response function is being linked to a vital rate in the life cycle model, describe how the stressor-response curve is linked

to the life stage (which vital rate is affected). This column can be left blank if a user is only interested in running the Joe Model

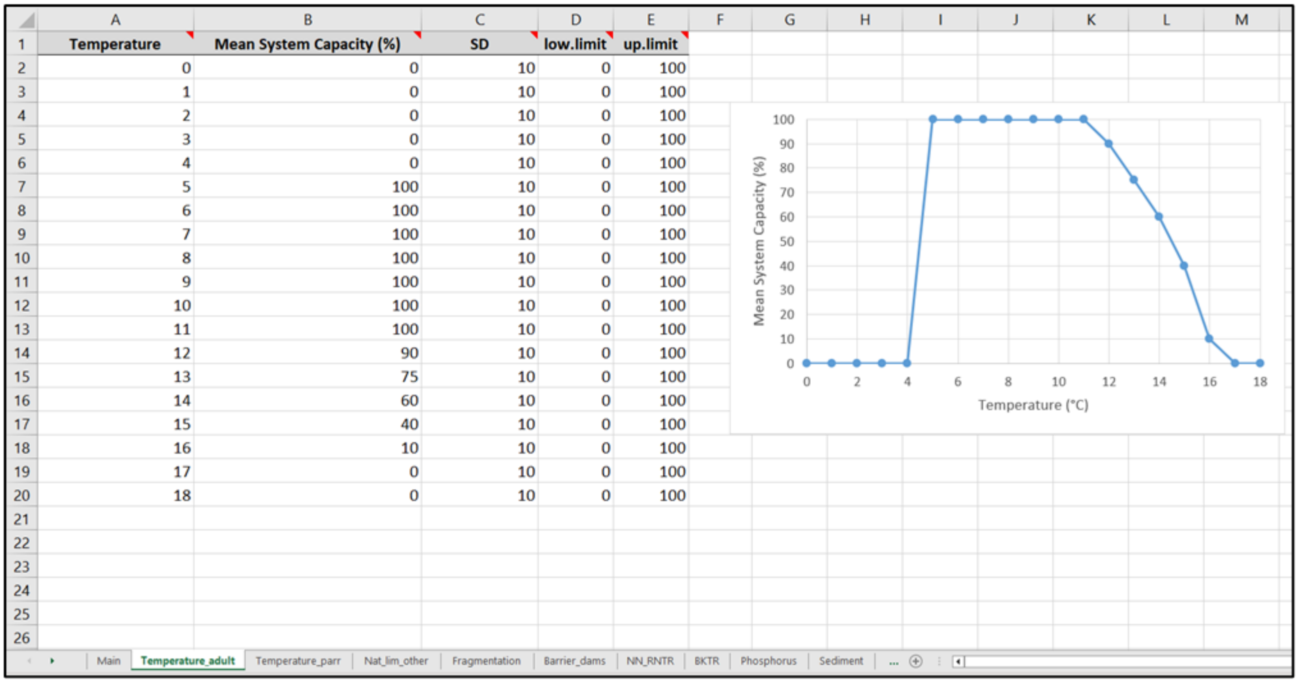
Possible mechanisms include “**capacity**”, “**survival,**” or “**fecundity**”. Once set, the response component of the stressor-response curve will act as a multiplier to the specified vital rate. For example, if a stressor is linked to “survival” of stage X and the stressor-response curve estimates a scaled response of 0.8 based on the stressor magnitude, then the default baseline survivorship will be multiplied by 0.8 in the simulation (e.g., original survivorship of life stage X: 0.34; adjusted survivorship of life stage X: 0.8*0.34 = 0.27).

- **survival**: Stressors linked to “survival” modify the default survivorship of a given life stage transition.
- **capacity**: Stressors linked to “capacity” will modify the stage-specific capacity values by a multiplier (0 – 1) based on the stressor-response relationship, regardless of the mechanism used to represent density-dependent constraints on growth (*compensation ratios vs location and stage-specific K values, See Section 7*). Capacity adjustments will be implemented as the final step in the calculation. Stressor-response functions linked to capacity will only have an effect if the life stage is parameterized with density-dependent constraints (e.g., compensation ratios ≠ 1.0 or Beverton-Holt K values enabled). If a stressor-response function is linked to a life stage, but there are no density-dependent constraints on that life stage in the life cycle profile, then the stressor will not have an effect on the simulation.
- **fecundity:** Stressor-response functions linked to fecundity will adjust the (eps) eggs per spawner (female) using the response value as a multiplier on the default input fecundity. For example, if the default fecundity is 3,000 and a stressor linked to fecundity has an effect of 0.6 (response), then the resulting fecundity (epf) in the life cycle model will be 3,000 * 0.6 = 1,800. Note that the stressor-response relationships must be linked to a sexually mature life stage for the fecundity multiplier to be meaningful. For example, if a fecundity multiplier is linked to an early life stage (e.g., stage_1), but the target species does not become mature until stage_3, then the stressor will have no observable effect on the population.

	Response (effect) estimates from the stressor-response curves will always be adjusted to 0 if biological response values are below zero, and adjusted to 100 (1.0) if response values exceed 100% in the stressor-response workbook. This ensures that survivorship remains between 0 and 1.
Units	**Optional:** include units as meta data (a friendly reminder). This column is not used in any calculations but is included as metadata and displayed on summary plots in the tool (where available).
Model	***Set all values to “Joe Model” if not using the life cycle model.*** This column is used to define which assessment endpoint the stressor-response curve should be applied to. Possible options include “All”, “Joe Model,” or “Life cycle model”. “All” specifies that the stressor-response curve is generic enough to be used in both the Joe Model and the life cycle model. “Joe Model” specifies that the stressor-response relationship should only be used in the simplified Joe Model roll-up summary. “Life cycle model” specifies that the stressor-response relationship is only applicable to the life cycle model.

Individual Stressor-Response Curve Worksheets

The remaining worksheets in the Stressor-response workbook are all used to describe the relationships between raw stressor values (on the x-axis) and the biological response (on the y-axis) (i.e., the stressor-response data).



Temperature	Mean System Capacity (%)	SD	low.limit	up.limit
0	0	10	0	100
1	0	10	0	100
2	0	10	0	100
3	0	10	0	100
4	0	10	0	100
5	100	10	0	100
6	100	10	0	100
7	100	10	0	100
8	100	10	0	100
9	100	10	0	100
10	100	10	0	100
11	100	10	0	100
12	90	10	0	100
13	75	10	0	100
14	60	10	0	100
15	40	10	0	100
16	10	10	0	100
17	0	10	0	100
18	0	10	0	100

Figure 9. Example of an individual stressor worksheet within the stressor response Excel workbook.

Individual stressor worksheets within the stressor-response file contain stressor-response curves for each of the stressor-response functions outlined in the “Main” worksheet (one worksheet for each row). **The spelling of the worksheet name must exactly match the spelling of the stressor on the ‘Main’ worksheet.** Additional rows can be added (as needed) to increase the resolution and specify the shape of a complex non-linear stressor response curve. Threshold stressor-response functions or stressor-response functions with discrete values may have relatively few rows.

The columns within each stressor-response worksheet are formatted as follows:

Column Name	Description
[Stressor]	The raw value of the stressor (on the x-axis). For example, the temperature would likely be valued in degrees Celsius, but the units and additional metadata are declared on the ‘Main’ worksheet.
Mean System Capacity (%)	The Mean System Capacity (on the y-axis) is associated with the raw stressor value on the x-axis. This column is the response component of the stressor-response curve. Values should be entered as a percentage and range from 0 to 100. Note that, for the life cycle model, this value may be the life-stage-specific dose-response curve for capacity or survival. *Mean system capacity is user as a generic term across the Joe Model and Life Cycle Model.*

SD	Standard Deviation (SD) is used to resample the mean system capacity (response) values based on a given stressor level. In the CEMPRA (Joe Model) simulations, environmental parameters are resampled for each batch replicate, year and location based on SD values in the stressor magnitude workbook. This column is used to represent uncertainty in the stressor-response relationship. Resampling will be based on the mean system capacity and SD values with a normal distribution (linear) or log scale (log) according to the value specified in the main worksheet under the Stress_Scale column. Regardless of the parameters for resampling, the lower limit and upper limit values will constrain values to fixed limits. If the SD value is set to zero, then no resampling will take place within the stressor-response curve.
low.limit	The lower limit for stressor-response resampling (see SD column distribution). Set to 0 as a default.
up.limit	The upper limit for stressor-response resampling (see SD column distribution). Set to 100 as a default.

5.1.3 How to build your own stressor-response function

CEMPRA (Joe Model) users can either select pre-assembled stressor-response functions or define their own stressor-response functions for use in the model. A library is currently being developed to host pre-assembled stressor-response functions and their associated documentation (Appendix A). Users who develop their own stressor-response functions are encouraged to fill out the appropriate documentation for their function and upload it to this public stressor-response library for future use (Appendix A). Please see MacPherson et al. (2018) and Rosenfeld et al. 2022 for further discussion on fundamental considerations in the development of customized stressor-response functions.

Methodology

Depending on data availability, stressor-response functions may be developed from available empirical data, from the elicitation of experts and stakeholders, or from existing literature.

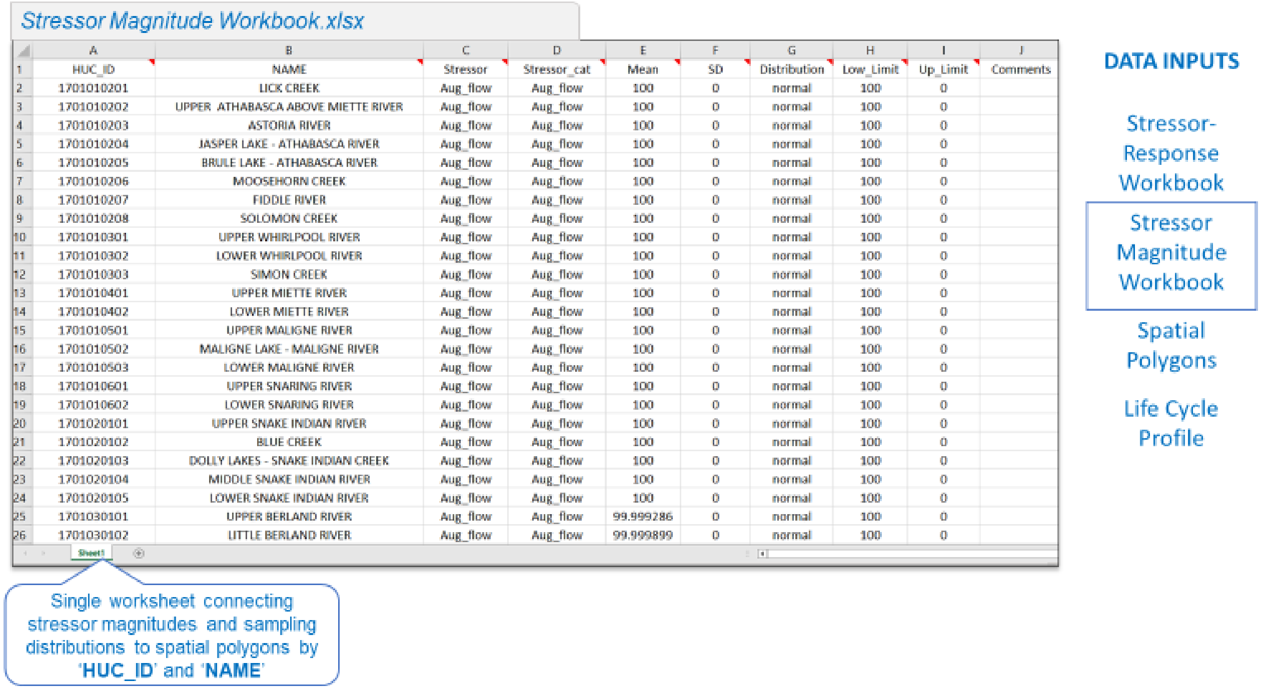
When using empirical data, the response component of the stressor-response relationship must be standardized to a common response metric. For convenience, the CEMPRA tool uses *Mean System Capacity* as a generic nickname to represent a standardized response scaled from 0% to 100%. In the CEMPRA tool (both the Joe Model and Life Cycle Model assessment endpoints), stressor-response relationships are reported as a scaled response value from 0% to 100%. These values are reported in the stressor-response workbook. The default assumption is that when a stressor is at its lowest level (or least harmful state if the scale is reversed), the mean system capacity will be set to 100% for that stressor. If this adjustment is not made, scenarios will be ranked differently simply by including specific stressors for comparison regardless of ecological state.

5.2 Stressor Magnitude File

5.2.1 Purpose

The stressor magnitude file is an Excel worksheet which defines and bounds each stressor within the individual locations (i.e., spatial units) being assessed. Similar to the stressor-response functions, stressor magnitude values are sampled across locations with stochasticity. The stressor magnitude workbook is structured accordingly, with data for each stressor and location entered into the dataset in a long table format (as opposed to the standard wide table format). Each row specifies a unique stressor for each unique location. Therefore, the number of rows in this dataset should be equal to the number of stressors multiplied by the number of locations. For each simulation in the CEMPRA tool, values are sampled at random with stochasticity for each stressor and location. For each year and batch replicate, stressor magnitude values will be drawn from each normal (or lognormal) distribution based on the Mean and SD (standard deviation) and then further constrained based on the specified lower and upper limits.

If there is no suspected uncertainty or interannual variability in stressor magnitude values for a given stressor and/or location, then the SD value can be set to 0, and the lower limit, upper limit and distribution type can be ignored.



Stressor Magnitude Workbook.xlsx

	A	B	C	D	E	F	G	H	I	J
1	HUC_ID	NAME	Stressor	Stressor_cat	Mean	SD	Distribution	Low_Limit	Up_Limit	Comments
2	1701010201	LICK CREEK	Aug_flow	Aug_flow	100	0	normal	100	0	
3	1701010202	UPPER ATHABASCA ABOVE MIETTE RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
4	1701010203	ASTORIA RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
5	1701010204	JASPER LAKE - ATHABASCA RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
6	1701010205	BRULE LAKE - ATHABASCA RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
7	1701010206	MOOSEHORN CREEK	Aug_flow	Aug_flow	100	0	normal	100	0	
8	1701010207	FIDDLE RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
9	1701010208	SOLOMON CREEK	Aug_flow	Aug_flow	100	0	normal	100	0	
10	1701010301	UPPER WHIRLPOOL RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
11	1701010302	LOWER WHIRLPOOL RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
12	1701010303	SIMON CREEK	Aug_flow	Aug_flow	100	0	normal	100	0	
13	1701010401	UPPER MIETTE RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
14	1701010402	LOWER MIETTE RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
15	1701010501	UPPER MALIGNE RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
16	1701010502	MALIGNE LAKE - MALIGNE RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
17	1701010503	LOWER MALIGNE RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
18	1701010601	UPPER SNARING RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
19	1701010602	LOWER SNARING RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
20	1701020101	UPPER SNAKE INDIAN RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
21	1701020102	BLUE CREEK	Aug_flow	Aug_flow	100	0	normal	100	0	
22	1701020103	DOLLY LAKES - SNAKE INDIAN CREEK	Aug_flow	Aug_flow	100	0	normal	100	0	
23	1701020104	MIDDLE SNAKE INDIAN RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
24	1701020105	LOWER SNAKE INDIAN RIVER	Aug_flow	Aug_flow	100	0	normal	100	0	
25	1701030101	UPPER BERLAND RIVER	Aug_flow	Aug_flow	99.999286	0	normal	100	0	
26	1701030102	LITTLE BERLAND RIVER	Aug_flow	Aug_flow	99.999899	0	normal	100	0	

Single worksheet connecting stressor magnitudes and sampling distributions to spatial polygons by 'HUC_ID' and 'NAME'

Figure 10: Example of a stressor magnitude workbook.

5.2.2 Layout

Each row in the stressor magnitude workbook specifies a relationship between a unique stressor and a unique location. Locations (discussed in detail in the next section) are specified by a unique ID (HUC_ID) and NAME. The HUC_ID column is a legacy from an older version of the Joe Model that referenced Hydrological Unit Codes as ID values, but HUC_ID can be any set of unique IDs specified by the user for their spatial units of interest. The NAME column can be blank but is included for convenience since many users find it challenging to cross-reference ID values between different datasets.

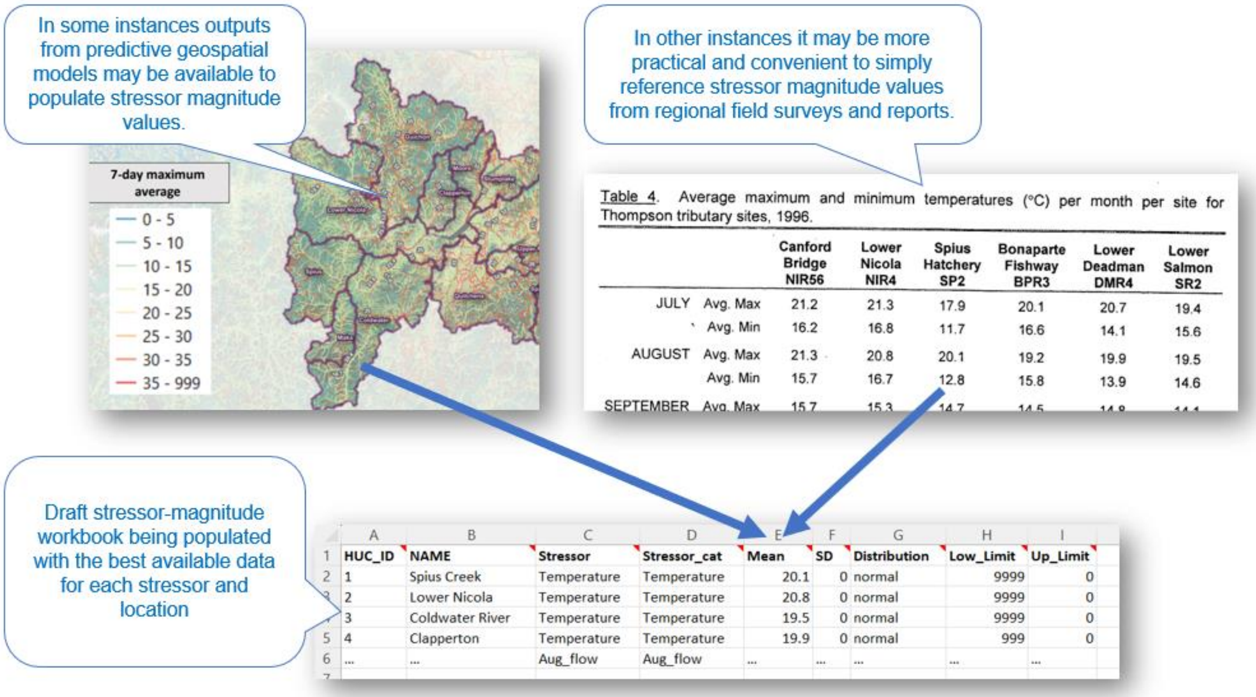
Column Name	Description
HUC_ID	An ID field is used to represent a unique location (spatial unit). The HUC_ID column is a legacy from an older version of the Joe Model that referenced Hydrological Unit Codes as ID values, but HUC_ID can be any set of unique IDs specified by the user for their spatial units of interest. The NAME column can be blank but is included for convenience since many users find it challenging to cross-reference ID values between different datasets. The HUC_ID field must match the feature column HUC_ID in the spatial polygons file.
NAME	The NAME of the polygon in the spatial polygons file. This field must match the NAME column in the spatial polygons file, but it can be blank if a spatial unit does not have a defining name.

Stressor	Stressor Name. This must match the information and sheet names in the stressor-response workbook.
Stressor_cat	Name of the Stressor category. This column must match the spelling used in the stressor-response workbook.
Mean	The mean value of the stressor for the spatial unit (location) HUC_ID (polygon).
SD	Standard deviation (SD) of stressor values for the target HUC_ID. During the simulation, stressor values for each HUC_ID are resampled from a distribution with mean and SD values. Setting the SD to zero means that there will be no variability in the stressor value during the simulation.
Distribution	Type of distribution to use for resampling. Either “normal” or “lognormal”. We recommend using a “normal” distribution where possible. Testing coverage is incomplete for “lognormal”.
Low_Limit	The lower limit for resampling.
Up_Limit	The upper limit for resampling.
Comments	Internal comments by the user for personal reference. It can be blank.

5.2.3 Assembling your own stressor-magnitude data

Data for stressor magnitude estimates can come from a variety of sources, including GIS data, modelled data, field data, expert opinion, regional trends, or estimates from the literature. When assembling your own stressor magnitude data, stressor magnitudes and ranges will need to be assigned to the individual locations (spatial units) being represented in the model. Stressors within each location must be assigned a magnitude, and distribution or the SD value must be set to 0 for simulations with no stochasticity. Within each spatial unit, users can specify the mean value, standard deviation, distribution (normal or lognormal), and upper and lower limits of each stressor, as discussed in the previous section.

