NOAA Technical Memorandum NMFS-XXX-##

CPS Acoustic Classification

Alice Beittel

January 2023

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northwest Fisheries Science Center

CPS Acoustic Classification

Alice Beittel[1]

1. NOAA Southwest Fisheres Science Center, Southwest Fisheries Science Center

Table of contents

List of Figures

List of Tables

Welcome

Last updated: 2024-11-21 23:59:22 UTC

[[[enter cool image of CPS survey]]]

West Coast coastal pelagic species play an important role in the California Current ecosystem. They're food sources for marine mammals, sea birds, and larger fish, and they support commercial and recreational fisheries. The biomass and abundance estimates derived from this project are used in stock assessment models to support sustainable fisheries.

Document Objective:

This resource will serve as a tutorial to demonstrate how the SWFSC uses acoustic data generate biomass estimates of Coastal Pelagic Species from Baja, Mexico to Vancouver, Canada.

As part of our commitment to open science, reproducibility, and transparency, we provide this metadata guide to compliment our public-domain data.

Please consider this resource to be a **Living Document**. The code in this repository is regularly being updated and improved.

Do not hesitate to reach out (to us at either alice.beittel@noaa.gov or GitHub issues, especially if you find discrepancies in the data or want to suggest improvements to infrastructure. Thank you in advance for your collaboration and partnership with us as we develop our future data universe.

User Resources

Cite This Data

[enter text on how to do this]

NOAA README

This repository is a scientific product and is not official communication of the National Oceanic and Atmospheric Administration, or the United States Department of Commerce. All NOAA GitHub project code is provided on an 'as is' basis and the user assumes responsibility for its use. Any claims against the Department of Commerce or Department of Commerce bureaus stemming from the use of this GitHub project will be governed by all applicable Federal law. Any reference to specific commercial products, processes, or services by service mark, trademark, manufacturer, or otherwise, does not constitute or imply their endorsement, recommendation or favoring by the Department of Commerce. The Department of Commerce seal and logo, or the seal and logo of a DOC bureau, shall not be used in any manner to imply endorsement of any commercial product or activity by DOC or the United States Government.

NOAA License

1 Survey Background

1.1 Who conducts the survey?

The California Current Ecosystem Survey is conducted by researchers at the NOAA Southwest Fisheries Science Center from the Fisheries Resources Division. The survey is also made possible by volunteers from additional NOAA line offices and science centers, universities, international partners, NOAA interns, and inter-agency employees.

