Summary:

Juvenile Atlantic cod (*Gadus morhua*) in coastal Newfoundland settle to nearshore habitats in multiple (3-6) discrete events, or pulses, annually. These pulses produce a size-structured population in the first year of life that varies from year to year. We use finite mixture distribution models to classify pulse structure for age-0 and age-1 Atlantic cod using length data from coastal Newfoundland, Canada. Improved understanding of fish ecology requires knowledge of size structure within a population context, and finite mixture distribution models provide a reliable technique for determining distinct groups within a multi-modal distribution across time-series data (Macdonald and Pitcher, 1979).

Mixture models were derived using the *mixdist* package in R (Macdonald and Du, 2018; R Core Team, 2022), which fits finite mixture distributions using MLE and a combination of Newton-type and expected-maximization (EM) algorithms. Mixture models offer an important tool in assessing size-structure within a population because of their cluster-based approaches for instances with undefined subpopulations (McLachlan and Peel, 2000). We estimated mixture distributions (μ, σ, and π) to age-0 and age-1 cod for each sampling interval across each cohort. To illustrate the approach, we review the 2021 cohort assignment, which provided a model year with clearly defined pulse structure. This document contains one example of age-1 mixture distributions and two examples of age-0 mixture distributions. We used the mixture distribution groups as a baseline for pulse assignment and determined final pulse assignment based on growth trajectories and abundance.