The use of spatial units allows users to assign aggregated values from modelled data, field data, or from the literature, where high spatial accuracy may not be available.



In some instances outputs from predictive geospatial models may be available to populate stressor magnitude values.

In other instances it may be more practical and convenient to simply reference stressor magnitude values from regional field surveys and reports.

Table 4. Average maximum and minimum temperatures (°C) per month per site for Thompson tributary sites, 1996.

		Canford Bridge NIR56	Lower Nicola NIR4	Spius Hatchery SP2	Bonaparte Fishway BPR3	Lower Deadman DMR4	Lower Salmon SR2
JULY	Avg. Max	21.2	21.3	17.9	20.1	20.7	19.4
	Avg. Min	16.2	16.8	11.7	16.6	14.1	15.6
AUGUST	Avg. Max	21.3	20.8	20.1	19.2	19.9	19.5
	Avg. Min	15.7	16.7	12.8	15.8	13.9	14.6
SEPTEMBER	Avg. Max	15.7	15.3	14.7	14.5	[illegible]	[illegible]

Draft stressor-magnitude workbook being populated with the best available data for each stressor and location

	A	B	C	D	E	F	G	H	I
1	HUC_ID	NAME	Stressor	Stressor_cat	Mean	SD	Distribution	Low_Limit	Up_Limit
2	1	Spius Creek	Temperature	Temperature	20.1	0	normal	9999	0
3	2	Lower Nicola	Temperature	Temperature	20.8	0	normal	9999	0
4	3	Coldwater River	Temperature	Temperature	19.5	0	normal	9999	0
5	4	Clapperton	Temperature	Temperature	19.9	0	normal	999	0
6	...	...	Aug_flow	Aug_flow	...	...	...	...	...
7									

Figure 11: Example of assigning stressor magnitudes values (stream temperature) to spatial units (locations). Stressor magnitude values should be aggregated to locations to represent location-averaged estimates. Locations can be split and aggregated as needed such that each location represents averaged generalized conditions.

5.3 Locations (Spatial Polygons)

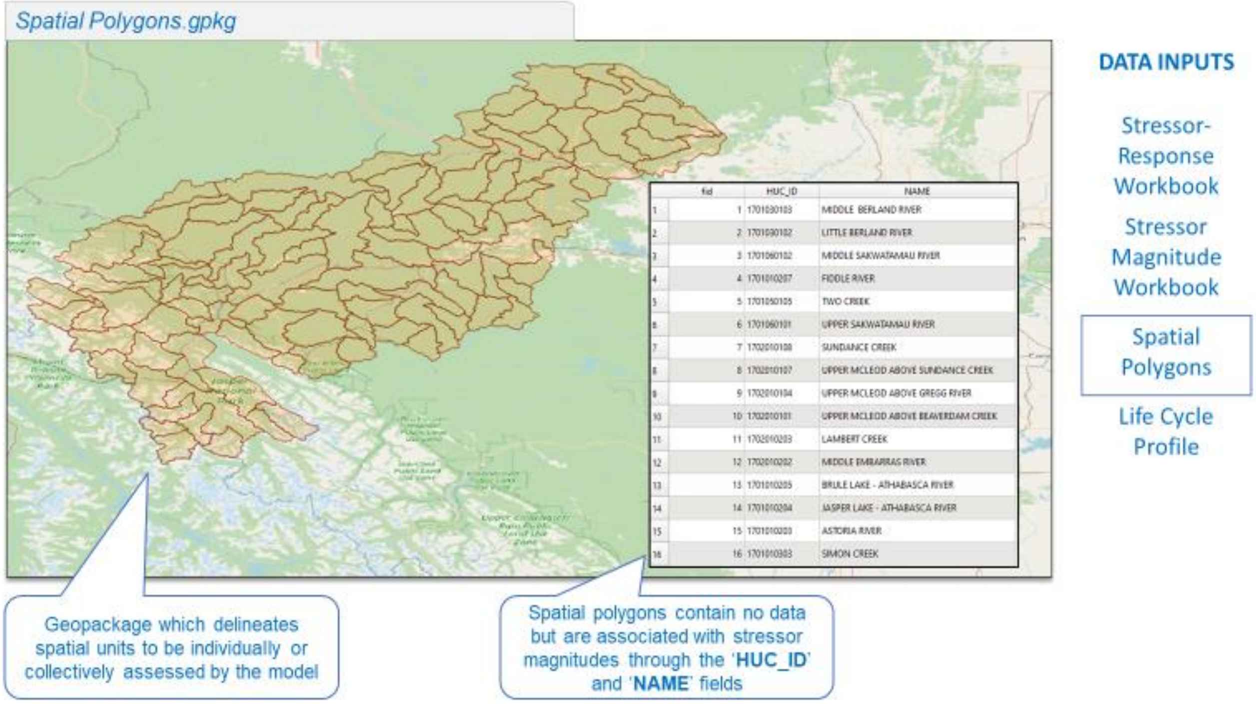
Locations are represented in the CEMPRA tool as spatial polygons. Locations should be defined to reflect heterogeneity in stressor values across the study area.

Spatial polygons are imported into the CEMPRA tool as a GIS data file. Possible formats include either geopackage (.gpkg) or shapefile (.shp) format. The spatial units GIS data file does not contain any stressor magnitude data for modelling but consists of just the geometry and fields for the location ID (HUC_ID) and NAME. The polygon geometry included in this file is used for display purposes only and is joined to the stressor magnitude file when imported into the CEMPRA. Therefore, the size and shape of each spatial unit do not influence any components of the assessment.

Regardless of the file format, the GIS (locations) spatial polygon data file must meet the following criteria:

- **"**HUC_ID" field: The locations data layer must have an attributed field labelled '**HUC_ID**' with a unique identifier for each spatial unit (location).
- "NAME" field: The locations data layer must have an attributed field labelled '**NAME**'**.** This field is included for convenience. Values in this field can be blank, but users are encouraged to populate this field.

- The HUC_ID and NAME fields must match values (and spelling) in the stressor magnitude Excel input so data can be cross-referenced and joined in the tool.
- The spatial polygons file must be imported in the standard latitude/longitude projection (CRS: “EPSG:4326 – WGS84”). External GIS software such as QGIS can be used to transform projections as needed.
- There should be no geometry errors (invalid geometries) in the polygon geometry data. Run check geometry and fix errors in programs such as QGIS.
- Simplify geometries: we recommended running functions like ‘Simplify geometries’ in programs such as QGIS to reduce the file size before importing data into the CEMPRA tool. Ideally, the locations spatial polygon data file should be under 10MB for the best performance in the Shiny application. Processing larger datasets is possible by using the R package version of the CEMPRA tool.



	fid	HUC_ID	NAME
1	1	1701030103	MIDDLE BERLAND RIVER
2	2	1701030102	LITTLE BERLAND RIVER
3	3	1701060102	MIDDLE SAKWATAMAU RIVER
4	4	1701010207	FIDDLE RIVER
5	5	1701050105	TWO CREEK
6	6	1701060101	UPPER SAKWATAMAU RIVER
7	7	1702010108	SUNDANCE CREEK
8	8	1702010107	UPPER MCLEOD ABOVE SUNDANCE CREEK
9	9	1702010104	UPPER MCLEOD ABOVE GREGG RIVER
10	10	1702010101	UPPER MCLEOD ABOVE BEAVERDAM CREEK
11	11	1702010203	LAMBERT CREEK
12	12	1702010202	MIDDLE EMBARRAS RIVER
13	13	1701010205	BRULE LAKE - ATHABASCA RIVER
14	14	1701010204	JASPER LAKE - ATHABASCA RIVER
15	15	1701010203	ASTORIA RIVER
16	16	1701010303	SIMON CREEK

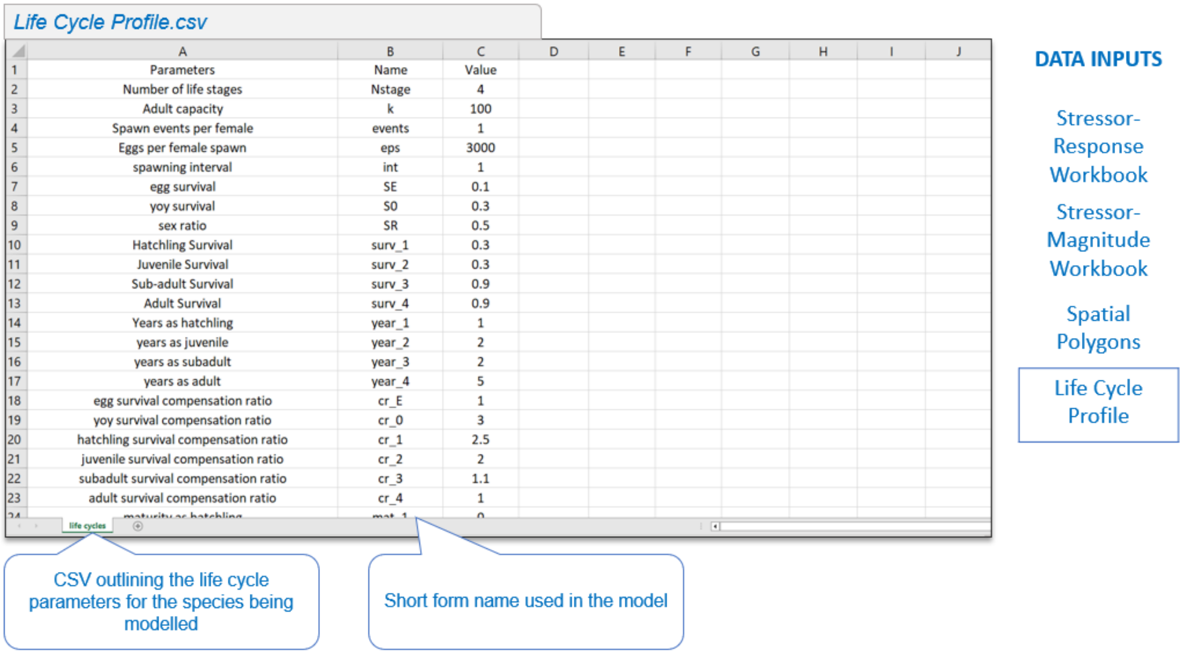
Geopackage which delineates spatial units to be individually or collectively assessed by the model

Spatial polygons contain no data but are associated with stressor magnitudes through the ‘**HUC_ID**’ and ‘**NAME**’ fields

Figure 12: Example of a spatial polygons (locations) file and its associated attribute table with fields HUC_ID and NAME.

5.4 Life Cycles Profile

The life cycles profile (csv file) is an optional input applicable to users who are interested in running the integrated life cycle model. The life cycles profile file specifies all input parameters required to run the life cycle model (e.g., number of stages, survival rates, fecundity, etc.). The format of the life cycles profile is a generic template, but once populated, it is used to parameterize the life cycle model for a specific study system. Usually, this consists of a target population (e.g., Athabasca Rainbow Trout, Nicola Basin Chinook Salmon etc.).



Life Cycle Profile.csv

	A	B	C
1	Parameters	Name	Value
2	Number of life stages	Nstage	4
3	Adult capacity	k	100
4	Spawn events per female	events	1
5	Eggs per female spawn	eps	3000
6	spawning interval	int	1
7	egg survival	SE	0.1
8	yoy survival	S0	0.3
9	sex ratio	SR	0.5
10	Hatchling Survival	surv_1	0.3
11	Juvenile Survival	surv_2	0.3
12	Sub-adult Survival	surv_3	0.9
13	Adult Survival	surv_4	0.9
14	Years as hatchling	year_1	1
15	years as juvenile	year_2	2
16	years as subadult	year_3	2
17	years as adult	year_4	5
18	egg survival compensation ratio	cr_E	1
19	yoy survival compensation ratio	cr_0	3
20	hatchling survival compensation ratio	cr_1	2.5
21	juvenile survival compensation ratio	cr_2	2
22	subadult survival compensation ratio	cr_3	1.1
23	adult survival compensation ratio	cr_4	1

CSV outlining the life cycle parameters for the species being modelled

Short form name used in the model

Figure 13. Example life cycles profile CSV file

A detailed discussion of the life cycles profile (csv file) is included in Section 7.2 with accompanying background information. It's possible to run the Joe Model and omit the life cycle model entirely. Therefore, the life cycles profile csv should be considered as an optional input for advanced use cases of the CEMPRA tool.

6 R-Shiny Application Walkthrough

The R Shiny web application is a user-friendly interface for the CEMPRA powered by the *JoeModelCE* R package. Users can access a working version R Shiny application hosted here: https://essa.shinyapps.io/JoeModelCEShiny. The following section is intended to walk new users through the components of the R Shiny web application.

6.1 About

When you first access the R Shiny web application for the CEMPRA, you will be automatically directed to the **About** page. This page contains a brief introduction to the application and the CEMPRA (Joe Model) tool, a list of contributors and a list of features for users to explore. A tutorial video is also embedded on this webpage to guide users through the application.

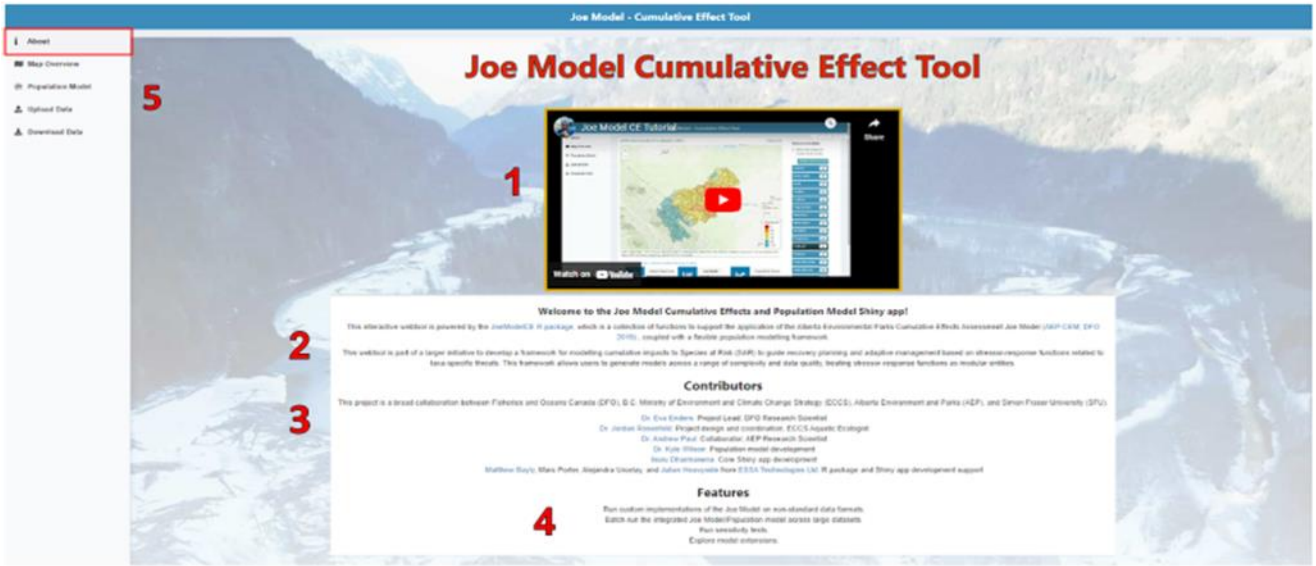
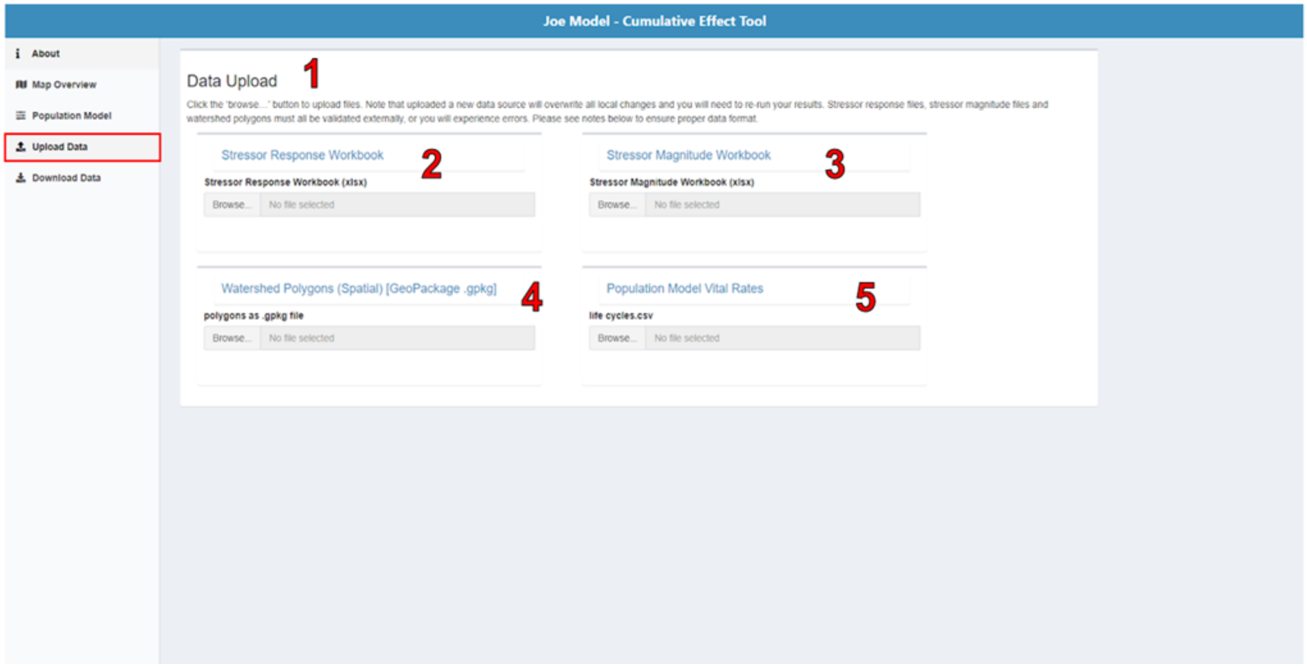


Figure 14. About page in the R Shiny web application. 1) Tutorial video for using the R Shiny application. 2) Brief introduction to the CEMPRA and associated life cycle model. 3) List of contributors to the CEMPRA. 4) List of features included in the R Shiny web application. 5) Menu for the R Shiny application

6.2 Upload Data

If you would like to import a custom set of stressor-response functions, spatial polygons, or vital rates (for the life cycle model), select the **Upload Data** page in the menu on the left side of the screen. This page allows you to upload four key input files from your computer: the stressor-response workbook (.xlsx), the stressor-magnitude workbook (.xlsx), a spatial polygons file (.gpkg or .shp), and a life cycle profile (.csv). To upload a file from your computer, click “Browse…” on the toolbar below the associated file heading. Next, select the desired file from the file window and click “Open”. The file will begin to upload. A status bar will appear below the file name, indicating the upload status. Once the file has been successfully uploaded, the status bar will display “Upload complete”.



Data Upload

Click the 'browse...' button to upload files. Note that uploaded a new data source will overwrite all local changes and you will need to re-run your results. Stressor response files, stressor magnitude files and watershed polygons must all be validated externally, or you will experience errors. Please see notes below to ensure proper data format.

Stressor Response Workbook (xlsx)

Stressor Magnitude Workbook (xlsx)

polygons as .gpkg file

life cycles.csv

Figure 15. Data upload page of the R Shiny web application. 1) Instructions for uploading files. 2) Location to upload stressor-response workbook (.xlsx). 3) Location to upload stressor magnitude workbook (.xlsx). 4) Location to upload spatial polygons file (.gpkg or .shp). 5) Location to upload life cycle profile (.csv).

If an error occurs, the file will not upload correctly, and a red box with an error message will appear below the file name. Note that stressor-response files, stressor magnitude files and watershed polygons must all be validated externally before uploading; otherwise, you will experience errors. In addition, required data formats are listed in the headings for each file (note: shapefiles are also accepted for the Watershed Polygons (Spatial) file). Please ensure that you are using the correct data formats, or you will experience errors.

Uploading a new data source will overwrite all local changes, and you will need to re-run your results if you have previously run the tool (see next section).