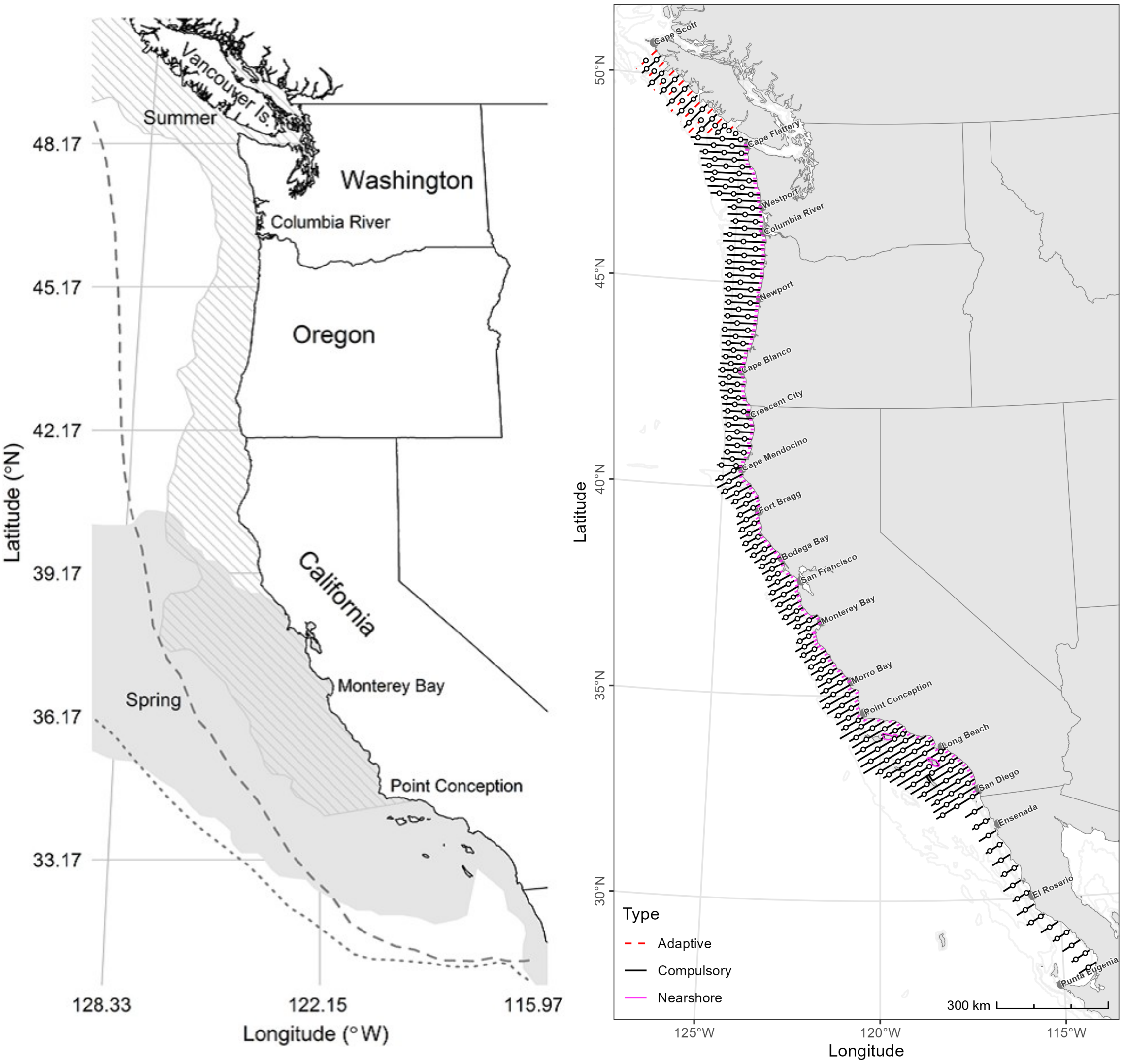
1.2 Where does the survey take place?

1.3 Research objectives:

1. Acoustically map the distributions, measure the species compositions and size-frequency distributions, and estimate the abundances and biomasses of CPS present in the survey area, e.g., Pacific Sardine Sardinops sagax, Northern Anchovy (*Engraulis mordax*), Pacifc Herring (*Clupea pallasii*), Round Herring (*Etrumeus acuminatus*), Pacific Mackerel (*Scomber japonicus*), and Jack Mackerel (*Trachurus symmetricus*)
2. Characterize and investigate linkages to their biotic and abiotic environments
3. Gather information regarding their life histories
4. Compare the species composition and size distributions of trawls and near shore vessel purse seine sets.

1.4 Survey History:

The SWFSC's ATM surveys of CPS in the CCE began in 2006 with a focus on the northern stock of Pacific Sardine. Since then, they have expanded in scope and objectives to



(a) Sardine distribution

(b) General sampling scheme

Figure 1.1: On left, a conceptual spring (shaded region) and summer (hashed region) distributions of potential habitat for the northern stock of Pacific Sardine along the west coasts of Mexico, the United States, and Canada. On right, the general sampling scheme of planned core-region (solid black lines), adaptive (dashed red lines), and nearshore lines (pink).

include the larger forage-fsh assemblage and krill. This evolution, and the migratory behavior of Pacific Sardine, serve to explain the present survey region and design.