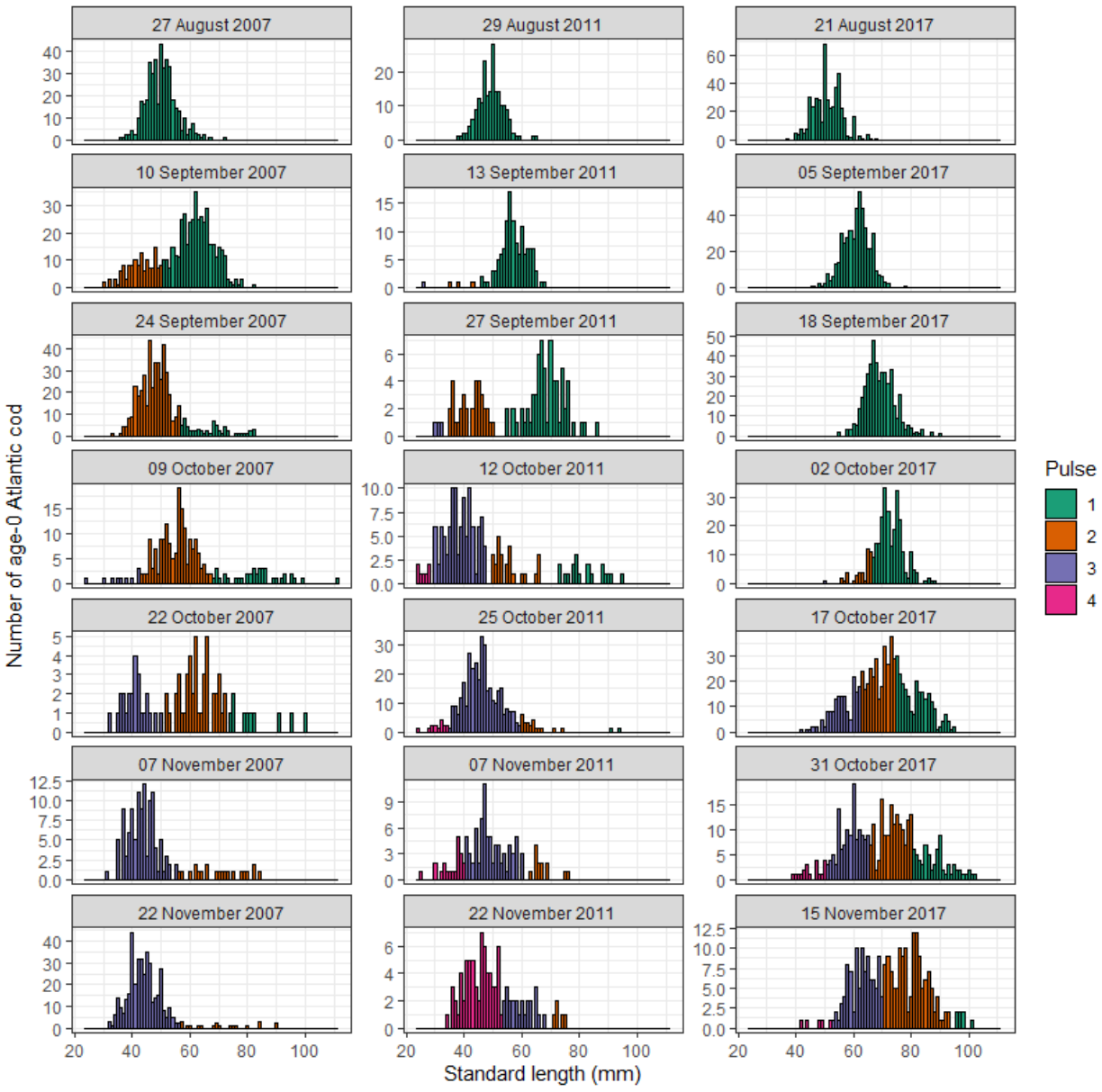

Figure S1. Size frequency histogram of age-0 Atlantic cod (*Gadus morhua*) in coastal Newfoundland, Canada, from August to November in 2007, 2011, and 2017 (Gregory et al. 2017; 2019). Standard length is measured in mm from the snout to the caudal peduncle. Pulse represents the settlement group for each season. Pulse 1 represents the first settlement wave and Pulse 4 represents the fourth settlement wave.

Section S1: Finite Mixture Distribution Models

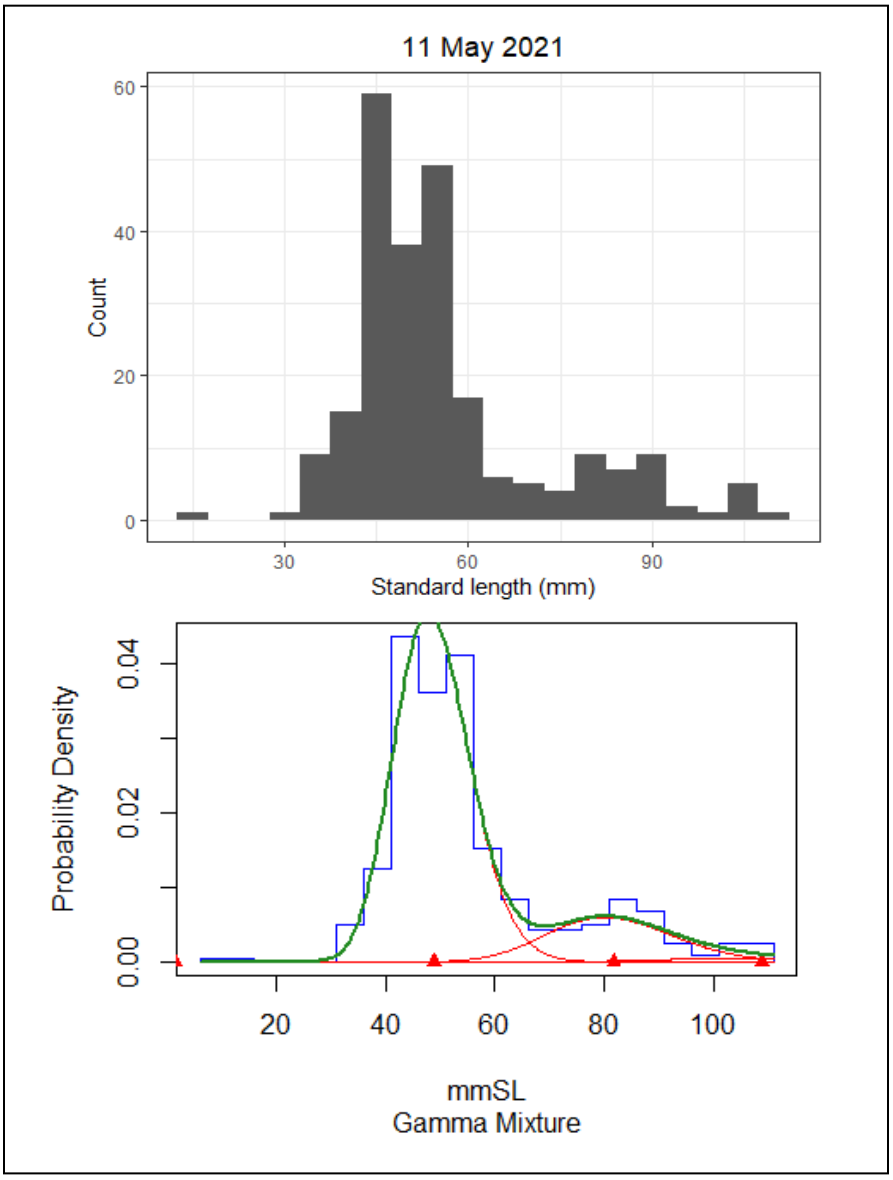

Figure S2. Finite mixture distribution model with Gamma distribution for age-1 Atlantic cod on 11 May 2021 (Gregory et al., unpubl.) using mixdist in R (Macdonald & Du 2018). Size is measured as standard length (mm). The red triangle represents the mean of the distribution, and the red line represents the standard deviation. The green line indicates model fit; curvature closely mirroring the probability density distribution indicates a strong model fit.