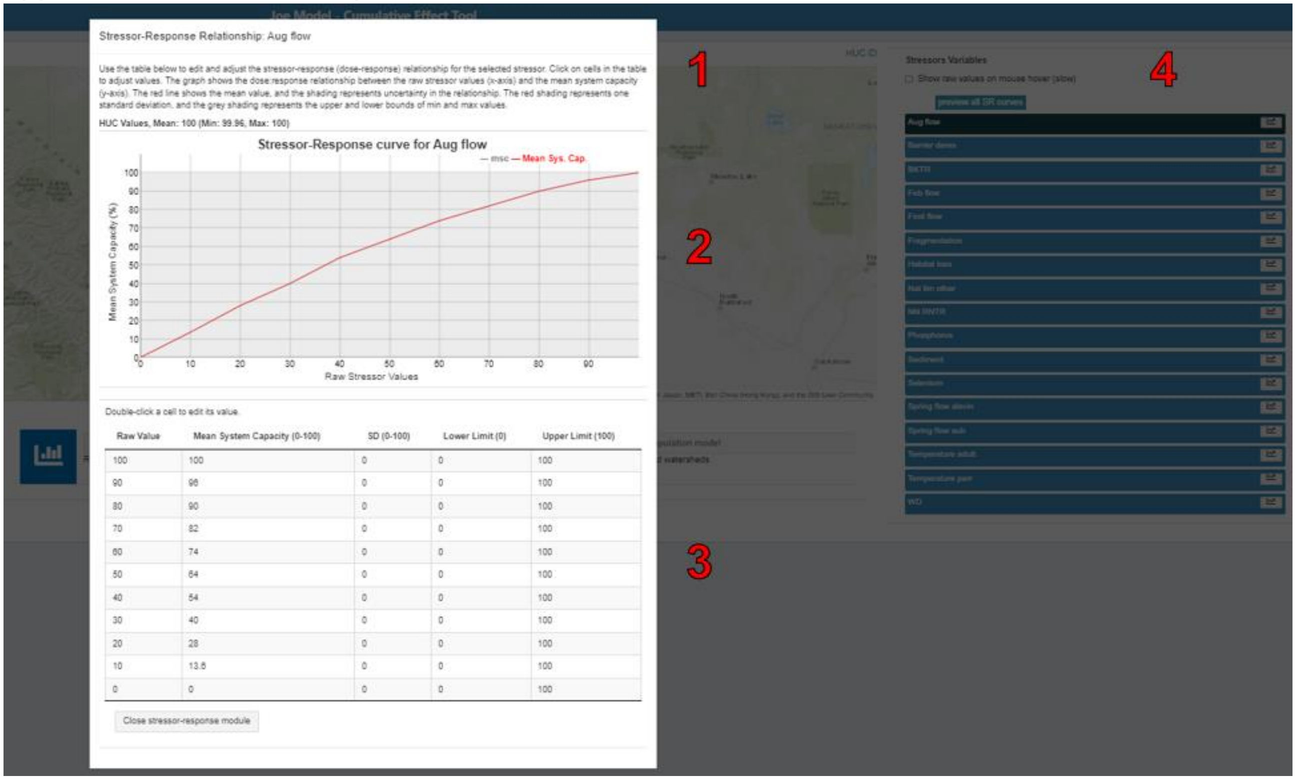
6.3 Map Overview

6.3.1 Map Window

The map window on the **Map Overview** page allows users to view stressor magnitudes and model results in a choropleth map based on the spatial polygons layer you imported. The map in the centre of the screen is interactive and allows users to scroll, zoom, and select individual polygons. When hovering over a polygon, users will see the name and HUC_ID of that polygon displayed directly above the map window. To select a polygon, click on it once. Selected polygons will appear in blue. Clicking on additional polygons will add to your selection. To deselect a polygon, click it again. The number of selected polygons (HUCs) will appear below the map window on the left-hand side. Click the “deselect all” button beside this text string to deselect all polygons.

6.3.2 Stressors

On the **Map Overview** page, users can view stressor-response relationships for each of the stressors included in the input stressor-response Excel workbook. A list of these stressors is on the screen's right side. Clicking on the chart icon 📈 next to one of the stressors opens a pop-up window to view the stressor-response curve and raw stressor-response table associated with that specific stressor. You can edit cell values within the stressor-response table by double-clicking on the cell you want to edit. Any edits to the stressor-response table will automatically appear in the stressor-response curve.



Stressor-Response Relationship: Aug flow

Use the table below to edit and adjust the stressor-response (dose-response) relationship for the selected stressor. Click on cells in the table to adjust values. The graph shows the dose:response relationship between the raw stressor values (x-axis) and the mean system capacity (y-axis). The red line shows the mean value, and the shading represents uncertainty in the relationship. The red shading represents one standard deviation, and the grey shading represents the upper and lower bounds of min and max values.

HUC Values, Mean: 100 (Min: 99.96, Max: 100)

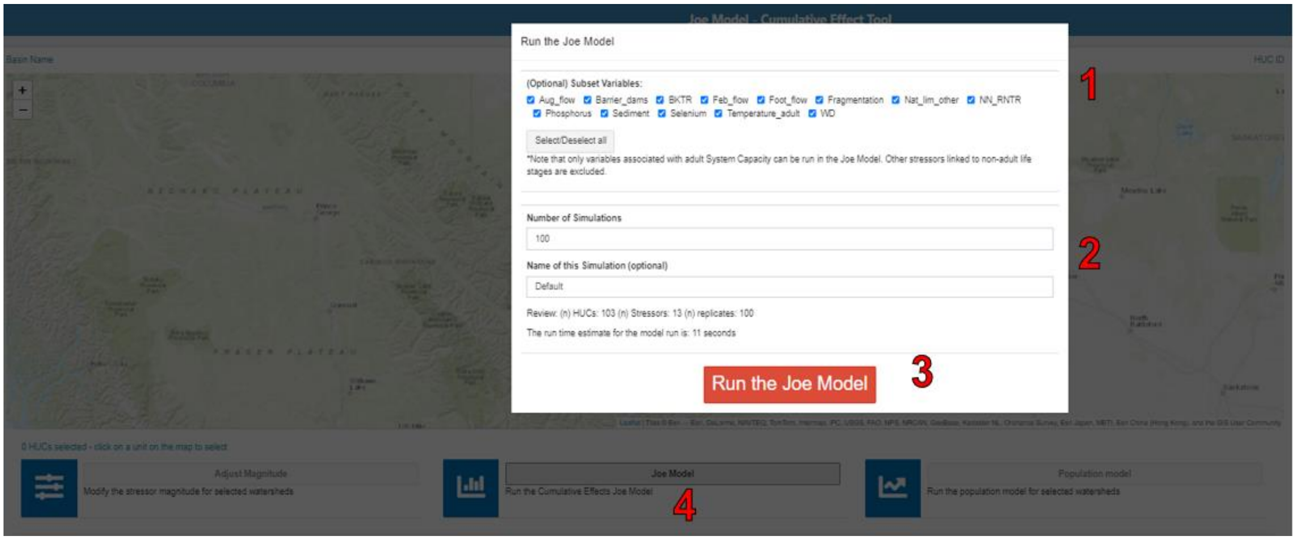
Double-click a cell to edit its value.

Raw Value	Mean System Capacity (0-100)	SD (0-100)	Lower Limit (0)	Upper Limit (100)
100	100	0	0	100
90	96	0	0	100
80	90	0	0	100
70	82	0	0	100
60	74	0	0	100
50	64	0	0	100
40	54	0	0	100
30	40	0	0	100
20	28	0	0	100
10	13.6	0	0	100
0	0	0	0	100

Figure 16: Stressor-response relationship pop-up window example. 1) Description of table and stressor-response curve. 2) Stressor-response curve plot. 3) Stressor-response curve input table. 4) List of stressors used in the CEMPRA (Joe Model).

6.3.3 Joe Model

To run the Joe Model, click the “Joe Model” button in the centre of the lower panel. In the pop-up window, you can select/deselect variables that you want to include/exclude in the Joe Model run. In addition, you can specify the number of simulations or batch runs you would like to conduct (in the “Number of Simulations” box), and you can provide a name for the simulation (in the “Name of this Simulation” box). When you are ready to run the Joe Model for all spatial polygons in the study area, click “Run the Joe Model”. Note there is currently no option to run the Joe Model for individual spatial polygons.



Run the Joe Model

(Optional) Subset Variables:

☑ Aug_flow ☑ Barrier_dams ☑ BKTR ☑ Feb_flow ☑ Foot_flow ☑ Fragmentation ☑ Nat_lim_other ☑ NN_RNTR ☑ Phosphorus ☑ Sediment ☑ Selenium ☑ Temperature_adult ☑ WD

Select/Deselect all

*Note that only variables associated with adult System Capacity can be run in the Joe Model. Other stressors linked to non-adult life stages are excluded.

Number of Simulations: 100

Name of this Simulation (optional): Default

Review: (n) HUCs: 103 (n) Stressors: 13 (n) replicates: 100

The run time estimate for the model run is: 11 seconds

Run the Joe Model

Figure 17: Pop-up window for the Joe Model. 1) Variables to include/exclude from cumulative effects modelling. 2) Inputs for the number of simulations and the name of the simulation. 3) Button to run the Joe Model. 4) Button to bring up the pop-up window.

6.3.4 System Capacity Plots

After running the Joe Model, a new section labelled “System Capacity Plots” will appear at the bottom of the map page. You can view tables and plots of the cumulative system capacity for selected watersheds (click “Selected watersheds”) or for all watersheds (click “All watersheds”) compared to the global mean system capacity across all simulations.

System capacity plots show the scaled probability output from the Joe Model from stochastic simulations.

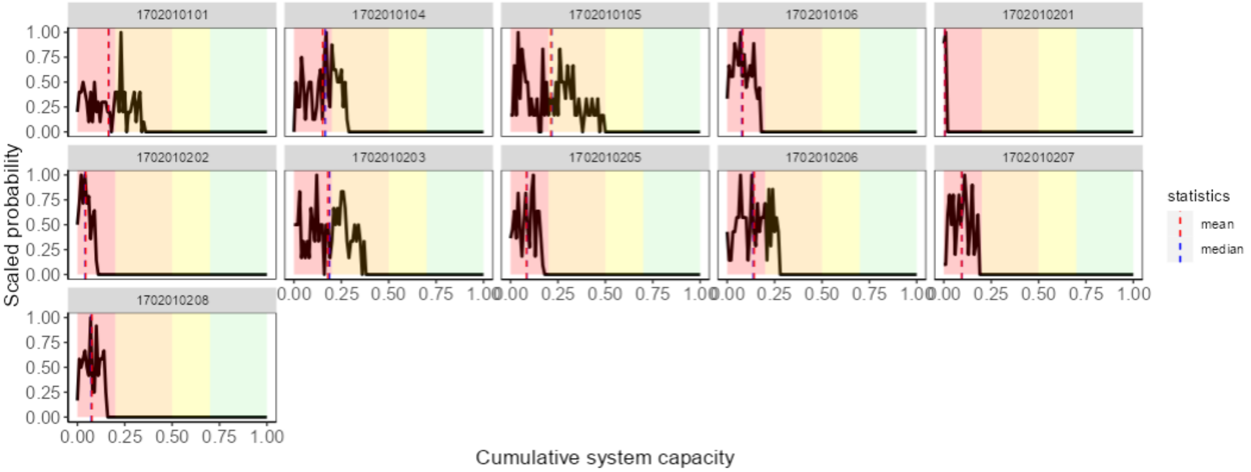
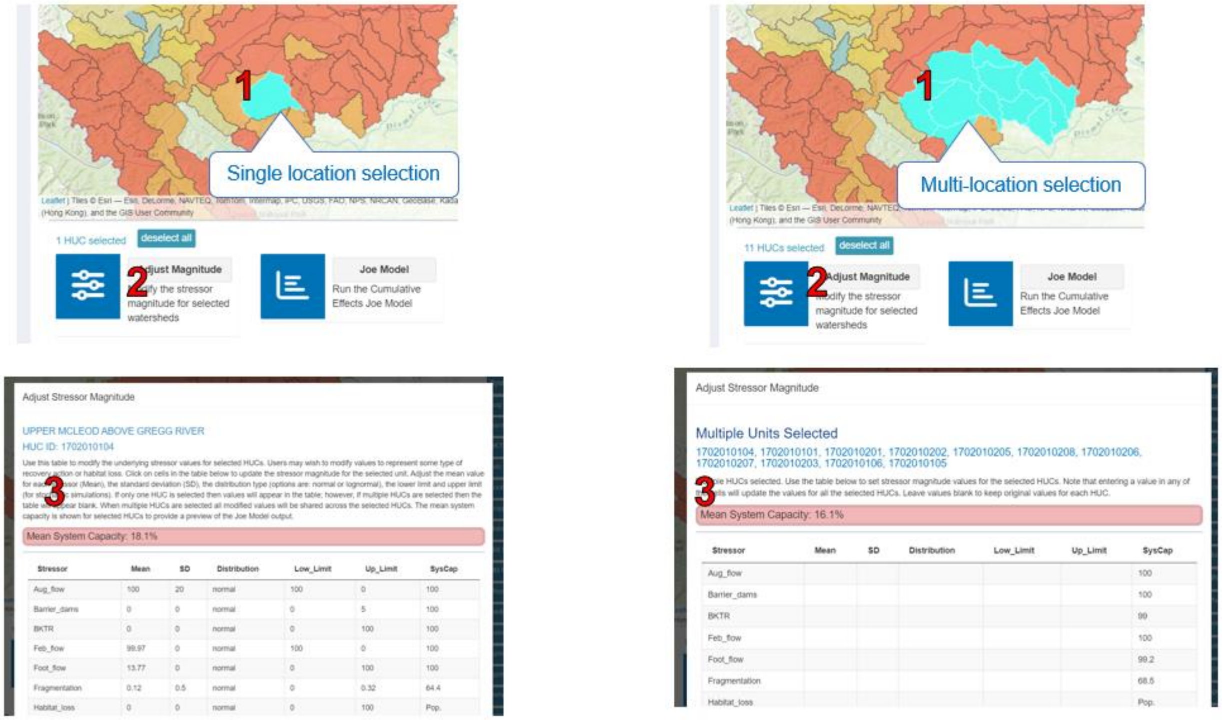


Figure 18. System capacity plots for eleven spatial units (locations) showing the cumulative system capacity scores (mean, median and scaled probability distributions).

6.3.5 Adjust Magnitude

Users may wish to modify stressor values to represent recovery action or habitat loss. To view and edit stressor magnitudes for an individual spatial polygon, select the spatial polygon of interest and click “Adjust Magnitude” on the right-hand side of the lower panel. Next, review the mean system capacity in the pop-up window for one or more selected locations. Below this, a table containing all the stressors, their magnitudes and distributions will be displayed. Double-click on the desired cell to edit any of the values in this table. Adjust the mean value for each stressor (Mean), the standard deviation (SD), the distribution type (options are: normal or lognormal), and the lower limit and upper limit (for stochastic simulations). Note that stressor names and system capacity values cannot be edited. If only one HUC is selected, values will appear in the table; however, if multiple HUCs are selected, the table will appear blank. When multiple HUCs are selected, all modified values will be shared across the selected HUCs. The mean system capacity is shown for selected HUCs to preview the model output.



Adjust Stressor Magnitude

UPPER MCLEOD ABOVE GREGG RIVER

HUC ID: 1702010104

Use this table to modify the underlying stressor values for selected HUCs. Users may wish to modify values to represent some type of recovery action or habitat loss. Click on cells in the table below to update the stressor magnitude for the selected unit. Adjust the mean value for ea[illegible]ssor (Mean), the standard deviation (SD), the distribution type (options are: normal or lognormal), the lower limit and upper limit (for sto[illegible]c simulations). If only one HUC is selected then values will appear in the table; however, if multiple HUCs are selected then the table w[illegible]ppear blank. When multiple HUCs are selected all modified values will be shared across the selected HUCs. The mean system capacity is shown for selected HUCs to provide a preview of the Joe Model output.

Mean System Capacity: 18.1%

Stressor	Mean	SD	Distribution	Low_Limit	Up_Limit	SysCap
Aug_flow	100	20	normal	100	0	100
Barrier_dams	0	0	normal	0	5	100
BKTR	0	0	normal	0	100	100
Feb_flow	99.97	0	normal	100	0	100
Foot_flow	13.77	0	normal	0	100	100
Fragmentation	0.12	0.5	normal	0	0.32	64.4
Habitat_loss	0	0	normal	0	100	Pop.

Adjust Stressor Magnitude

Multiple Units Selected

1702010104, 1702010101, 1702010201, 1702010202, 1702010205, 1702010208, 1702010206, 1702010207, 1702010203, 1702010106, 1702010105

[illegible]ple HUCs selected. Use the table below to set stressor magnitude values for the selected HUCs. Note that entering a value in any of th[illegible]s will update the values for all the selected HUCs. Leave values blank to keep original values for each HUC.

Mean System Capacity: 16.1%

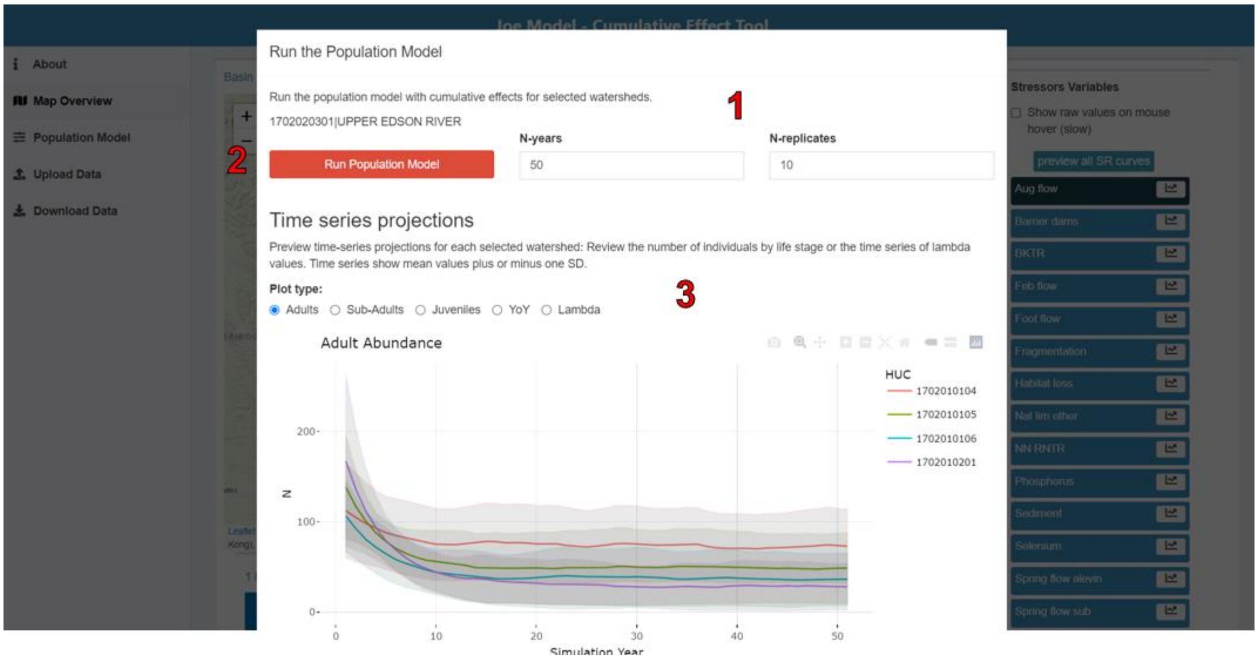
Stressor	Mean	SD	Distribution	Low_Limit	Up_Limit	SysCap
Aug_flow						100
Barrier_dams						100
BKTR						99
Feb_flow						100
Foot_flow						99.2
Fragmentation						68.5
Habitat_loss						Pop.

Figure 19. Example of single or multi-location stressor magnitude adjustment across spatial units.

6.3.6 Life cycle model

To run the life cycle model associated with the CEMPRA, first select (click) the desired spatial polygons for the run. Once the desired spatial polygons are selected, click the “Life cycle model” button on the left side of the lower panel. In the pop-up window that appears, enter the number of years (in the “N-years” box) and the number of replicates or batch runs (in the “N-replicates” box). When ready to run the life cycle model for the selected spatial polygons, click “Run Life cycle model”.

After the run is complete, three graph panels will appear in the pop-up window. The first panel displays time series projections of the number of individuals by life stage and a time series of lambda values for adults. The second panel plots abundance at different life stages in the selected spatial polygon(s) compared to a baseline condition. The third panel plots stressor magnitude against system capacity (0-1) for each of the stressors used in the model. You can view individual data points by hovering over the plot with your cursor. When hovering over the plot, you will also see a menu appear in the top right-hand corner of the plotting window. This menu allows you to pan, zoom around the plot and download a snapshot of the plot onto your computer.



Run the Population Model

Run the population model with cumulative effects for selected watersheds.

1702020301|UPPER EDSON RIVER

N-years: 50

N-replicates: 10

Time series projections

Preview time-series projections for each selected watershed: Review the number of individuals by life stage or the time series of lambda values. Time series show mean values plus or minus one SD.