1.5 Code of Conduct

2 Data Acquisition

2.1 Survey Equipment

2.1.1 Acoustic Instruments

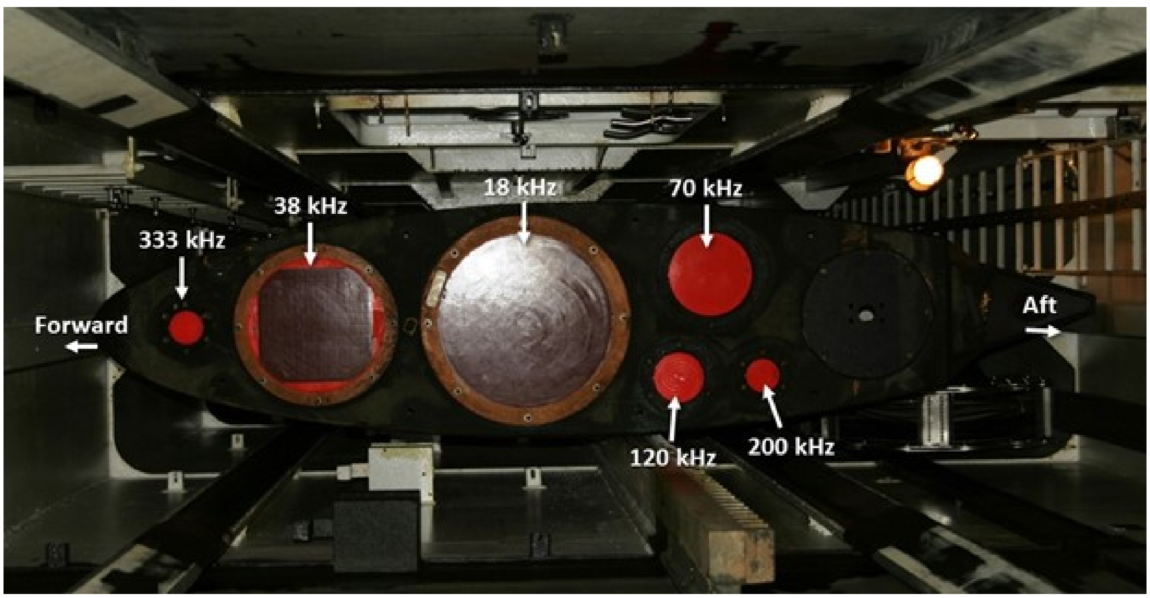


Figure 2.1: Transducer locations on the bottom of the centerboard aboard Lasker.

On *Lasker* and *Shimada*, multi-frequency Wideband Transceivers (Simrad EK80 WBTs; Kongsberg) were confgured with split-beam transducers (Simrad ES18, ES38-7, ES70-7C, ES120-7C, ES200-7C, and ES333- 7C on *Lasker* and ES18, ES38B, ES70-7C, ES120-7C, and ES200-7C on *Shimada*; Kongsberg). The transducers were mounted on the bottom of a retractable keel or "centerboard". The keel was retracted (transducers ~5-m depth) during calibration, and extended to the intermediate position (transducers ~7-m depth) during the survey. Exceptions were made during shallow water operations, when the keel was re- tracted; or during times of heavy weather, when the keel was extended (transducers ~9-m depth) to provide extra stability and reduce the efect of weather-generated noise. Transducer position and motion were measured at 5 Hz using an inertial motion unit (Applanix POS-MV; Trimble).

2.1.2 Underway CTD

On *Lasker* and *Shimada*, conductivity and temperature profiles were measured down to 300 m using calibrated sensors on a probe cast from the vessel while underway (UnderwayCTD, or UCTD; Teledyne Ocean- science). Casts were typically conducted between two to four times along each transect. These data indicate the depth of the surface mixed layer, above which most epipelagic CPS reside during the day. These data were also used to estimate the time-averaged sound speed (Demer, 2004), for estimating ranges to the sound scatterers, and frequency-specifc sound absorption coefcients, for compensating signal attenuation of the sound pulse between the transducer and scatterers (Simmonds and MacLennan, 2005).

2.2 Software

2.2.1 Echosounder Software

EK80

2.2.2 NetTime

On *Lasker* and *Shimada*, the computer clocks were synchronized with the GPS clock (UTC) using a synchronization software called NetTime.

2.2.3 EAL

The 38-, 70-, 120-, 200-, and 333-kHz echosounders were controlled by the EK80 Adaptive Logger (EAL2, Renfree and Demer, 2016). The EAL optimizes the pulse interval based on the seabed depth, while avoiding aliased seabed echoes, and was programmed such that once an hour the echosounders would record three pings in passive mode, for obtaining estimates of the background noise level.

2.2.4 K Sync

To minimize acoustic interference on *Lasker* and *Shimada*, transmit pulses from the EK80s, acoustic Doppler current profler and echosounder (Simrad-Kongsberg EC150-3C), multibeam echosounder (Simrad- Kongsberg ME70), imaging sonar (Simrad-Kongsberg MS70), scanning sonar (Simrad-Kongsberg SX90), and a separate acoustic Doppler current profler (Teledyne RD Instruments OS75 ADCP) were triggered using a synchronization system (Simrad K-Sync; Kongsberg). The K-Sync trigger rate, and thus the echosounder ping interval, was modulated by the EAL using the 18-kHz seabed depth provided by the Scientifc Computing System (SCS).