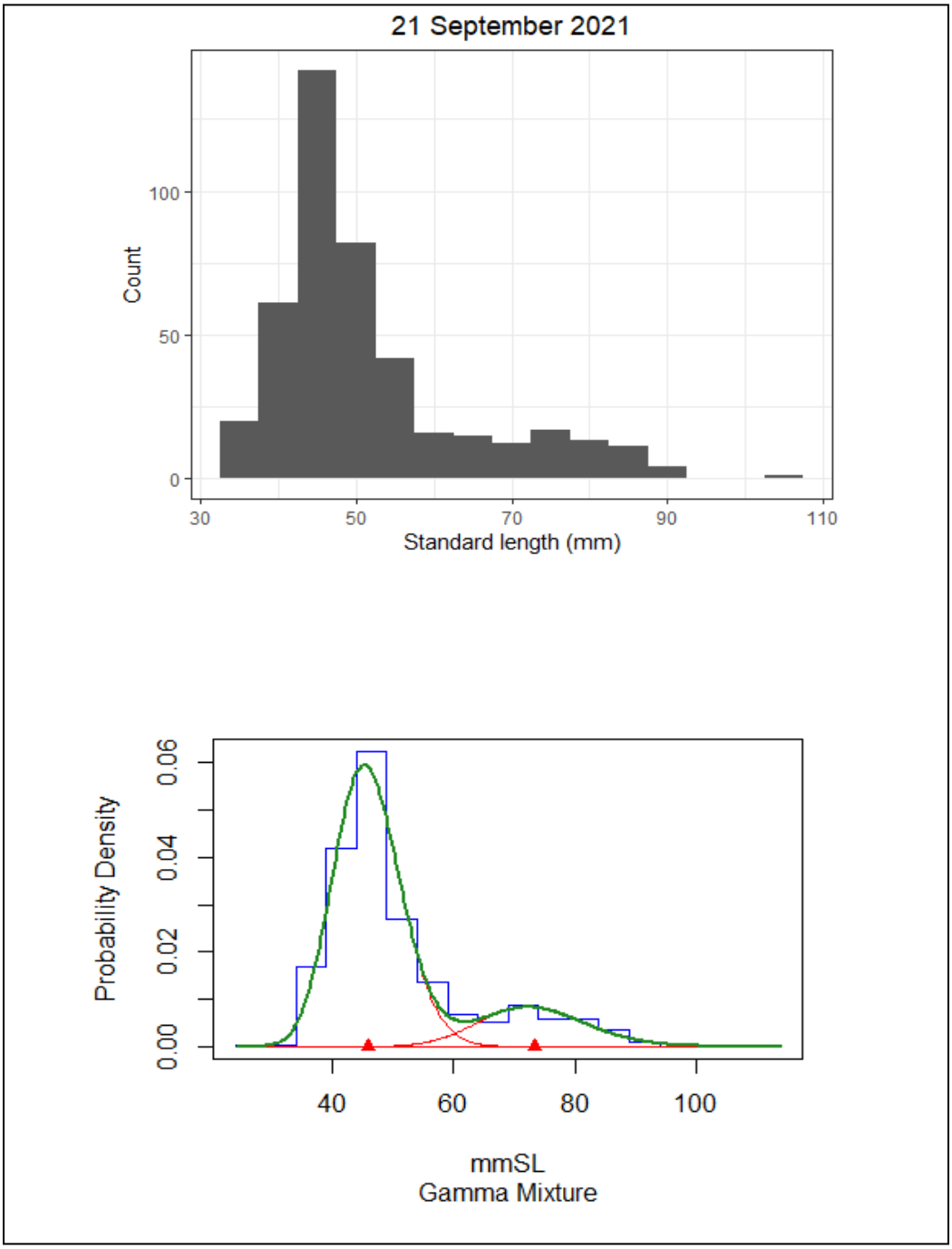

Figure S3. Finite mixture distribution model with Gamma distribution for age-0 Atlantic cod on 21 September 2021 (Gregory et al., unpubl.) using mixdist in R (Macdonald & Du 2018). Size is measured as standard length (mm). The red triangle represents the mean of the distribution, and the red line represents the standard deviation. The green line indicates model fit; curvature closely mirroring the probability density distribution indicates a strong model fit.