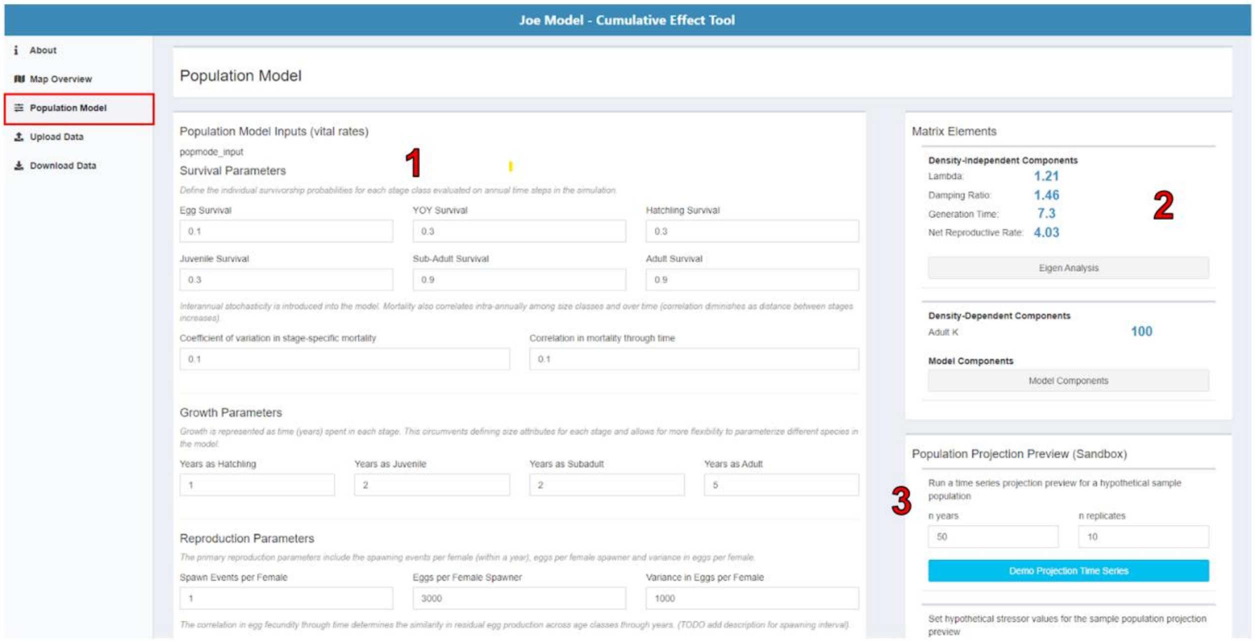
Plot type: ● Adults ○ Sub-Adults ○ Juveniles ○ YoY ○ Lambda

Figure 20: Pop-up window for the life cycle model. 1) Input boxes for the number of years and number of replicates in the simulation. 2) Button to run the life cycle model. 3) Plotting window for time series projection plots.

6.4 Life cycle model

The life cycle model is included as an extension to the original CEMPRA. It leverages a flexible stage-structured life cycle modelling framework that can be modified for different target systems. The underlying life cycle model code was originally developed by van der Lee and Koops (2020) and then further modified by Dr. Kyle Wilson to incorporate environmental stressors that are linked to survivorship, capacity, and growth.

If you intend to use the coupled life cycle model along with the Joe Model, click on the **Life cycle model** page in the menu on the left-hand side of the screen. This page will display all of the parameters and vital rates being used as inputs for the life cycle model. To edit any of these parameters, click anywhere inside the box for a given parameter and either manually enter a new number, or use the up and down arrows that appear on the right side of the box to incrementally adjust the value. Refer to Section 7 for background and additional details on the life cycle modelling component of the CEMPRA tool.



Population Model

Population Model Inputs (vital rates)

popmode_input

Survival Parameters

Define the individual survivorship probabilities for each stage class evaluated on annual time steps in the simulation.

Egg Survival: 0.1
YOY Survival: 0.3
Hatchling Survival: 0.3
Juvenile Survival: 0.3
Sub-Adult Survival: 0.9
Adult Survival: 0.9

Interannual stochasticity is introduced into the model. Mortality also correlates intra-annually among size classes and over time (correlation diminishes as distance between stages increases).

Coefficient of variation in stage-specific mortality: 0.1
Correlation in mortality through time: 0.1

Growth Parameters

Growth is represented as time (years) spent in each stage. This circumvents defining size attributes for each stage and allows for more flexibility to parameterize different species in the model.

Years as Hatchling: 1
Years as Juvenile: 2
Years as Subadult: 2
Years as Adult: 5

Reproduction Parameters

The primary reproduction parameters include the spawning events per female (within a year), eggs per female spawner and variance in eggs per female.

Spawn Events per Female: 1
Eggs per Female Spawner: 3000
Variance in Eggs per Female: 1000

The correlation in egg fecundity through time determines the similarity in residual egg production across age classes through years. (TODO add description for spawning interval).

Matrix Elements

Density-Independent Components

Lambda: 1.21
Damping Ratio: 1.46
Generation Time: 7.3
Net Reproductive Rate: 4.03

Density-Dependent Components

Adult K: 100

Model Components

Population Projection Preview (Sandbox)

Run a time series projection preview for a hypothetical sample population

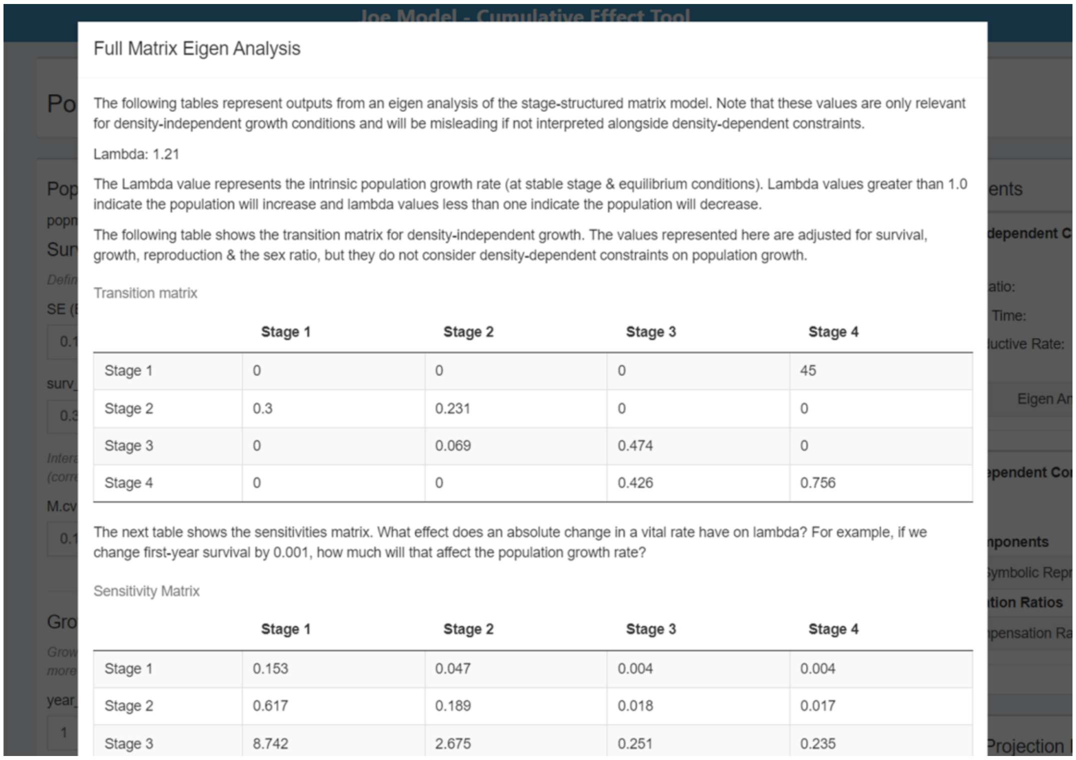
n years: 50
n replicates: 10

Set hypothetical stressor values for the sample population projection preview

Figure 21. Life cycle model page of the R Shiny application. 1) Input section for population vital rates. 2) Parameters for density-independent and density-dependent components of the model. 3) Sandbox for conducting sample runs of the life cycle model

Matrix elements and derived parameters from the life cycle model are shown in the upper right corner of the page. These values will update in real-time as users change the inputs. Lambda values are useful because they show if population parameters will cause the population to grow (lambda greater than 1) or shrink over time (lambda less than 1). Generation time can also be useful to reference against the target population of interest.

Clicking on the “Eigen Analysis” button will bring up a full modal display showing the derived transition matrix, elasticities, sensitives, net reproductive rates and stable stage distributions.



Full Matrix Eigen Analysis

The following tables represent outputs from an eigen analysis of the stage-structured matrix model. Note that these values are only relevant for density-independent growth conditions and will be misleading if not interpreted alongside density-dependent constraints.

Lambda: 1.21

The Lambda value represents the intrinsic population growth rate (at stable stage & equilibrium conditions). Lambda values greater than 1.0 indicate the population will increase and lambda values less than one indicate the population will decrease.

The following table shows the transition matrix for density-independent growth. The values represented here are adjusted for survival, growth, reproduction & the sex ratio, but they do not consider density-dependent constraints on population growth.

Transition matrix

	Stage 1	**Stage 2**	**Stage 3**	**Stage 4**
Stage 1	0	0	0	45
Stage 2	0.3	0.231	0	0
Stage 3	0	0.069	0.474	0
Stage 4	0	0	0.426	0.756

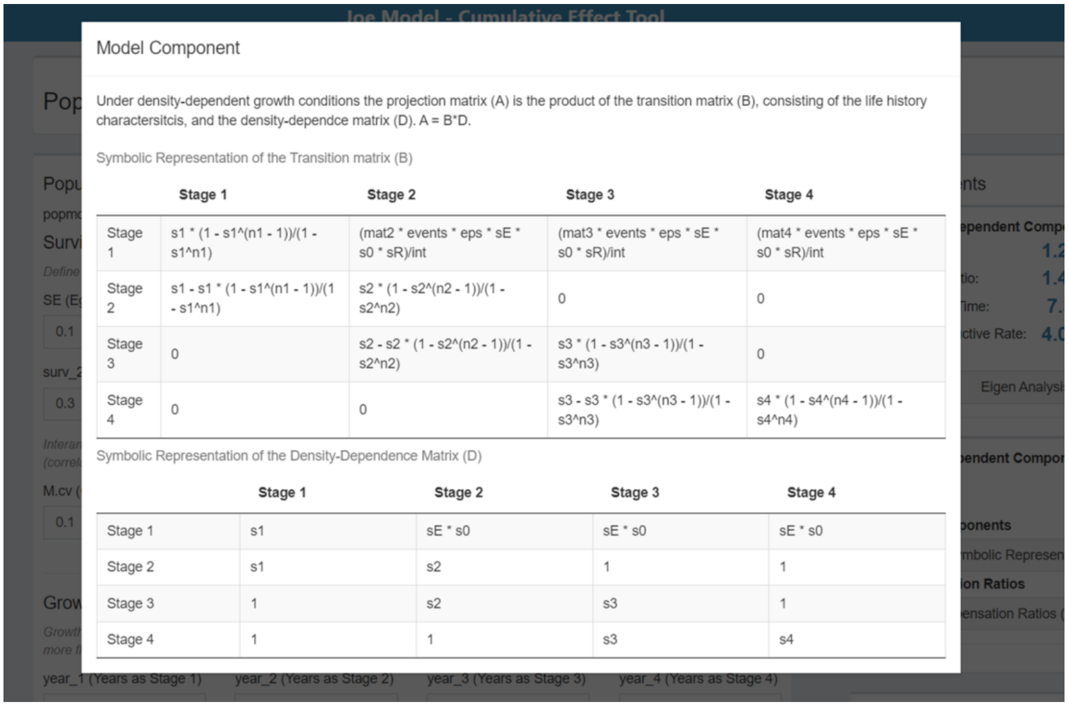
The next table shows the sensitivities matrix. What effect does an absolute change in a vital rate have on lambda? For example, if we change first-year survival by 0.001, how much will that affect the population growth rate?

Sensitivity Matrix

	Stage 1	**Stage 2**	**Stage 3**	**Stage 4**
Stage 1	0.153	0.047	0.004	0.004
Stage 2	0.617	0.189	0.018	0.017
Stage 3	8.742	2.675	0.251	0.235

Figure 22. Full eigen analysis showing the transition matrix, sensitivity matrix, elasticity matrix, stable stage distribution estimate of the intrinsic population growth rate (lambda), generation time and damping ratio.

The symbolic representation of the transition matrix can also be reviewed by clicking on the “Symbolic Representations” button. Additional explainer popups are also available throughout the page.



Model Component

Under density-dependent growth conditions the projection matrix (A) is the product of the transition matrix (B), consisting of the life history charactersitcis, and the density-dependce matrix (D). A = B*D.

Symbolic Representation of the Transition matrix (B)

	Stage 1	Stage 2	Stage 3	Stage 4
Stage 1	s1 * (1 - s1^(n1 - 1))/(1 - s1^n1)	(mat2 * events * eps * sE * s0 * sR)/int	(mat3 * events * eps * sE * s0 * sR)/int	(mat4 * events * eps * sE * s0 * sR)/int
Stage 2	s1 - s1 * (1 - s1^(n1 - 1))/(1 - s1^n1)	s2 * (1 - s2^(n2 - 1))/(1 - s2^n2)	0	0
Stage 3	0	s2 - s2 * (1 - s2^(n2 - 1))/(1 - s2^n2)	s3 * (1 - s3^(n3 - 1))/(1 - s3^n3)	0
Stage 4	0	0	s3 - s3 * (1 - s3^(n3 - 1))/(1 - s3^n3)	s4 * (1 - s4^(n4 - 1))/(1 - s4^n4)

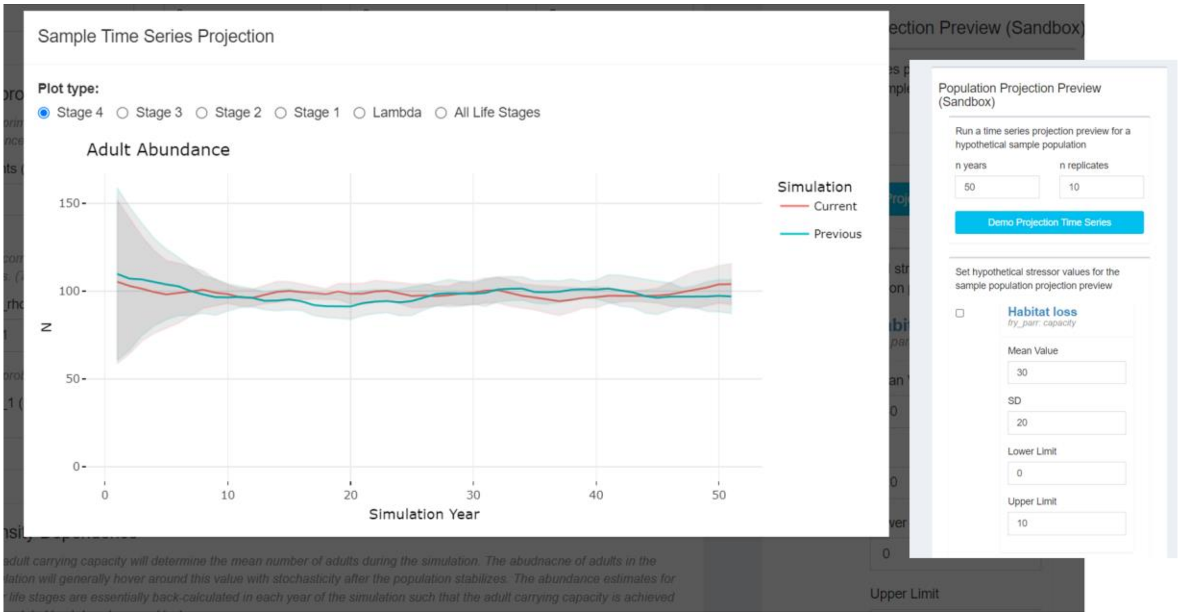
Symbolic Representation of the Density-Dependence Matrix (D)

	Stage 1	Stage 2	Stage 3	Stage 4
Stage 1	s1	sE * s0	sE * s0	sE * s0
Stage 2	s1	s2	1	1
Stage 3	1	s2	s3	1
Stage 4	1	1	s3	s4

Figure 23. Symbolic (equation) representation of the transition matrix and density dependence matrix.

In the bottom left-hand corner of the page is a population projection “sandbox” module. This section allows the user to run sample projections of the population forward through time over a set number of years and batch replicates. Adjusting any of the life cycle parameters and re-running the sandbox projection will allow the user to see how their changes influence the growth trajectory and variability of the population. Running simulations is valuable since it allows users to capture density-dependent and density-independent components of the life cycle model. Separate time series are drawn for the current (red) and previous (teal) projections. The mean value (across batch replicates) is shown as a line, and the grey shading shows the variability across each batch replicate (Figure 24).

The sandbox module can also be used to run sample projections for hypothetical combinations of stressor values. Stressor variables linked to the life cycle model and uploaded by the user will appear below the “Demo Projection Time Series” button. Users can populate these values to demonstrate how an individual stressor (or group of stressors) will influence the population at a set of hypothetical values. Be sure to expand the panel for each stressor and values for the mean, SD, lower limit, and upper limit. Check or uncheck the checkbox to include or exclude the stressor in the sandbox simulations.


Sample Time Series Projection

Plot type:
● Stage 4 ○ Stage 3 ○ Stage 2 ○ Stage 1 ○ Lambda ○ All Life Stages

Population Projection Preview (Sandbox)

Run a time series projection preview for a hypothetical sample population

n years: 50
n replicates: 10

Set hypothetical stressor values for the sample population projection preview

☐ **Habitat loss**
fry_parr: capacity

Mean Value: 30
SD: 20
Lower Limit: 0
Upper Limit: 10

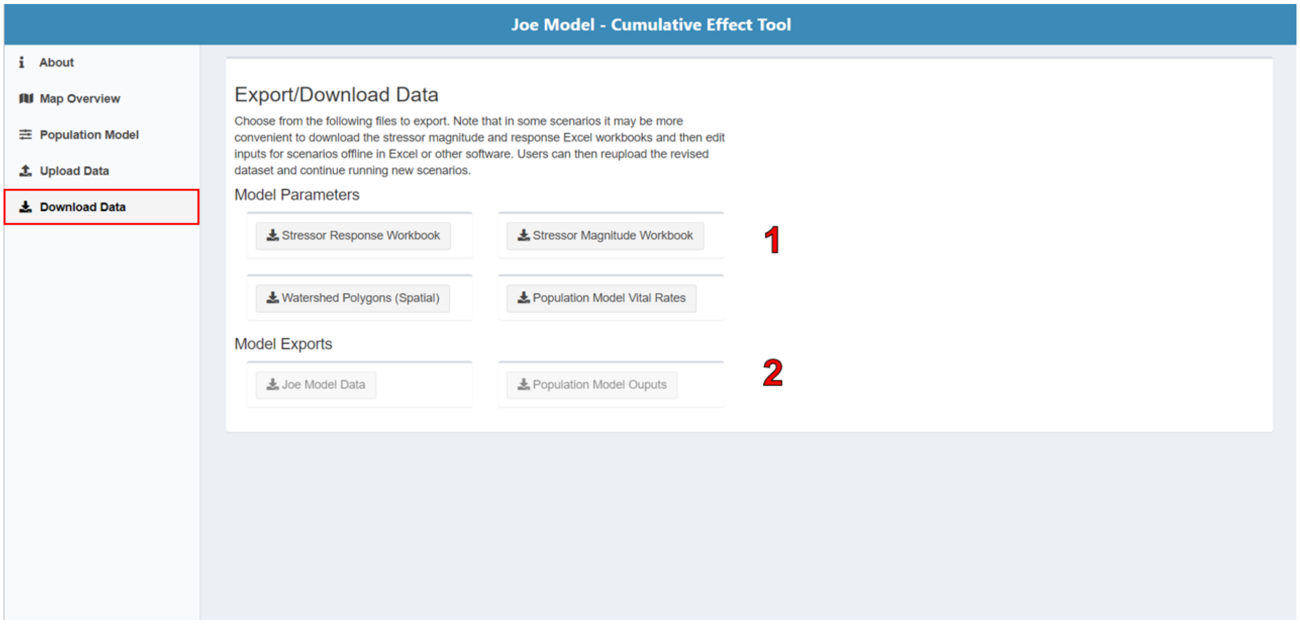
Figure 24. Population projection “sandbox” showing stage-specific abundance values for current and previous model runs. Users may change values such as fecundity or a stressor to visualize how the population will respond to hypothetical changes to parameters (outside of formalized scenario runs).

6.5 Download Data

On the **Download Data** page, you can either download the model input parameter files (i.e., the stressor-response workbook, stressor magnitude workbook, spatial polygons file, and life cycle profile) for offline revisions, or you can directly download the CEMPRA, or Life cycle model outputs for reporting purposes.

Model Parameters: In some circumstances, it may be more convenient to download the stressor magnitude and response Excel workbooks, edit inputs for scenarios offline in Excel or other software and reupload.

Model Exports: The model results can only be exported after the model is run. Export buttons will be disabled until the Joe Model or life cycle model (population model) has been run.



Choose from the following files to export. Note that in some scenarios it may be more convenient to download the stressor magnitude and response Excel workbooks and then edit inputs for scenarios offline in Excel or other software. Users can then reupload the revised dataset and continue running new scenarios.

Figure 25: Download the Data page of the R Shiny Application. 1) Model parameter files for download. 2) Model output files for download (CEMPRA and life cycle model).