2.3 Raw Acoustic Data Format

Measurements of volume backscattering strength (S_v; dB re 1 m2 m-3) and target strength (*TS*; dB re 1 m2), indexed by time and geographic positions provided by GPS receivers, were stored in Simrad-Kongsberg .raw format with a 1-GB maximum fle size. During daytime, the echosounders operated in CW mode and logged to 60 m beyond the detected seabed range or to a maximum range of 500, 500, 500, 300, and 150 m for 38, 70, 120, 200, and 333 kHz, respectively. During nighttime, the echosounders operated in FM mode and logged to 100 m. For each acoustic instrument, the prefx for each fle name is a concatenation of the survey name (e.g., 2307RL), the operational mode (CW or FM), and the logging commencement date and time from the EK80 software (v21.15.1). For example, a fle generated by the Simrad-Kongsberg EK80 software for a WBT operated in CW mode is named 2307RL_CW-D20220826-T155651.raw.

3 Data Workflow

3.0.1 Data pipeline from boat to shore to report

4 Data Preparation

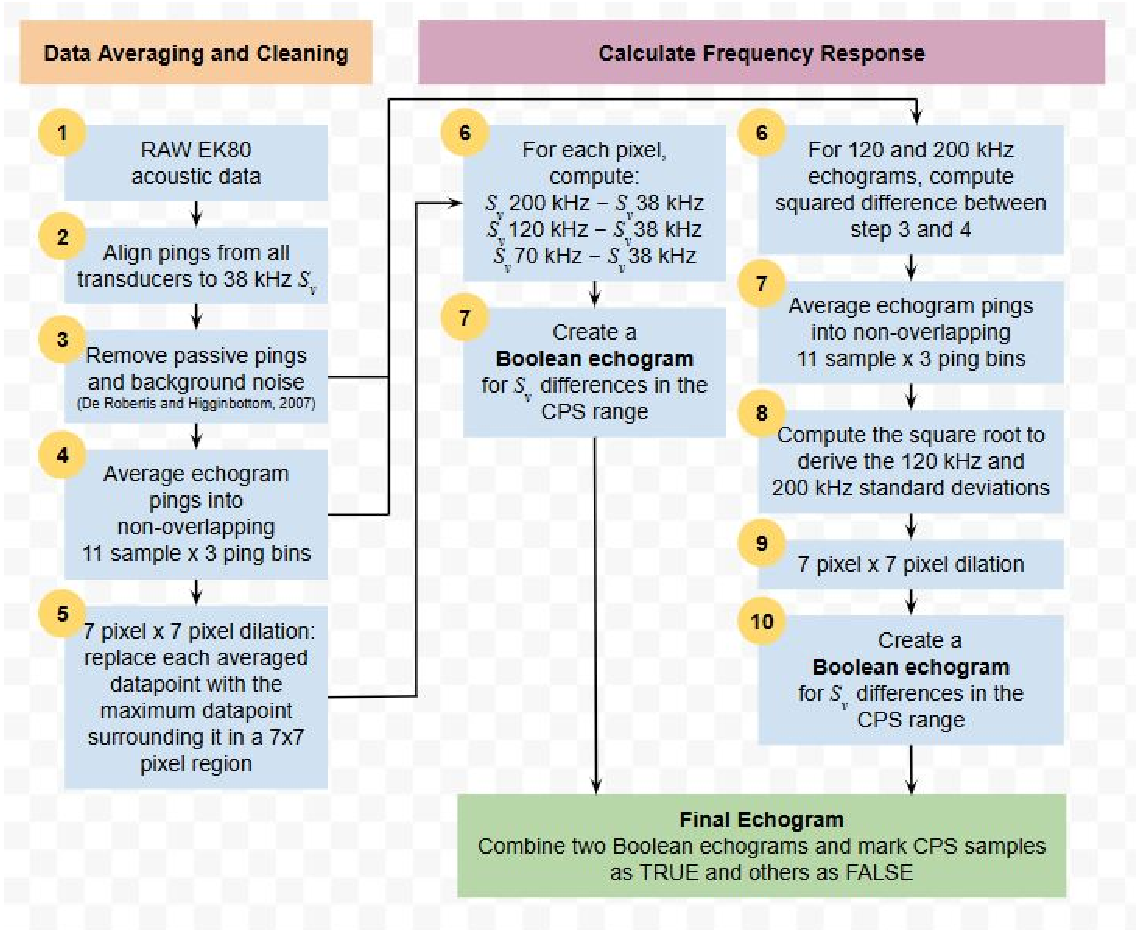
4.0.1 Select regions of interest

4.0.1.1 Select data on transect lines

4.0.1.2 Integration stop and start lines

5 Data Processing

We identify acoustic echos of schooling CPS using a semi-automated data processing algorithm implemented using Echoview software and in-house `Posit` code in the `estimATM` repository. The filters and thresholds were based on a subsample of echoes from randomly selected CPS schools. The aim of the filter criteria is to retain at least 95% of the noise-free backscatter from CPS while rejecting at least 95% of the non-CPS backscatter .



5.0.1

5.0.2 Super Detailed Processing Steps

5.0.2.1 *Automated Processing in Echoview*

5.0.2.2 Data Averaging and Cleaning:

1. Enter RAW EK80 acoustic data into Echoview and load ECS file.
2. Match geometry of all Sv variables to the 38-kHz Sv. This aligns pings from all transducers to 38-kHz.
3. Remove passive-mode pings and background noise:Estimate and subtract background noise using the background noise removal function (De Robertis and Higginbottom, 2007) in Echoview (Figs. 7b, e).

4. Average the noise-free Sv echograms using non-overlapping 11-sample by 3-ping bins.
5. Expand the averaged, noise-reduced Sv echograms with a 7 pixel x 7 pixel dilation.

5.0.2.3 Calculate Frequency Response:

6a. For each pixel, compute: $S_{v,200kHz} - S_{v,38kHz}$, $S_{v,120kHz} - S_{v,38kHz}$, and $S_{v,70kHz} - S_{v,38kHz}$;

7a. Create a Boolean echogram for Sv differences in the CPS range: $-13.85 < S_{v,70kHz} - S_{v,38kHz} < 9.89$ and $-13.5 < S_{v,120kHz} - S_{v,38kHz} < 9.37$ and $-13.51 < S_{v,200kHz} - S_{v,38kHz} < 12.53$

6b. For 120 and 200 kHz, compute the squared difference between the noise-filtered Sv (Step 3) and averaged Sv (Step 4)

7b. Average the results using an 11-sample by 3-ping window to derive variance

8b. Compute the square root to derive the 120- and 200-kHz standard deviations (σ_{120kHz} and σ_{200kHz}, respectively)

9b. Expand the standard deviation echograms with a 7 pixel x 7 pixel dilation;

10. Create a Boolean echogram based on the standard deviations in the CPS range: σ_{120kHz} > -65 dB and σ_{200kHz} > -65 dB. Diffuse backscattering layers have low σ (Zwolinski et al., 2010) whereas fish schools have high σ. Intersect the two Boolean echograms to create an echogram with "TRUE" samples for candidate CPS schools and "FALSE" elsewhere. Mask the noise-reduced echograms using the CPS Boolean echogram (Figs. 7c, f);

5.0.2.4 *Manual Processing in Echoview*

13. Create an integration-start line 5 m below the transducer (~10 m depth);
14. Create an integration-stop line 3 m above the estimated seabed (Demer et al., 2009), or to the maximum logging range (e.g., 350 m), whichever is shallowest;
15. Set the minimum Sv threshold to -60 dB (corresponding to a density of approximately three 20-cm-long Pacific Sardine per 100 m^3);

16. Integrate the volume backscattering coefficients (sV , m2 m-3) attributed to CPS over 5-m depths and averaged over 100-m distances;
17. Output the resulting nautical area scattering coefficients (sA; m2 nmi-2) and associated information from each transect and frequency to comma-delimited text (.csv) files.

6 Figures and Tables

Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

6.1 Code

You can embed an R code chunk like this:

```
summary(cars)
```

```
     speed           dist       
 Min.   : 4.0   Min.   :  2.00  
 1st Qu.:12.0   1st Qu.: 26.00  
 Median :15.0   Median : 36.00  
 Mean   :15.4   Mean   : 42.98  
 3rd Qu.:19.0   3rd Qu.: 56.00  
 Max.   :25.0   Max.   :120.00  
```

6.2 Including Plots

You can also embed plots and reference them, like so Figure 6.1.