6.6 Interactive Formula Builder

**** (March 22, 2023): This component is still in development and not yet available on the GitHub repository application ****

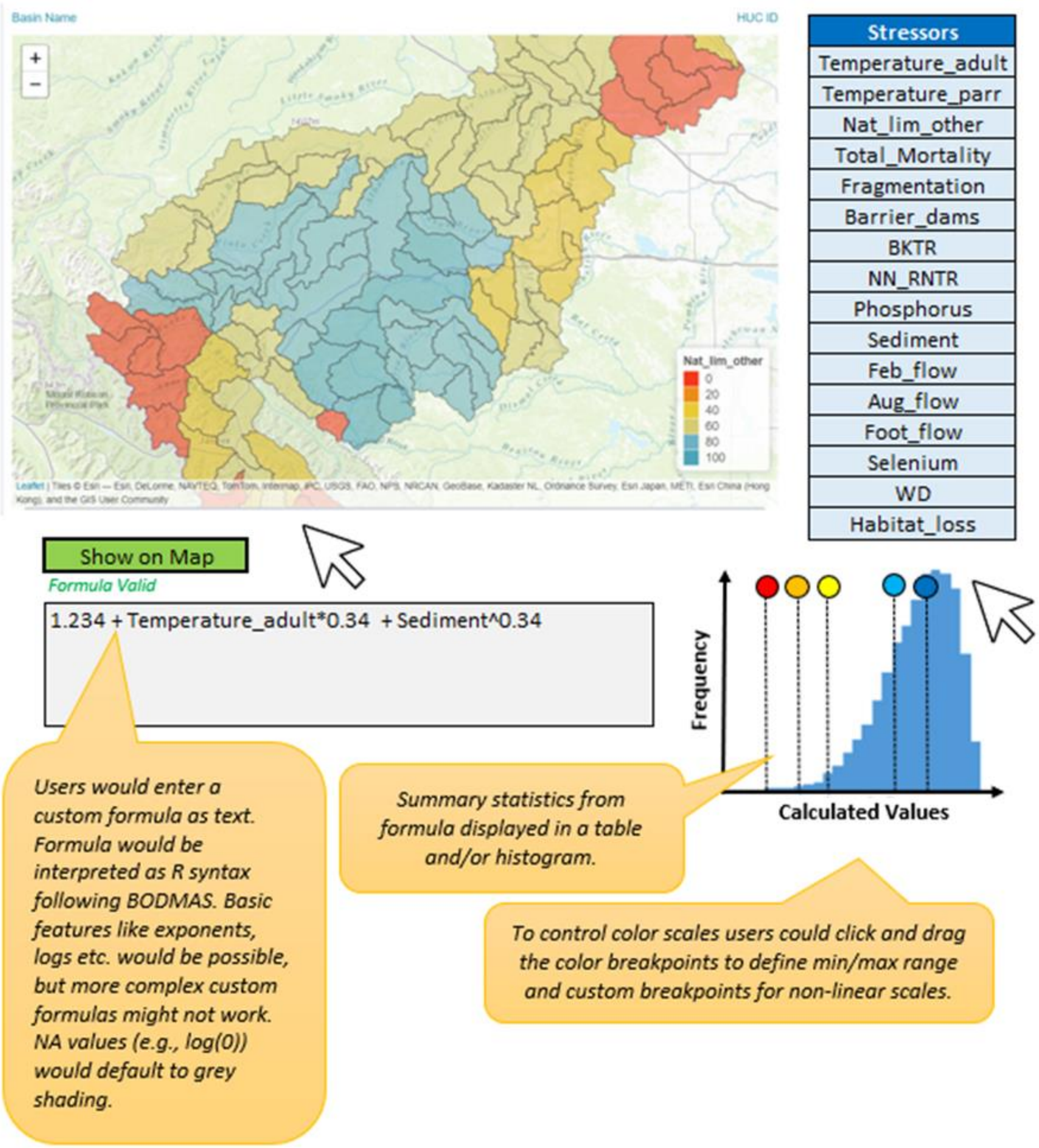
The interactive formula builder is used to generate custom plots of stressor variables with customized functions and potential interactions between variables. The interactive formula builder provides users with the ability to display outputs from customized regression equations.

A dialogue box is available, allowing users to drag and drop multiple stressors into a formula builder to display relationships and interactions (e.g., 0.45 + stressor_A^0.3 + stressor_B*stressor_C). Various mathematical operators can be included in the formula builder:

- Addition: +; Stressor_A + Stressor_B
- Subtraction: -; Stressor_A - Stressor_B
- Constants: 0.783 + Stressor_A
- Multiplication: *; Stressor_A * Stressor_B
- Division: /; Stressor_A / Stressor_B
- Brackets: (); (0.783 + Stressor_A) / Stressor_B
- Exponents: ^; Stressor_A^(0.34 * Stressor_B)
- Logarithms:
    - Natural: log(Stressor_A)
    - Base-10: log10(Stressor_A)

The interactive formula builder can be useful to display outputs from various habitat suitability models, predictive regression equations, alternative cumulative effects roll-up functions or building novel stressors.

Building Novel Stressors: The interactive formula builder can also be used to build novel stressor variables that consist of combinations of individual stressors. Users enter values into the formula builder to create a synthetic stressor variable. Once satisfied, users can download the novel stressor in the standardized stressor magnitude and stressor-response workbook format. The novel data can then be reimported into the tool to appear on the main map panel for use in the Joe Model and Life Cycle Model.



1.234 + Temperature_adult*0.34 + Sediment^0.34

Users would enter a custom formula as text. Formula would be interpreted as R syntax following BODMAS. Basic features like exponents, logs etc. would be possible, but more complex custom formulas might not work. NA values (e.g., log(0)) would default to grey shading.

Summary statistics from formula displayed in a table and/or histogram.

To control color scales users could click and drag the color breakpoints to define min/max range and custom breakpoints for non-linear scales.

Figure 26: Interactive formula builder module

7 Life Cycle Modeling Framework

7.1 Overview

The integrated life cycle model is a core component of the CEMPRA tool. The life cycle modelling component is a valuable endpoint to evaluate and understand cumulative effects through the lens of demographic rates and population ecology. Some user groups may be satisfied with the simplified Joe Modelling (stressor roll-up) component of the CEMPRA tool and, therefore, not wish to interact with the life cycle model. However, other user groups may benefit substantially from working with the life cycle modelling component. Framing cumulative effects through an integrated life cycle model allows us to understand critical bottlenecks to the productivity and capacity of a target study system. In the CEMPRA life cycle modelling component, stressor-response relationships are linked to vital rates such as life-stage-specific survivorship, fecundity, and carrying capacity. Therefore, the life cycle modelling component can be used to make relative comparisons between locations (spatial units), scenarios, and stressors to understand limiting factors and design recovery action strategies.

At its core, the life cycle modelling component of the CEMPRA tool is a stage-structured matrix model (see Caswell 2000). A simplified life cycle profile csv data input file (described below) is populated by the user and then imported to parameterize and construct components of the matrix model (e.g., number of stages, stage-specific survivorship, years in each stage, etc.).

When the life cycle model is run, a hypothetical population is projected forward in time though simulations. The stage-structured matrix model governs the behaviour of the simulated population. Density-dependent growth constraints are implemented using either compensation ratios (if the location and stage-specific capacities are unknown) or location and stage-specific Beverton-Holt functions (discussed further below). Location-specific stressor values will interact with the simulated population to curtail or enhance stage-specific survivorship, fecundity, or habitat capacities. Population projections are then compared across scenarios and/or locations to evaluate the relative change in equilibrium abundance estimates for a target life stage (i.e., carrying capacities) and/or the intrinsic productivity (i.e., growth rates) possible at low densities.

The life cycle modelling component of the CEMPRA tool performs a large number of calculations behind the scenes. While convenient, the embedded complexity can create misleading results if input values and assumptions are not carefully considered. It is assumed that users of the life cycle model have an understanding of basic concepts in population ecology (e.g., population growth rates, carrying capacities etc.) and a familiarity with matrix life cycle models. The following resources provide useful refreshers for interested individuals:

- Basic refresher on matrix life cycle models: https://compadre-db.org/Education/article/what-is-a-matrix-model
- An in-depth overview of stage-structured matrix models: https://blog.uvm.edu/tdonovan-vtcfwru/files/2020/06/12-Donov-pages-322-CB.pdf
- Density-dependent and density-independent constraints on population growth: https://www.nature.com/scitable/knowledge/library/population-limiting-factors-17059572/
- Density-dependent growth functions (review section on the Beverton-Holt function): http://courses.ecology.uga.edu/ecol4000-fall2018/wp-content/uploads/sites/22/2018/08/Chapter-3-complex-dynamics.pdf

The model code for the life cycle modelling component of the CEMPRA tool follow a similar structure to the codebase used by Van der Lee and Koops (2020). The underlying code and assessment framework was modified substantially by Dr. Kyle Wilson and Matthew Bayly (M.J. Bayly Analytics Ltd.) throughout 2022 and 2023. Code snippets, functional forms, and rationale largely follow conventional workflows demographic modelling outlined in Schaube and Kéry 2021. Users are encouraged to review these resources for additional background and rationale.

7.2 Data Input: Life Cycle Profiles

7.2.1 Purpose

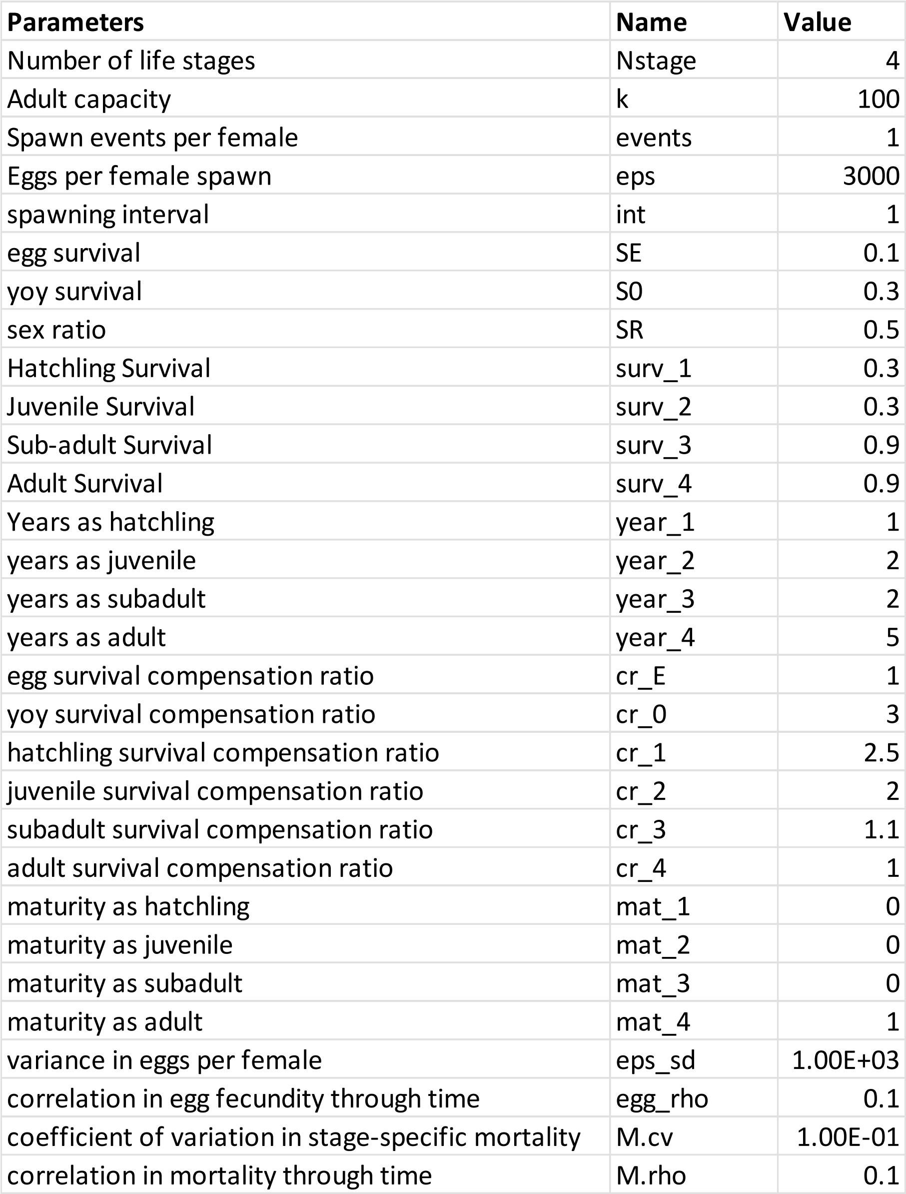
The life cycle profile is the main input file for the life cycle model. It provides the names and values of key life cycle parameters and vital rates, including parameters for survival, growth, reproduction, and density-dependent effects. This file makes it easy for users to store and edit life cycle parameter values outside of the R Shiny web application. The following sections break down components of the life cycle profile file with illustrative examples. The intent of the following sections is to provide a detailed explanation of how the life cycle model works with a description of each component of the input file so that users may create their own life cycle profile for a target species of interest.

7.2.2 Layout

The life cycle profile is a comma-separated values (CSV) file that contains the names and values of each of the parameters within the life cycle model. Life cycle profiles will be unique to each species or life history variant. The life cycle profile file contains three columns:

- **Parameters:** The full name/description of the parameter. This column can be adjusted by the user to provide more relevant nicknames for each stage (e.g., fry survival, smolt survival etc.)

- **Name:** The short form name of the parameter used in the model. The names of these parameters are referenced by the model code and should not be modified (apart from adding or removing stage classes).
- **Value:** The numeric value of the parameter used in the model. The values are adjusted for each species profile.



Parameters	Name	Value
Number of life stages	Nstage	4
Adult capacity	k	100
Spawn events per female	events	1
Eggs per female spawn	eps	3000
spawning interval	int	1
egg survival	SE	0.1
yoy survival	S0	0.3
sex ratio	SR	0.5
Hatchling Survival	surv_1	0.3
Juvenile Survival	surv_2	0.3
Sub-adult Survival	surv_3	0.9
Adult Survival	surv_4	0.9
Years as hatchling	year_1	1
years as juvenile	year_2	2
years as subadult	year_3	2
years as adult	year_4	5
egg survival compensation ratio	cr_E	1
yoy survival compensation ratio	cr_0	3
hatchling survival compensation ratio	cr_1	2.5
juvenile survival compensation ratio	cr_2	2
subadult survival compensation ratio	cr_3	1.1
adult survival compensation ratio	cr_4	1
maturity as hatchling	mat_1	0
maturity as juvenile	mat_2	0
maturity as subadult	mat_3	0
maturity as adult	mat_4	1
variance in eggs per female	eps_sd	1.00E+03
correlation in egg fecundity through time	egg_rho	0.1
coefficient of variation in stage-specific mortality	M.cv	1.00E-01
correlation in mortality through time	M.rho	0.1

Figure 27: Example of a life cycle profile CSV input file with four stage classes.

7.3 Matrix Life cycle model

The stage-structured matrix modelling framework, embedded within the CEMPRA tool, can be represented graphically by a life cycle diagram (Figure 28) or a transition matrix. The transition matrix can be represented symbolically with either equations (Table 1) or absolute values (Table 2).

Figure 28 shows the sample life cycle profile. Figure 27 is represented as a life cycle diagram. In the diagram and input file, we see that there are four main stages (stage_1 to stage_4). stage_1 individuals can become stage_4 individuals after three years in the simulation, but it is also possible for some individuals to spend more than one year in stages 2, 3 and 4 (denoted by the circular loop). We also see that stage_4 individuals are sexually mature and have the capacity to generate new stage_1 individuals. There are also special year 0 (Age-0) events that occur before new stage_1 (Age-1) individuals are secured in the simulation. These events include egg survival (SE) and Age-0 fry survival (S0).

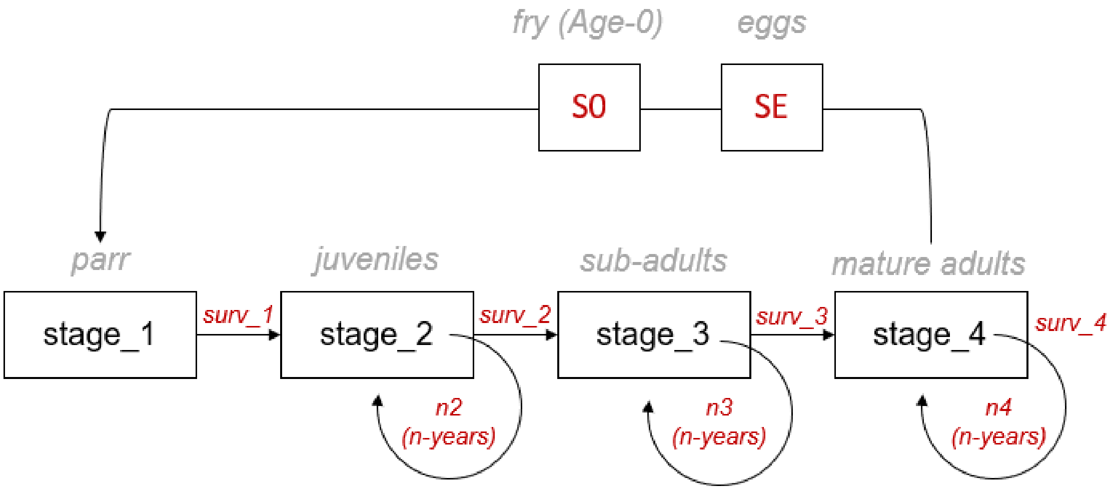


Figure 28. The life cycle diagram for the demo life cycle profile is shown in Figure 27.

The stage-structured matrix life cycle model in the CEMPRA tool is set up as a pre-birth pulse census (see Caswell 2000). Since the design of stage-structured matrix models does not easily allow for the initial number of eggs and fry to be represented as independent matrix elements (cells), their transitions are included within the fecundity term. In a pre-birth pulse census, we assume that the demographic census takes place immediately before spawning (fecundity), meaning that yearlings of the previous spawning year have survived a full time-step (Age-0/stage-0 to Age-1/stage-1). Yearlings (Age-0: egg & fry) must survive the entire census period to the start of the next census. Therefore, the Age-0 transitions (egg-to-fry survivorship: SE and fry-to-parr survivorship: S0) are accounted for within the fecundity element (cells) of the transition matrix (Table 1).

Vital Rates for Survivorship and Growth

Parameter	Description
Nstage	The number of stages in the transition matrix (excluding Stage-0/Age-0). Each stage must span one or more years in the life cycle. In the reference example, there are four stages: stage_1, stage_2, stage_3 and stage_4.
surv_1 **surv_2** **surv_3** **surv_4**	Mean annual survivorship of individuals in each stage. If Nstage value was higher than four, additional rows would be added to the life cycle profile csv file for surv_5, surv_6 etc. If the Nstage value is less than four, then the addition surv_X rows should be deleted from the life cycle profile. *These survivorship estimates should be estimates of intrinsic density-independent survival (in the absence of density-dependent constraints).*
year_1 **year_2** **year_3** **year_4**	The number of years spent in each stage. Individuals in the simulation can spend more than one year in each stage. In the example input file, mature adults can spend up to five years in stage_4. If we were to set the year values to 1 for all stages, then we would have an age-based Leslie matrix model. If additional stages are added (or removed), be sure to modify rows in the life cycle profile csv file accordingly (i.e., delete or add new rows).
SE	Egg survivorship (density-independent).
S0	Age-0 fry survivorship (density-independent).

Ensure that all survivorship estimates represent hypothetical density-independent survivorship in the absence of density-dependent constraints. Density-dependent survivorship is accounted for in the next section. If density-independent survivorship is unknown, but strong, density-dependent constraints are to be included in the species profile, then it might be possible to simply set the density-independent survivorship estimate to a value close to 1.0 (e.g., S0: 0.999).

For fecundity, we have to consider the proportion of each age class that is sexually mature (mat), the proportion of the population that is female (SR: 0.5), the fecundity (eps: eggs per spawner) per spawning event, the spawning events per year (events) and the spawning interval (int). The calculation of individuals in stage class 1 (stage_1) also must account for the Age-0 survivorship of eggs and fry.

Sample fecundity function for stage class 4:

$$\frac{(mat4\ \times events\ \times eps\ \times SE \times S0 \times SR)}{int}$$

Vital Rates for Fecundity

Parameter	Description
mat_1 **mat_2** **mat_3** **mat_4**	The proportion of each stage class that is sexually mature (0 – 1). For example, in the demo species profile, 100% of the individuals become sexually mature at stage class 4, and the sexual maturity is 0% for all other stage classes. It is also possible for a stage class to have partial maturity (e.g., 0.85). If the Nstage value is different than four, then add or remove rows in the species profile so that the number of mat values matches the number of stage classes (Nstage value).
events	Spawning events per female per year. This parameter will almost always be set to 1 to indicate one spawning event per year per mature female. For populations with complex life history variants (e.g., systems with Spring Chinook & Fall Chinook), we recommend keeping this value at one and using two different species profiles to represent each life history variant.
eps	Eggs per spawning female. The mean fecundity per female per spawning event.
SR	The sex ratio is represented as the proportion of the population that is female. This value will almost always be set to 0.5 to indicate an equal proportion of males and females in the population.
int	Spawning interval (in years). This value will also be set to 1 for most species indicating that mature individuals spawn each year.

We can combine all parameters discussed in this section along with the example species profile (Figure 27) to construct a symbolic (mathematical) representation of the transition matrix (Table 1). The stage-to-stage transitions account for the probability of staying within each stage or advancing to the next stage based on the surv_X and n-year spent within a stage year_X. The fecundity element of the matrix (top row) includes elements for fecundity and Age-0 survivorship.

Table 1. Symbolic Representation of the Transition Matrix

	stage_1	stage_2	stage_3	stage_4
stage_1	surv_1 * (1 - surv_1^(year_1 - 1))/(1 - surv_1^ year_1)	(mat2 * events * eps * sE * s0 * sR)/int	(mat3 * events * eps * sE * s0 * sR)/int	(mat4 * events * eps * sE * s0 * sR)/int
stage_2	surv_1 - surv_1 * (1 - surv_1^(year_1 - 1))/(1 - surv_1^ year_1)	surv_2 * (1 - surv_2^(year_2 - 1))/(1 - surv_2^ year_2)	0	0
stage_3	0	surv_2 - surv_2 * (1 - surv_2^(year_2 - 1))/(1 - surv_2^ year_2)	surv_3 * (1 - surv_3^(year_3 - 1))/(1 - surv_3^ year_3)	0
stage_4	0	0	surv_3 - s3 * (1 - surv_3^(year_3 - 1))/(1 - surv_3^ year_3)	surv_4 * (1 - surv_4^(year_4 - 1))/(1 - surv_4^ year_4)

The stage-to-stage transition probabilities are also expressed as functions from surv_X (annual survivorship with stage X) and year_X (number of simulation years within stage X). surv_X is the total annual probability of survival (i.e., regardless of staying within the current stage OR advancing to the next subsequent stage). The example below illustrates how the combined probability of staying within a stage or advancing to the next stage always equals surv_X regardless of n-years in each stage. *Note that the sum of the yellow cells equals 0.6 (for both fates of staying within stage or advancing to next stage).*

Probability of staying in stage X as stage X

		Years in Stage X			
		1	**2**	**3**	**4**
Survival (surv_X)	**0**	-	0.000	0.000	0.000
	0.2	0.000	0.167	0.194	0.199
	0.6	0.000	0.375	0.490	0.540
	0.8	0.000	0.444	0.590	0.661
	1	-	-	-	-

Probability of advancing from stage X to stage Y

		Years in Stage X			
		1	**2**	**3**	**4**
Survival (surv_X)	**0**	-	0.000	0.000	0.000
	0.2	0.200	0.033	0.006	0.001
	0.6	0.600	0.225	0.110	0.060
	0.8	0.800	0.356	0.210	0.139
	1	-	-	-	-

We can continue with the working example to represent the transition matrix numerically (Table 2). The fecundity element for stage_4 is set at 45 since it accounts for the vital rates relating to maturity and Age-0 survivorship.

Net Fecundity (stage-4) = (mat4 * events * eps * sE * s0 * sR)/int
45 = (1 * 1 * 3,000 * 0.1 * 0.3 * 0.5)/1

Table 2. Numerical representation of the transition matrix

	stage_1	**stage_2**	**stage_3**	**stage_4**
stage_1	0	0	0	45
stage_2	0.3	0.231	0	0
stage_3	0	0.069	0.474	0
stage_4	0	0	0.426	0.756

The derived lambda value of the projection matrix (intrinsic rate of growth) in this example species profile is 1.21 (above 1.0), meaning that the population will continue to grow exponentially in the absence of density-dependent constraints.

7.4 Density-Dependent Constraints on Growth

It is rare for natural populations to grow in perpetuity without any constraints on growth, survival, and reproduction. Therefore, life cycle models will typically include a mechanism (or multiple mechanisms) to constrain population growth or limit high densities.

Stressor-response relationships can be incorporated into the life cycle model without accounting for density-dependent constraints, but the interpretation of the results will be limited to eigen analyses of intrinsic growth rates and sensitivities/elasticities assessments. Not only are these outputs difficult to communicate to diverse working groups, but they may also be misleading. If there are key demographic bottlenecks in the life cycle, then a density-independent model may inappropriately lead users to focus on stressors linked to fecundity or early life-stage survivorship, before a key bottleneck (e.g., egg-to-fry survivorship) is experience. However, if (in reality) a hypothetical population experiences strong density-dependant constraints on growth, then factors limiting habitat availability or productivity of a key life stage will become more influential. A common example of density-independent and density-dependent constraints on growth can be found in the transition between early life stages of Steelhead (*Oncorhynchus mykiss*) as individuals transition from egg-to-fry (a density independent life stage) and then from fry-to-smolts/parr (a density-dependent life stage) (Ward and Slaney, 1993).

The life cycle modelling component of the CEMPRA tool has two different mechanisms to incorporate density-dependent growth constraints. The first mechanism makes use of adult carrying capacity and compensation ratios (discussed below), while the second mechanism considers stage and location-specific carrying capacities. Incorporating density-dependent constraints with compensation ratios is convenient because users are only required to input a single value (an estimate of the adult carrying capacity); however, incorporating density-dependent constraints with stage and location-specific carrying capacities can provide more accurate results in systems with ample habitat data and knowledge of life stage-specific density constraints across habitat types. Both mechanisms of density-dependent constraints utilize the Beverton-Holt function to constrain the transition probability of key demographic bottlenecks.

The Beverton-Holt function calculates the expected number of individuals in the next time step (N at time + 1, or density) as a function of the number of individuals in the current time step (N at time). In the case of stage-structured matrix models, this relationship is expressed as the number of individuals transitioning between two stages (e.g., from stage 2 to stage 3). In the Beverton-Holt function, input parameters governing this relationship include an estimate of stage-specific carrying capacity (K), a baseline estimate of density-independent survival (S) (surv_X from the species profile) for the transition probability and the number of individuals in the current stage class (Nt) for the simulation year.

Beverton-Holt function for density-dependent growth:

$$N_{t+1} = \frac{S \cdot N_t}{1 + (\frac{S}{K}) \cdot N_t}$$

Figure 29 provides an overview of the Beverton-Holt function showing the number of individuals at time (t) on the x-axis and the number of individuals at time + 1 on the y-axis. For example, this could be the number of Age-0 fry on the x-axis and the number of Age-1+ parr recruits on the y-axis. The curved black line shows the effects of density-dependent growth (limited recruitment as the number of individuals in the first stage is increased). The steep red line is the intrinsic productivity (survivorship, surv_X), which is 0.8 under hypothetical density-independent growth conditions. The blue line is the hypothetical carrying capacity of 100.

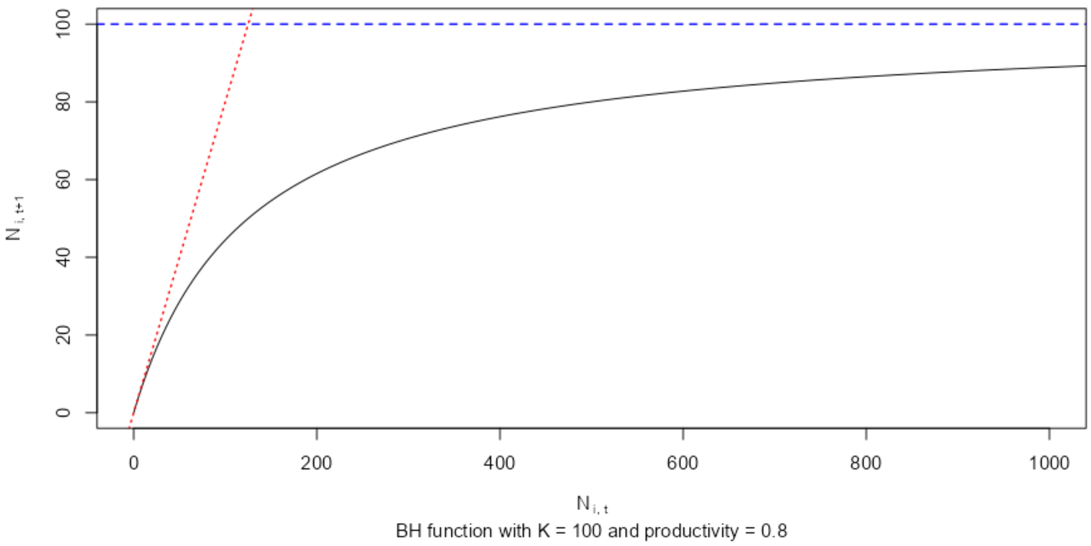
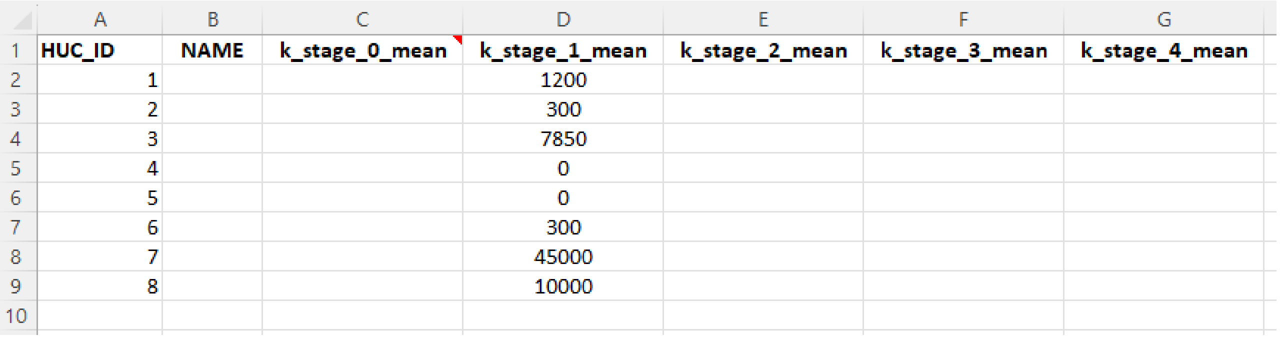


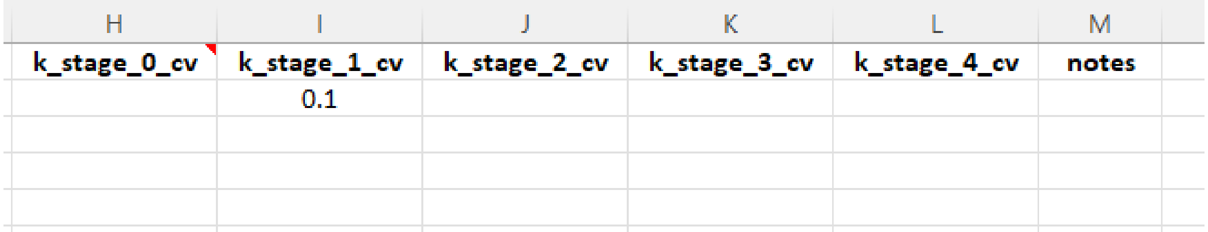
Figure 29. Beverton-Holt function for density-dependent growth showing the survivorship under density-independent growth (red line), the maximum carrying capacity (blue line) and the resulting relationship from the BH curve (black line).

7.4.1 Location and stage-specific carrying capacities:

For more systems with ample habitat data and known relationships between habitat availability and maximum densities, it is possible to use location-specific carrying capacities for one or more rate-limiting life stages (e.g., *Location X can produce up to 1,200 parr*). If this is the case, a special table can be included that specifies the maximum number of individuals per stage class per life stage per location:



	A	B	C	D	E	F	G
1	**HUC_ID**	**NAME**	**k_stage_0_mean**	**k_stage_1_mean**	**k_stage_2_mean**	**k_stage_3_mean**	**k_stage_4_mean**
2	1			1200			
3	2			300			
4	3			7850			
5	4			0			
6	5			0			
7	6			300			
8	7			45000			
9	8			10000			
10							



k_stage_0_cv	k_stage_1_cv	k_stage_2_cv	k_stage_3_cv	k_stage_4_cv	notes
	0.1				

Figure 30. Location and stage-specific carrying capacity input data table.

This table (Figure 30) is referred to as the **locations carrying capacity table** and exists as a special input file that can be used to control density-dependent growth in the life cycle model. Users must estimate the average carrying capacity for a given life stage at each location (e.g., k_stage_1_mean: 1,200) and the interannual variability (CV: coefficient of variation) in the carrying capacity (e.g., k_stage_1_cv: 0.1). If the location carrying capacity table is provide, any cells that are populated with values are assumed to have density-dependent constraints. Any cells that are left blank are assumed to be governed only by density-independent factors and do not have any density-dependent constraints.

In the previous example (Figure 30) the population model will run with constraints on stage_1 meaning that the survivorship from stage_0 to stage_1 will be governed by a Beverton-Holt relationship and the abundance (or density) of stage_1 individuals will be constrained for each location according the values provided in the locations carrying capacity table.

The CEMPRA tool does not support the development of these location and stage-specific carrying capacity estimates, but it's assumed that relevant reference literature will be used to develop appropriate input values (*e.g., if stage 1 is Steelhead parr; regional densities for Steelhead parr are roughly 1,500 parr/km of stream; and fish-accessible reaches within Rock Creek sum up to roughly 800m in length; then k_stage_1_mean should equal roughly 1,200 parr*). Developing these estimates alongside a species profile can be laborious, but the advantage is that projection from the CEMPRA tool will ultimately account for habitat quality and habitat availability.

If the locations carrying capacity table is provided, the population projections will implement density-dependent growth constraints for species-specific life stages according to the stage-specific carrying capacities (with Beverton-Holt constraint on growth), and intrinsic density-independent survivorship estimates (Figure 29). Figure 31 shows an example simulation for Steelhead from the CEMPRA tool with adult spawning on the y-axis and a stage-1 (parr) carrying capacity constraint set to 160,000 individuals. In the Steelhead example, the only density-dependent constraint is the parr carrying capacity of the system (set to 160,000). The number of adult spawners is, therefore, a derived metric from the life cycle model. Implementing density-dependent constraints with *location and stage-specific carrying capacities* is different from the approach with *compensation ratios,* where users are required to first input an estimate of the adult carrying capacity and work backwards from there.

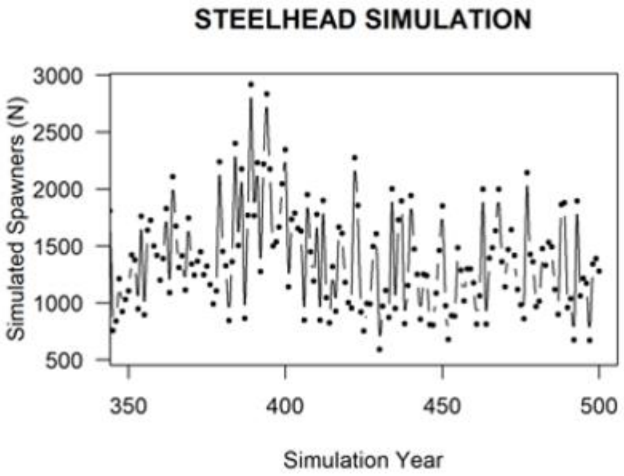


Figure 31. Simulation of Steelhead Spawners (y-axis) with stage_1 (parr) capacity set to 160,000.

Location and stage-specific carrying capacities constraints in the species profile

Parameter	Description
bh_stage_1 **bh_stage_2** **bh_stage_3** **bh_stage_4**	The location and stage-specific carrying capacities constraints can be toggled on and off for each life stage with bh_stage_X input in the species profile. Setting the Value to 0 (or omitting the row) means that there will be no density-dependent constraint on a given life stage. Setting the Value to 1 will implement a Beverton-Holt style density-dependent constraint for the given life stage. If implemented, the *location and stage-specific carrying capacity* input data table must be supplied, and values must be populated for each location.

7.4.2 Compensation Ratios for Density-Dependent Growth

It is possible to implement density-dependent constraints with both the compensation ratios and the location and stage-specific carrying capacities, but we recommend using only one mechanism for density-dependent growth to avoid confusion.

Compensation ratios (CR values) can be used in the CEMPRA life cycle model to parameterize and control density-dependent growth. Compensation ratios (described below) are a reparameterization of the classical Beverton-Holt function for density-dependent growth.

Compensation Ratios (CR) adjust the survivorship of each life stage based on the observed densities (abundance, N i,t) and stage-specific carrying capacities (Ki):

Compensation Ratio CR for life stage i:

$$S_{i,t} = \frac{S_{i,0} \cdot w_i}{1 + \frac{w_i - 1 \cdot N_{i,t}}{K_i}}$$

In the CR equation above, Si,0 is the baseline survivorship (surv_X) under density-independent growth conditions; wi is the compensation ratio (CR value) of life stage i; Ni,t is the current number of individuals in life stage i in a given time step (t); and Ki is the carrying capacity of life stage i. The compensation ratios, in essence, modify the survivorship of each life stage based on how far the stage-specific abundance (Ni,t) has departed from its assumed carrying capacity (Ki).

A plot of compensation ratios is provided below to illustrate their effects on stage-specific survivorship transitions. In this example, abundance values of a hypothetical stage class (i) are plotted along the x-axis with a carrying capacity (Ki) set to 100 individuals (blue vertical line). The hypothetical stage class (i) has a baseline survivorship (productivity) value of 0.8 in the absence of density-dependent growth conditions (horizontal red line). The y-axis on the plot shows how the default survivorship value of 0.8 is modified based on the stage-specific compensation ratio for stage class (CR i). The survivorship value for the stage class is suppressed as the abundance values exceed the carrying capacity K. The effects are amplified as compensation ratios are increased. Compensation ratios of 1.0 leave the vital rate unmodified. Compensation ratios less than 1.0 increase survivorship values (allowing for a potential positive effect of density). When the abundance of the age class is less than the carrying capacity, baseline survivorship values can actually increase. However, within the model code, adjusted survivorship values are fixed so that they never exceed 1.0 for any stage transition.

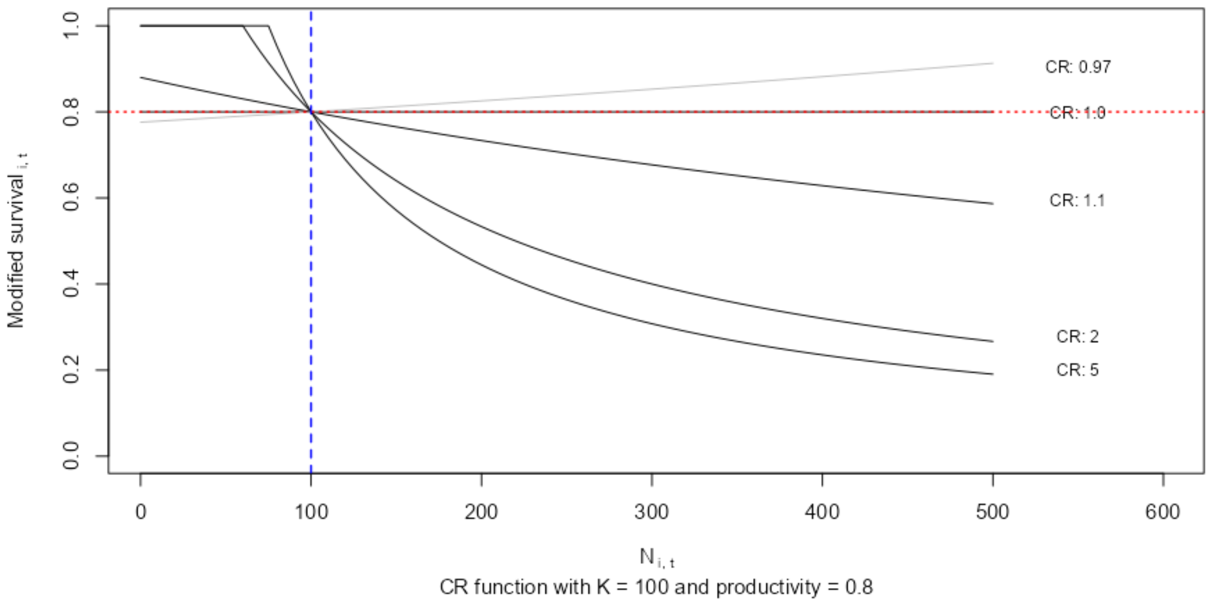


Figure 32. Influence of compensation ratios on stage-specific survivorship.

Stage-Specific Carrying Capacity Estimates:

When using compensation ratios, it is only possible to modify the carrying capacity for the adult age class via the K parameter in the life cycle profile. The K values for other age classes are derived from the stable-stage distribution of the underlying transition matrix. This means that users who implement density-dependent growth via compensation ratios can only modify stage-specific abundance relative to the stable stage distribution. If you know (or can estimate) critical location-specific and stage-specific capacities (e.g., maximum stage_1 parr at location X), then refer to Section **Error! Reference source not found.** to implement density-dependent growth constraints with location and stage-specific carrying capacities.