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

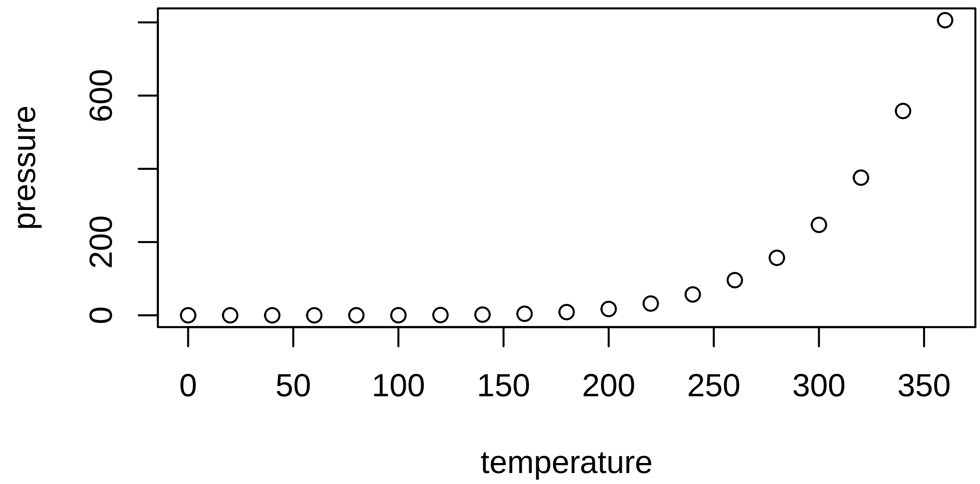


Figure 6.1: Plot of pressure

6.3 Including Tables

You can also embed tables and reference them with Table 6.1.

```
library(knitr)
kable(head(iris))
```

Table 6.1: Iris Data

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa

7 Rendering with Code

You can have code (R, Python or Julia) in your qmd file. You will need to have these installed on your local computer, but presumably you do already if you are adding code to your qmd files.

```r
x <- c(5, 15, 25, 35, 45, 55)
y <- c(5, 20, 14, 32, 22, 38)
lm(x ~ y)
```

```

Call:
lm(formula = x ~ y)

Coefficients:
(Intercept)            y
      1.056        1.326
```

7.1 Modify the GitHub Action

You will need to change the GitHub Action in `.github/workflows` to install these and any needed packages in order for GitHub to be able to render your webpage. The GitHub Action install R since I used that in `code.qmd`. If you use Python or Julia instead, then you will need to update the GitHub Action to install those.

If getting the GitHub Action to work is too much hassle (and that definitely happens), you can alway render locally and publish to the `gh-pages` branch. If you do this, make sure to delete or rename the GitHub Action to something like

```
render-and-publish.old_yml
```

so GitHub does not keep trying to run it. Nothing bad will happen if you don't do this, but if you are not using the action (because it keeps failing), then you don't need GitHub to run it.

7.2 Render locally and publish to gh-pages branch

To render locally and push up to the `gh-pages` branch, open a terminal window and then `cd` to the directory with the Quarto project. Type this in the terminal:

```
quarto render gh-pages
```

8 References

Quarto has powerful references functionality. You can easily insert citations from Zotero libraries that you maintain in the cloud (on Zotero). This allows the whole team to update the library and you can sync up to that library. Read about this on the Quarto documentation on citations. Google youtube videos on this also to see it in action.

Add a `.bib` file in to your project or add a linked Zotero library via RStudio in Visual mode with Tools > Project Options... > R Markdown > select custom libraries from the Zotero dropdown.

The you can type `@` and you will see a dropdown of the references in your libraries. You can then select the ones to add. If you don't see the one you need, you can paste in the DOI and it will be added to your references file (with all the info). The references will be added to your references section of your book automatically.

See the `references.qmd` file for how to include the references.

- `@ansley1981` will produce Ansley and Davis (1981)
- `[@ansley1981]` will produce (Ansley and Davis 1981).

References

Ansley, H. L. H., and C. D. Davis. 1981. "Migration and Standing Stock of Fishes Associated with Artificial and Natural Reefs on Georgia's Outer Continental Shelf." Brunswick, Georgia, USA.