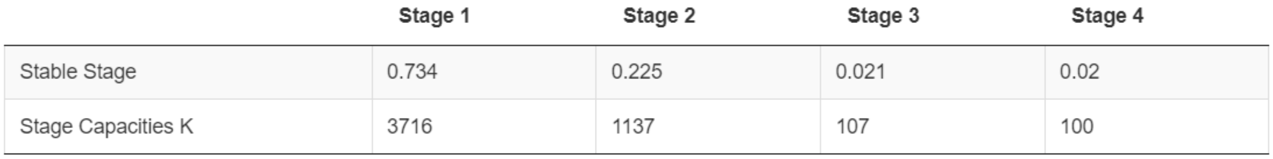
Stage-specific capacities (K) with compensation ratios:

- K (Stage-0, eggs): Calculated as the product of the number of individuals in all mature age classes, multiplied by maturation probabilities for each class, the number of spawning events, eggs per female, sex ratio and spawning interval.
- K (Stage-0, fry): K values for young-of-the-year (fry/Age-0) individuals are calculated as the product of K (Stage-0, eggs) * the egg survival (SE).
- K (Stage-1): Calculated from the stable-stage distribution of the transition matrix (B) after setting the adult stage (Stage-4) to K (e.g., 100).
- K (Stage-2): Calculated from the stable-stage distribution of the transition matrix (B) after setting the adult stage (Stage-4) to K (e.g., 100).

- K (Stage-3): Calculated from the stable-stage distribution of the transition matrix (B) after setting the adult stage (Stage-4) to K (e.g., 100).
- K (Stage-4): Manually input by the user for the population of interest. *This assumes that stage-4 is the only mature age class.*

Stable-stage distribution (portions) under the example species profile Figure 27:

Stable Stage Distributions (as a fraction 0 - 1) & Stage-specific Carrying Capacities (K):



	Stage 1	Stage 2	Stage 3	Stage 4
Stable Stage	0.734	0.225	0.021	0.02
Stage Capacities K	3716	1137	107	100

Density-Dependence Matrix (D):
Based on the derived stage-specific carrying capacities (K values), baseline survivorship values (SE, S0, surv_1, …) and the corresponding compensation ratios (cr_E, cr_0, cr_1, …), a density-dependence matrix (D) for a hypothetical population vector of eggs: 10,000,000, fry: 1,000,000; stage 1: 100,000, stage 2: 10,000, stage 3: 1,000 & stage 4: 100 will appear as follows:

The density-dependence matrix (D) contains vital rate modifiers for the estimated survivorship values at each stage transition. The density-dependence matrix (D) is multiplied with the corresponding transition matrix (B, Table 2) of density-independent transition probabilities. The finalized projection matrix (A) is the product of the density-dependent matrix (D), and the transition matrix (B) [A is a product of B*D = A]. The density-dependent matrix changes with each time step based on the number of individuals. The projection matrix (A) is therefore recalculated for each time step.

Compensation ratios are widely used as parameters in stock-recruitment functions, although they are admittedly less popular in classical matrix life cycle modelling. Steepness (*the proportion of recruitment produced when stock size is reduced to 20% of initial biomass*) is sometimes used in place of compensation ratios. Numerous other methods exist to introduce density dependence into stage-structured life cycle models. The compensation ratios are available as a default option for the CEMPRA tool to represent a versatile mechanism for applications to a large number of hypothetical species profiles. If location and stage-specific carrying capacities can be estimated, we recommend users set all compensation ratios to 1.0 for each stage class (therefore omitting compensation ratios) and refer to Section **Error! Reference source not found.** for classical density-dependent growth with Beverton-Holt functions. For additional background, please review the following references to learn more about compensation ratios and life cycle modelling with density-dependent growth.

Compensation Ratios in the Species Profile

Parameter	Description
cr_E **cr_0** **cr_1** **cr_2** **cr_3** **cr_4**	Compensation ratios for egg (cr_E), Age-0 fry (cr_0) and subsequent stage classes (cr_1 to cr_4). If the Nstage value is different than four, then add or remove rows accordingly. Compensation ratios can be set to 1.0 to omit the use of compensation ratios to govern density-dependent growth.

Useful references to understand Compensation Ratios:

- Goodyear, C. P. (1980). Compensation in fish populations. Biological monitoring of fish, 253-280.
- Myers, R. A. (2001). Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES Journal of Marine Science, 58(5), 937-951.
- Rose, K. A., Cowan Jr, J. H., Winemiller, K. O., Myers, R. A., & Hilborn, R. (2001). Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis. Fish and Fisheries, 2(4), 293-327.
- Myers, R. A., Bowen, K. G., & Barrowman, N. J. (1999). Maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences, 56(12), 2404-2419.
- Walters, C. J., & Martell, S. J. (2004). Fisheries ecology and management. Princeton University Press.
- Forrest, R. E., McAllister, M. K., Dorn, M. W., Martell, S. J., & Stanley, R. D. (2010). Hierarchical Bayesian estimation of recruitment parameters and reference points for Pacific rockfishes (Sebastes spp.) under alternative assumptions about the stock–recruit function. Canadian Journal of Fisheries and Aquatic Sciences, 67(10), 1611-1634.

7.5 Stochastic Simulations

Several additional parameters are available to influence the stochasticity (variability) of the population projections. Implementing these parameters is useful for understanding the viability of the population and (over many simulations) estimating the number of batch replicates that fall below a given critical threshold (e.g., X adults).

Standard Deviation in Eggs-per-Spawner

Parameter	Description
eps_sd	Standard Deviation in Eggs-per-Spawner controls the variability in fecundity across simulation years, and batch replicates. The example below shows a sample projection with eps_sd set to 250 and 750. In the

example with eps_sd set to 750, there are several years with very high fecundity. Density-dependent constraints (if implemented) may attenuate the apparent effect of high eps_sd inputs.

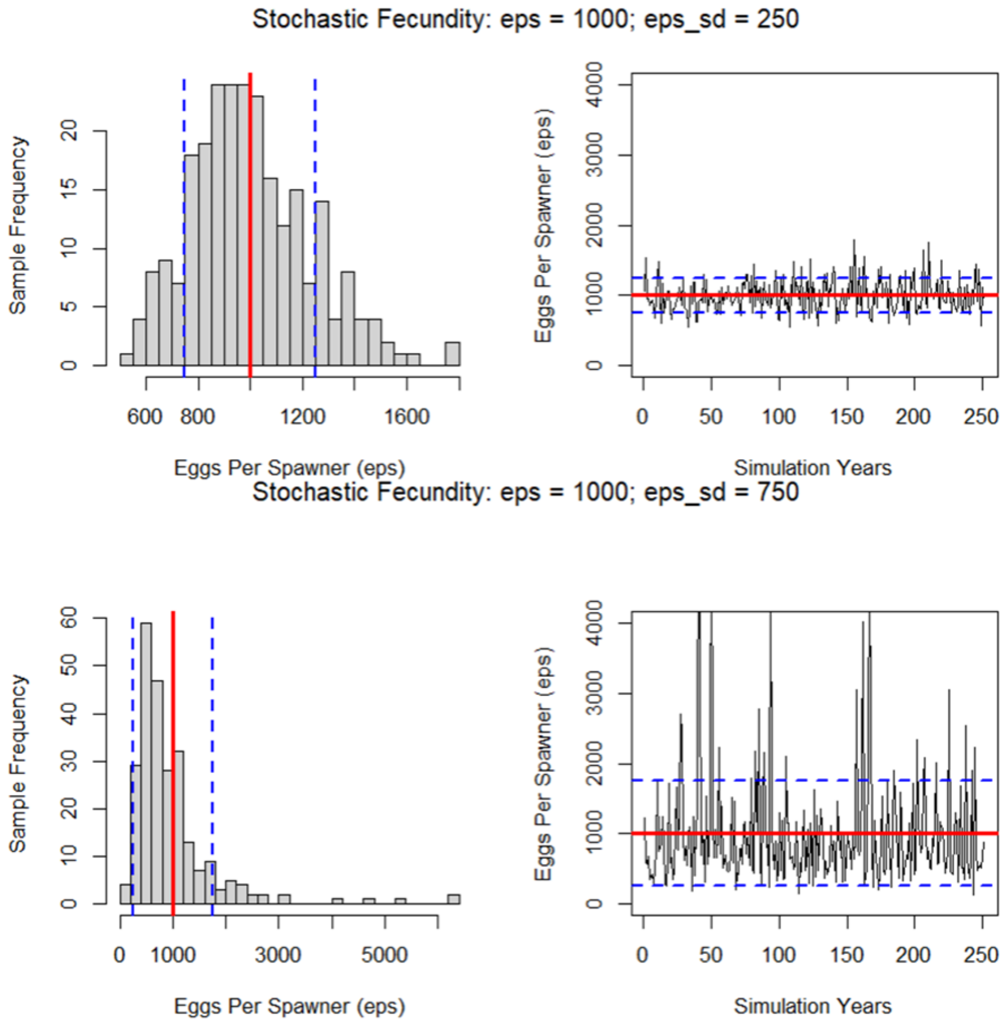
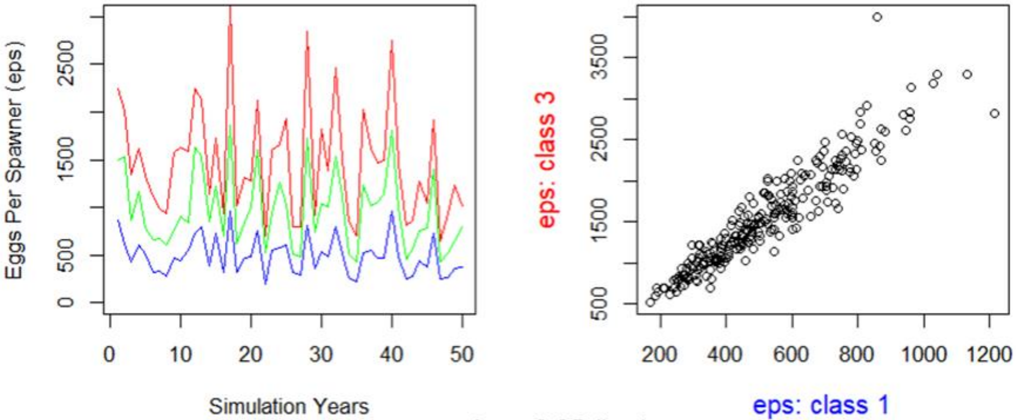


Figure 33. eps_sd Standard Deviation in Eggs-per-Spawner

Correlation in egg fecundity through time

Parameter	Description
egg_rho	egg_rho: correlation in egg fecundity through time: In natural populations, there will be good years and bad years. It’s assumed that good years will be good for large adults and small adults. If multiple mature stage classes contribute to spawning (fecundity) (i.e., maturity values are greater than 0), it is assumed that fecundity will be correlated between good and bad years across stage classes (i.e., stage_5, stage_6 & stage_7). egg_rho controls the degree of correlation in interannual fecundity between stage classes. See Figure 34 for an illustrative example. If egg_rho is low, and multiple stage classes contribute to spawning, then some stage classes may compensate for good/bad years. Conversely, if egg_rho is high, then the population may be highly volatile as all cohorts experience good/bad years simultaneously.


egg_rho = 0.95 (high)

egg_rho = 0.05 (low)

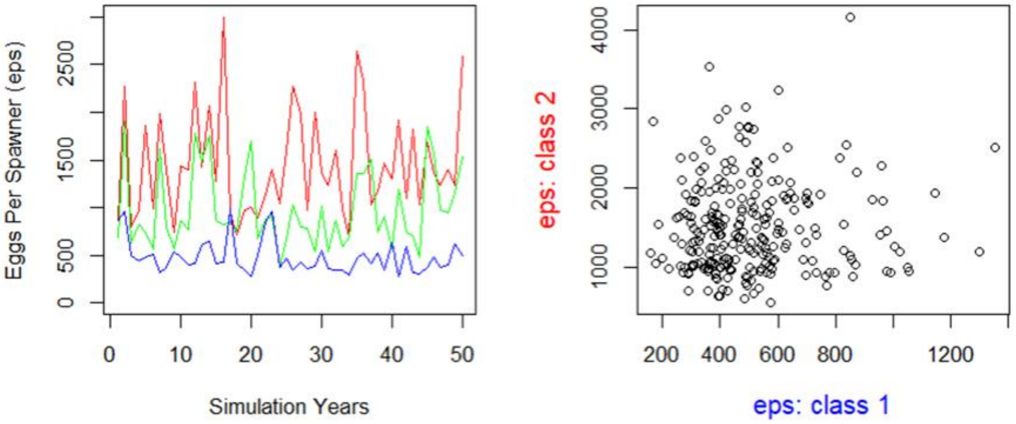


Figure 34. egg_rho: correlation in egg fecundity through time

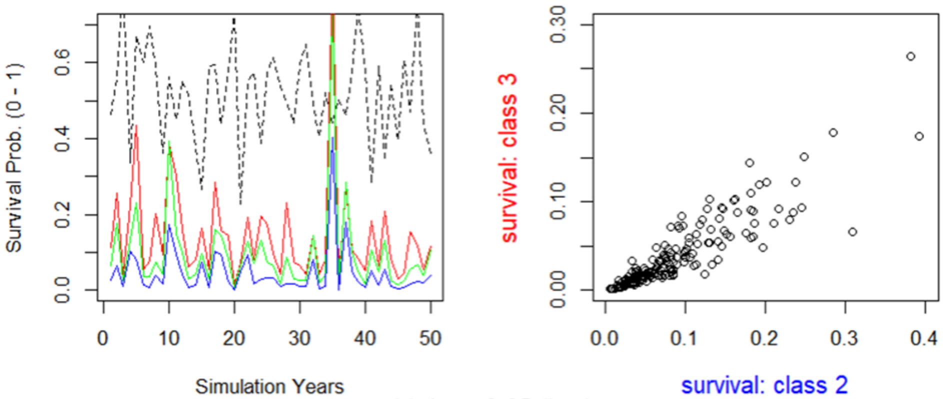
Coefficient of variation (CV) in interannual stage-specific mortality

Parameter	Description
M.cv	Coefficient of variation (CV) in stage-specific mortality based on a beta distribution.

Correlation in stage-class mortality through time

Parameter	Description
M.rho	M.rho, correlation in mortality through time: In natural populations, there will be good years and bad years. It's assumed that good and bad years will be correlated across all stages classes (excluding eggs, SE). M.rho controls the degree of correlation in interannual survivorship between stage classes. See Figure 34 for an illustrative example. If M.rho is low, then certain stage classes may compensate for good/bad years (based on a random sampling of survivorship). Conversely, if M.rho is high, then the population will become highly volatile as all cohorts experience good/bad years simultaneously.

M.rho = 0.95 (high)



M.rho = 0.05 (low)

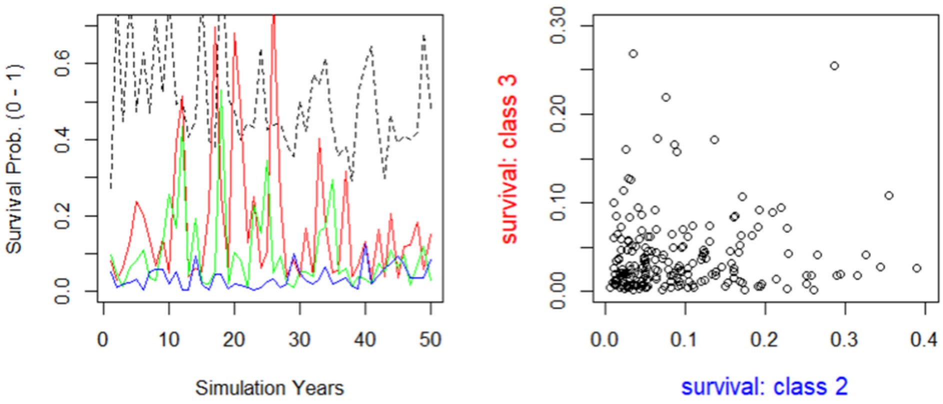


Figure 35. M.rho: correlation in survivorship through time

p.cat: Probability of Catastrophe per Generation

Parameter	Description
p.cat	p.cat: Probability of Catastrophe per Generation: Probability of catastrophe scaled to the average generation time of the population. *The annual probability of catastrophe is dependent on the generation time of the population.*

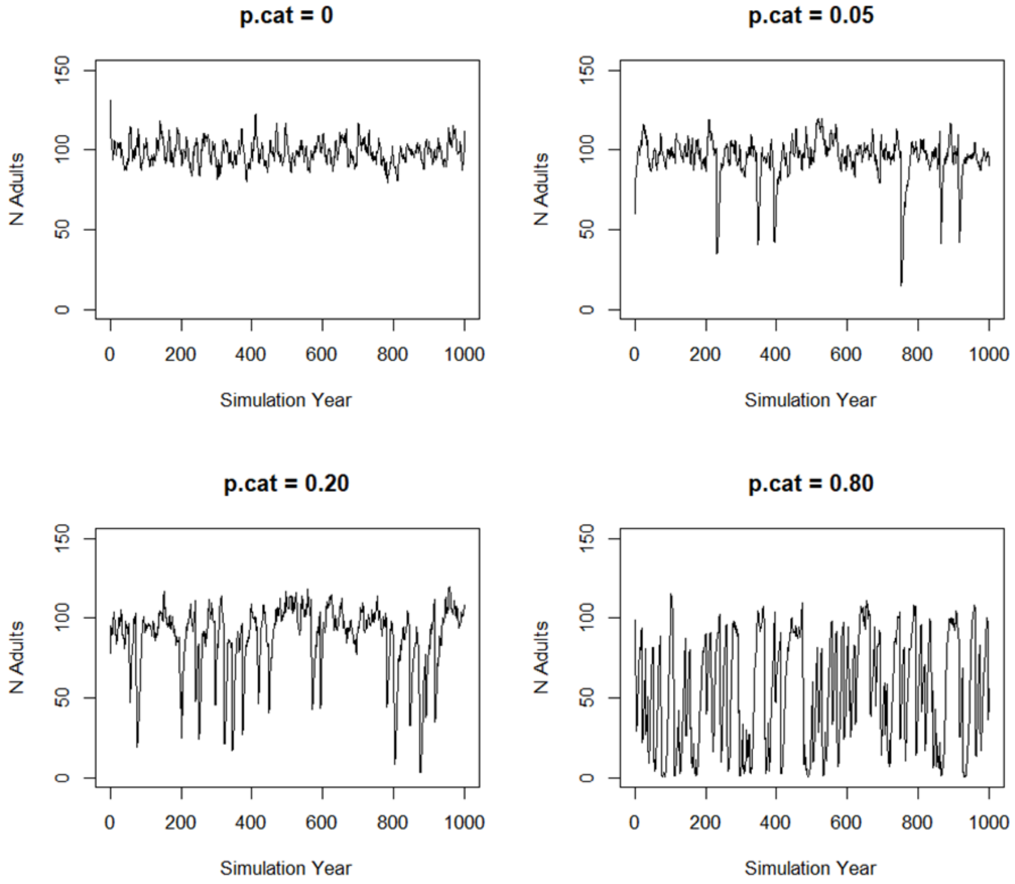


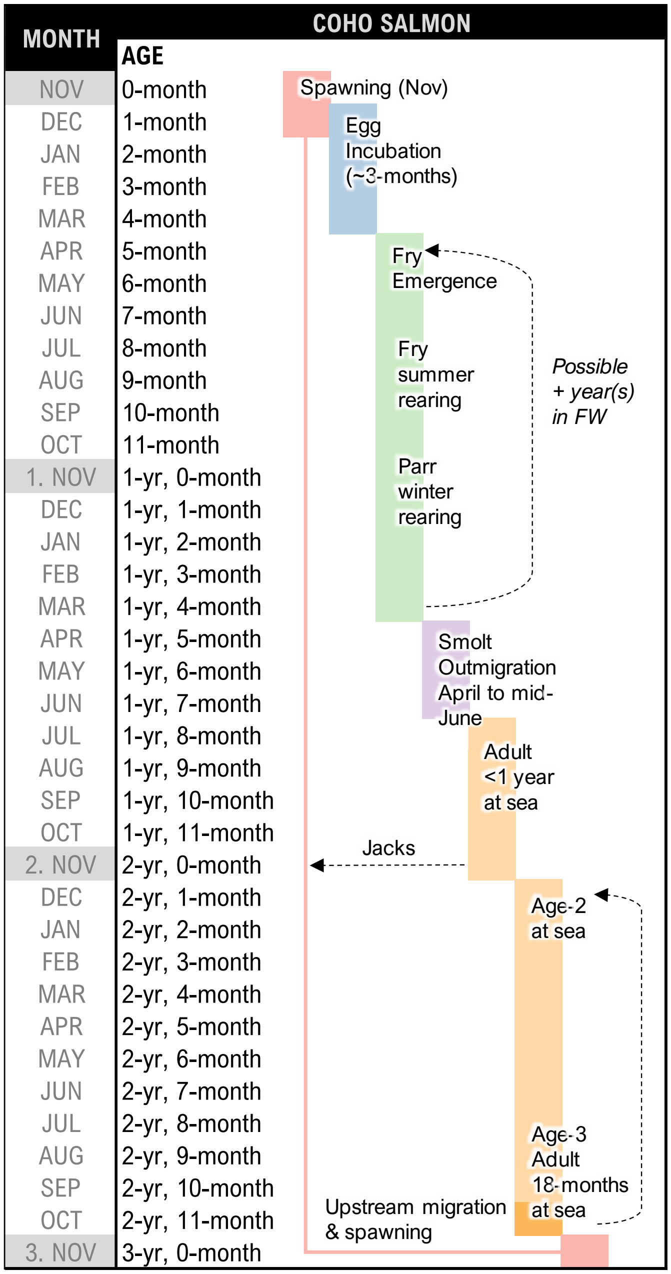
Figure 36. p.cat: Probability of Catastrophe per Generation

7.6 Building a Life Cycle Profile

A customized life cycle profile can be developed from the following template: https://github.com/essatech/JoeModelCEShiny/blob/main/data/demo/life%20cycles.csv

Appendix B includes sample species profiles for case study systems, including Nicola Basin Chinook Salmon, Steelhead and Coho Salmon. The COMPADRE (plant) and COMADRE (animal) online archives also have vital rates for 792 and 429 species as of March 2023 (https://compadre-db.org/Data/Comadre). The COMADRE archive can be a great place to review vital rate estimates for similar species and taxa to help get started.

We recommend first summarizing the species' life cycle using a simple diagram to represent key life stages. The diagram can then be converted to a periodicity table (or life history schedule) to represent time spent in each life stage (Figure 37). A life history diagram coupled with a periodicity table can be useful to help map out transitions between key stages in the life cycle model. Recall that the life cycle model works with annual time steps. “Census” periods within the matrix model do not need to fall precisely at the one-year interval, but care should be taken to ensure that key transitions are not missed or double-counted. It is best to start from the spawning period and work forward following a whole generation to its progeny.



MONTH	COHO SALMON AGE
NOV	0-month
DEC	1-month
JAN	2-month
FEB	3-month
MAR	4-month
APR	5-month
MAY	6-month
JUN	7-month
JUL	8-month
AUG	9-month
SEP	10-month
OCT	11-month
1. NOV	1-yr, 0-month
DEC	1-yr, 1-month
JAN	1-yr, 2-month
FEB	1-yr, 3-month
MAR	1-yr, 4-month
APR	1-yr, 5-month
MAY	1-yr, 6-month
JUN	1-yr, 7-month
JUL	1-yr, 8-month
AUG	1-yr, 9-month
SEP	1-yr, 10-month
OCT	1-yr, 11-month
2. NOV	2-yr, 0-month
DEC	2-yr, 1-month
JAN	2-yr, 2-month
FEB	2-yr, 3-month
MAR	2-yr, 4-month
APR	2-yr, 5-month
MAY	2-yr, 6-month
JUN	2-yr, 7-month
JUL	2-yr, 8-month
AUG	2-yr, 9-month
SEP	2-yr, 10-month
OCT	2-yr, 11-month
3. NOV	3-yr, 0-month

Figure 37. Sample combo life cycle diagram and periodicity table for a coastal population of Coho Salmon in British Columbia

A draft species profile can be loaded into the R-Shiny application and modified. User groups may find it valuable to reference the eigen analysis of the projection matrix to ensure that:

- Lambda (λ, instantaneous growth rates) estimates are reasonable (not substantially different from 1.0). If lambda estimates are lower than 1.0, review parameters.
- Generation time approximately matches the characterization of the species' life history in the literature.
- Stable-stage distributions from the projection matrix are not substantially different from expectations.

7.7 Benefits and Limitations

The life cycle modelling component of the CEMPRA tool is a valuable resource for understanding and quantifying cumulative effects, linking critical stressors to key life stages and supporting an understanding of the relevance of key drivers curtailing productivity and capacity. Related analyses have provided a useful framework to synthesize pathways for cumulative effects through the lens of population ecology (Beechie et al. 2021; Jorgensen et al. 2021; Sorel 2022; Kendall et al. 2023). Simulations allow user groups to play out hypothetical scenarios with multiple stressors, locations and species profiles, making it possible to perform many complex calculations within a simple user interface.

There are, however, key limitations to the CEMPRA tool and life cycle modelling writ large. Depending on how a species profile is parameterized, there can be key stages and vital rates that are highly sensitive to perturbation. It can be challenging to confirm if these sensitivities reflect key vulnerabilities in nature or artifacts of the modelling framework. Therefore, some researchers have emphasized the need to characterize life cycle models as hypothesis generators until predictions and causal pathways can be properly validated empirically (Roni et al. 2018).

The CEMPRA tool does not account for complex variations in life history strategies, seasonal movement, individual exposure and detailed habitat criteria. It's, therefore, up to the users to carefully design stressor variables, stressor magnitude datasets and stressor-response relationships such that linkages are already largely accounted for in the underlying input data. When summarizing results, we suggest that the user group focus on reviewing the relative difference between scenarios to a default (status-quo) reference scenario. Interpreting results as relative differences are far more relatable than trying to rationalize absolute values (e.g., *scenario A increased system capacity by ~10% relative to scenario B* vs *scenario A increased capacity by 321 fish relative to scenario B*).

User groups and practitioners should solicit life cycle models to help develop high-level goals for restoration and recovery programs (Roni et al., 2018). Life cycle models are useful for identifying key demographic bottlenecks and the sensitivity of those bottle necks to the range of stressor values observed on the landscape (or projected through simulation). Life cycle models will be less useful in efforts to design small-scale projects. Continued efforts should be made to validate and refine predictions with field data and empirical studies.

8 Stressor-Response Library/Database

8.1 About

The stressor-response library is designed as a public reference archive to house generalizable stressor-response functions for use in the Cumulative Effects Model for Prioritizing Recovery Actions (CEMPRA) tool. The functions archived in the stressor-response library are widely applicable to numerous applications and use cases.

Users of the CEMPRA tool can select a subset of stressor-response functions applicable to their study area and focal species of interest. Selected stressor-response functions can be used to construct a customized stressor-response input workbook for the CEMPRA tool. The stressor-response library is a valuable resource for reference as it can expedite assessments with shared knowledge of common impact pathways. Although a suite of default stressor-response functions is available, the stressor-response library will continually be updated with novel stressor-response functions uploaded, edited, and validated by the community of users.

Example stressor-response functions for Coho Salmon and Steelhead responses to temperature and sediment can be found in Appendix A.

8.1.1 Stressor-Response Library Schema

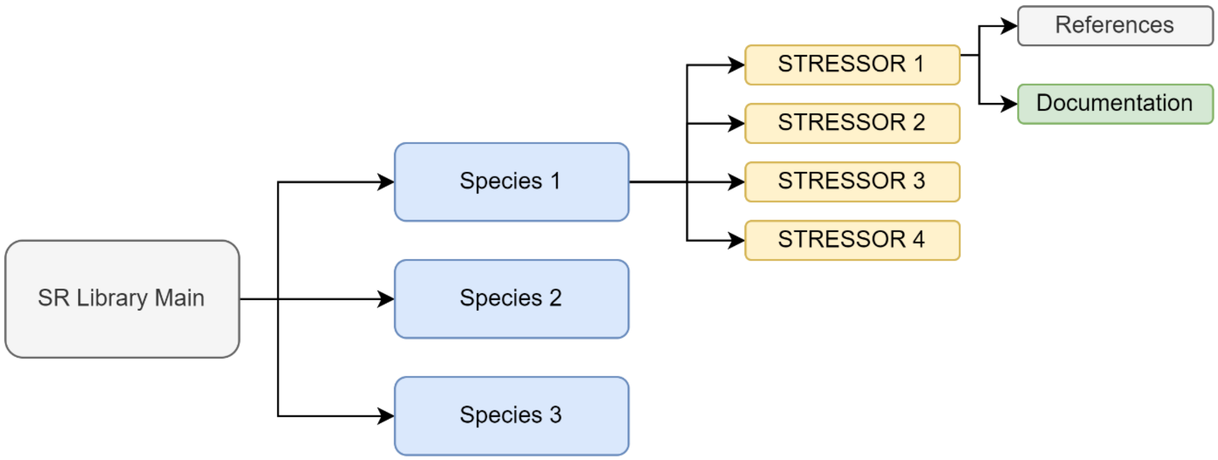


Figure 38: Data structure within the SR library, where the table hierarchy starts at the SR Library Main folder (left) and ends at the references and documentation for individual stressor-response functions (right).

Main Folder

The Main folder is the gateway to the stressor-response library. It contains the primary database and all of the species' folders (see below).

Species Folders

The main SR library is initially divided into folders for each of the currently defined species in the model. If a user develops a stressor-response function for a new species, a new folder/branch should be added to the library.

Stressor Folders

Stressor folders are located within each of the species' folders. The stressor folders contain the stressor-response function documentation for a given species and are named as follows:

SR curve ID_Stressor ID_Stressor Name

References

Within each of the stressor folders, the references folder contains copies of all the materials cited in the stressor-response function documentation.

8.2 Use in the Model

The primary stressor-response library (database) contains functions as individual Excel worksheets. These Excel worksheets are used to form a custom Excel SR workbook. The SR workbook is then used as one of the main inputs in the CEMPRA (see Data Inputs).

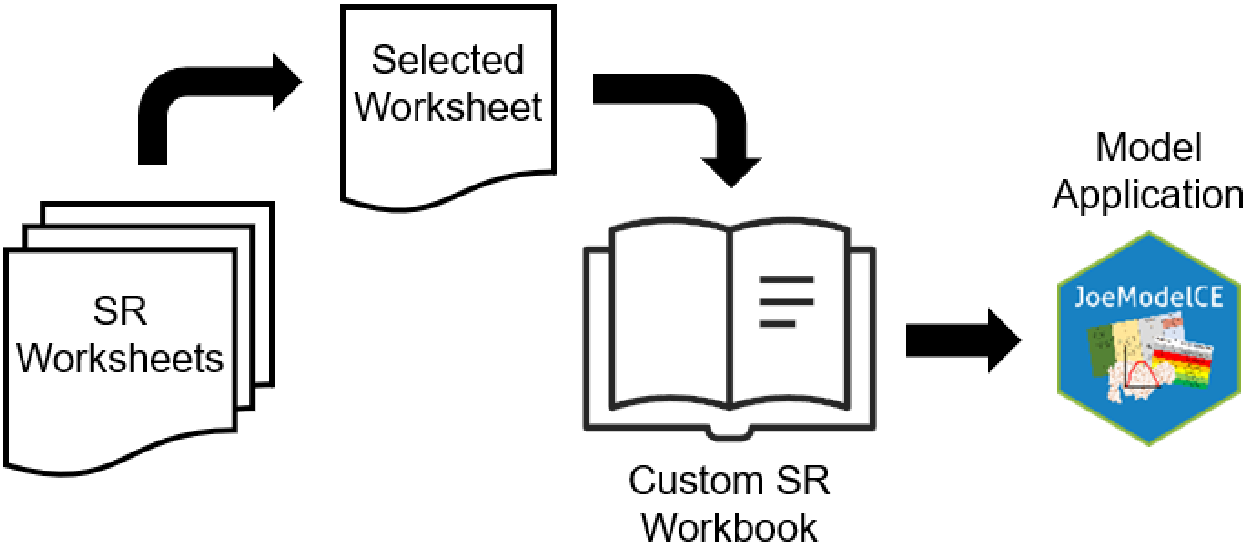


Figure 39: Use of the SR library in the CEMPRA (Joe Model). The SR library contains individual worksheets which can be selected to create a custom SR workbook for use in the model.

8.3 Stressor-Response Database

We are currently developing a formal web-based stressor-response database that users will be able to access remotely. A central online hub for all stressor-response functions will streamline collaboration and allow users around the world to share, create, and edit custom stressor-response functions. An interim stressor-response library is currently hosted on the following GitHub archive:

https://github.com/mattjbayly/stressor-response-library
** This resource link will be changed as the archive develops (please check for the most recent version of this documentation).*

Users can access the repository to get a general sense of the SR library schema and how it will connect to the CEMPRA in the future web-based database.

9 Case Study Applications

9.1 Athabasca Rainbow Trout

MacPherson et al. (2018) leveraged the Joe Modelling framework to support recovery action planning and threat analysis of Athabasca Rainbow Trout. The Joe Model was a valuable resource for synthesizing stressors due to its ease of and utility as a design and communication tool. Stressors were developed in a near real-time workshop-like setting to game out alternative management scenarios. The flexibility of the stressor-response input data allows for the inclusion of synthetic proxies and hypothetical impact pathways, weaving together academic knowledge, anecdotal descriptions and expert opinion. The Joe Model was ultimately used as a prioritization tool to help design actions, programs, and priorities across a large geographic and complex region.

The default study area, stressor list and species profile in the Shiny application are from the MacPherson et al. 2018 Athabasca Rainbow Trout project.

9.2 Nicola Basin Chinook Salmon, Coho Salmon, and Steelhead

The Nicola River provides important spawning, incubation, and rearing habitat to multiple salmonid species, including Chinook Salmon, Coho Salmon, and Steelhead. A case study application of the CEMPRA tool is being applied to Nicola Basin to support threat analysis and recovery action planning.

Appendix B includes species profiles for Coho and Steelhead in the Nicola Basin. An initial analysis is underway working with stream temperature, sedimentation and flows to link stressors to critical components of the life cycle for these species (Appendix A).

10 R Package Custom Applications

The *JoeModelCE* R Package (available on GitHub: https://github.com/essatech/JoeModelCE/) can be used for custom model applications and batch processing. The R Package allows users to run the model offline and customize the code, inputs, and outputs. Users can also conduct batch runs of the life cycle model, which allows for faster, automated processing over a large number of spatial units and/or scenarios. For more information about the use and installation of the *JoeModelCE* R Package, see the following tutorials:

- **Tutorial 1: Joe Model Overview:** https://essatech.github.io/JoeModelCE/articles/joe-model.html
- **Tutorial 2: Life cycle model Overview:** https://essatech.github.io/JoeModelCE/articles/population-model.html
- **Tutorial 3: Life cycle model Batch Run:** https://essatech.github.io/JoeModelCE/articles/population-model-batch-run.html
- **Tutorial 4: Joe Model Batch Run:** https://essatech.github.io/JoeModelCE/articles/joe-model.html

11 Concluding Remarks

11.1 Benefits and Limitations

The CEMPRA tool has the flexibility to perform cumulative effects assessments for both data-rich and data-poor species, allowing for the analysis of a range of study systems, including species-at-risk. Various assessment endpoints are available for different use cases. With only four generic input datasets, onboarding new users is intended to be relatively simple.

The Joe Modelling component of the CEMPRA tool is a generic assessment framework widely applicable to data-rich and data-limited species. The Joe Model acts as a simplified stressor roll-up framework for watershed assessments. Stressors can be added and manipulated in near real-time in a workshop-like setting. Proven use cases on numerous study systems show how it can be a critical asset to facilitate a shared understanding of watershed stewardship and high-level strategy development for watershed planning.

The life cycle modelling component of the CEMPRA tool offers the ability to link stressors to key vital rates in a demographic modelling framework. User groups and practitioners can solicit the life cycle models to help develop high-level goals for restoration and recovery programs (Roni et al., 2018). Life cycle models are useful to identify key demographic bottlenecks and the sensitivity of those bottlenecks to the range of stressor values observed on the landscape (or projected through scenario simulation).

The CEMPRA framework should be considered a simple approach to cumulative effects assessments. It provides the greatest utility for large-scale planning activities at the intersection between different user groups and stakeholders. The intent of the CEMPRA tool is to provide a working framework for both data-rich and data-limited species with alternative assessment endpoints. Stressor-response relationships are the underlying theme of the CEMPRA tool. Our hope is that the CEMPRA will support the development of a collaborative user community, reducing duplicated efforts in the development of stressor variables, species profiles and assessment frameworks.

13 References

Alexander, S. M., Provencher, J. F., Henri, D. A., Taylor, J. J., Lloren, J. I., Nanayakkara, L., ... & Cooke, S. J. 2019. Bridging Indigenous and science-based knowledge in coastal and marine research, monitoring, and management in Canada. Environmental Evidence, 8(1), 1-24.

Bear, E.A., T.T. McMahon, and A.V. Zale. 2007. Comparative thermal requirements for westslope cutthroat trout and rainbow trout: Implications for species interactions and development of thermal protection standards. Transactions of the American Fisheries Society 136:1113-1121. doi: 10.1577/T06-072.1

Beechie, T. J., C. Nicol, C. Fogel, J. Jorgensen, and B. Timpane-Padgham. 2021. Chum Salmon Life-Cycle Model Description and Results for the Chehalis River Basin. U.S. Department of Commerce, NOAA Contract Report NMFS-NWFSC-CR-2021-02.

Beechie, T. J., C. Nicol, C. Fogel, J. Jorgensen, J. Thompson, G. Seixas, J. Chamberlin, J. Hall, B. Timpane-Padgham, P. Kiffney, S. Kubo, and J. Keaton. 2021. Modeling Effects of Habitat Change and Restoration Alternatives on Salmon in the Chehalis River Basin Using a Salmonid Life-Cycle Model. U.S. Department of Commerce, NOAA Contract Report NMFS-NWFSC-CR-2021-01.

Houde, N., 2007. The six faces of traditional ecological knowledge: challenges and opportunities for Canadian co-management arrangements. Ecology and Society, 12(2).

Jensen, D. W., E. A. Steel, A. H. Fullerton, and G. R. Pess. 2009. Impact of fine sediment on incubation survival of Pacific salmon: a meta-analysis of published studies. Reviews in Fisheries Science 17(3):348-359. doi: 10.1080/10641260902716954.

Jorgensen, J. C., Nicol, C., Fogel, C., & Beechie, T. J. (2021). Identifying the potential of anadromous salmonid habitat restoration with life cycle models. Plos one, 16(9), e0256792.

Kendall, N. W., Unrein, J., Volk, C., Beauchamp, D. A., Fresh, K. L., & Quinn, T. P. (2023). Life Cycle Model Reveals Sensitive Life Stages and Evaluates Recovery Options for a Dwindling Pacific Salmon Population. North American Journal of Fisheries Management, 43(1), 203-230.

MacPherson, L., Sullivan, M., Reilly, J., & Paul, A. (2020). Alberta's Fisheries Sustainability Assessment: A Guide to Assessing Population Status, and Quantifying Cumulative Effects Using the Joe Modelling Technique. Canadian Science Advisory Secretariat (CSAS).

Murray, A.M., Sullivan, M.G., & Acorn, J. 2012. A tribute to Joseph Schieser Nelson, 1937–2011. Can. Field-Nat. 125: 373–380.

Pearsall et al. 2022. Nicola Watershed RAMS Final Report 2022: Nicola Basin Chinook Risk Assessment Final Report. A Process Developed with the Nicola Collaborative Research and Technical Committee.

Piet, G. J., Tamis, J. E., Volwater, J., de Vries, P., van der Wal, J. T., & Jongbloed, R. H. (2021). A roadmap towards quantitative cumulative impact assessments: every step of the way. Science of The Total Environment, 784, 146847.

Roni, P., Anders, P. J., Beechie, T. J., & Kaplowe, D. J. (2018). Review of tools for identifying, planning, and implementing habitat restoration for Pacific salmon and steelhead. North American Journal of Fisheries Management, 38(2), 355-376. Chicago

Schaub, M., & Kéry, M. (2021). Integrated life cycle models: Theory and ecological applications with R and JAGS. Academic Press.

Sorel, M. (2022). Informing salmon conservation with life cycle models that account for individual heterogeneity (Doctoral dissertation, University of Washington).

Sullivan, M. 2017. Adaptive Management of Athabasca Rainbow Trout; cumulative effects modelling and potential management actions. Alberta Environment and Parks; F isheries Management Branch. January 2017

Van der Lee, A. S., & Koops, M. A. (2020). Recovery Potential Modelling of Westslope Cutthroat Trout (Oncorhynchus clarkii lewisi) in Designatable Unit 1: Saskatchewan-Nelson River populations. Canadian Science Advisory Secretariat.

Walters, C. 1997. Challenges in adaptive management of riparian and coastal ecosystems. Conservation Ecology [online]1(2):1.

Ward, B. R., & Slaney, P. A. 1993. Egg-to-smolt survival and fry-to-smolt density dependence of Keough River Steelhead Trout. Canadian Special Publication of Fisheries and Aquatic Sciences, 209-217.

14 Appendices

Appendix A: Stressor-response function for Nicola Coho Salmon and Steelhead

Appendix B: Species profiles for Nicola Coho Salmon Steelhead

Appendix C: CEMPRA Extensions: Multi-stressor-interaction Matrix

Appendix D: R-package tutorials and package manual