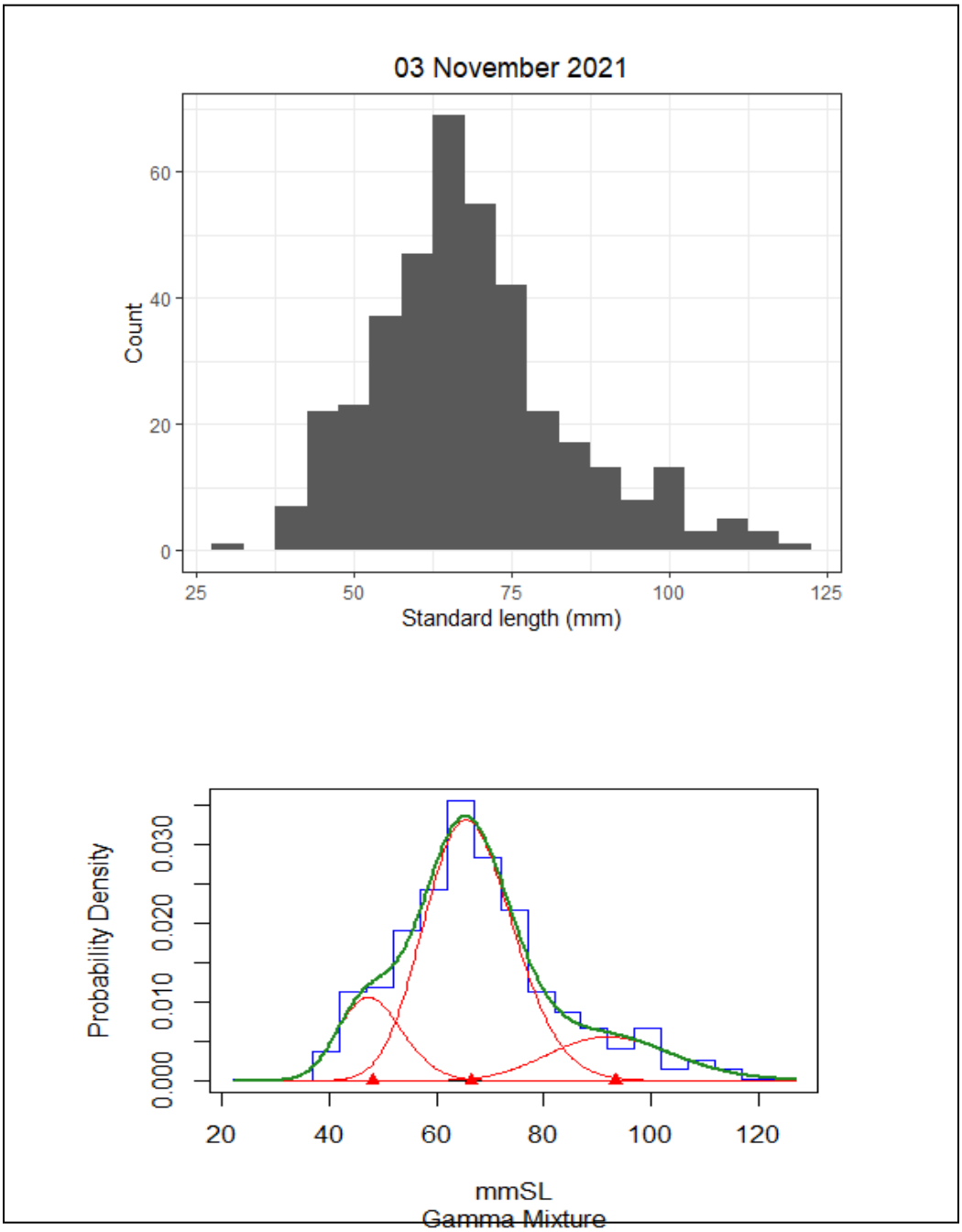

Figure S4. Finite mixture distribution model with Gamma distribution for age-0 Atlantic cod on 3 November 2021 (Gregory et al., unpubl.) using mixdist in R (Macdonald & Du 2018). Size is measured as standard length (mm). The red triangle represents the mean of the distribution, and the red line represents the standard deviation. The green line indicates model fit; curvature closely mirroring the probability density distribution indicates a strong model fit.

Section S2: Estimating Pulses from finite mixture distribution parameters

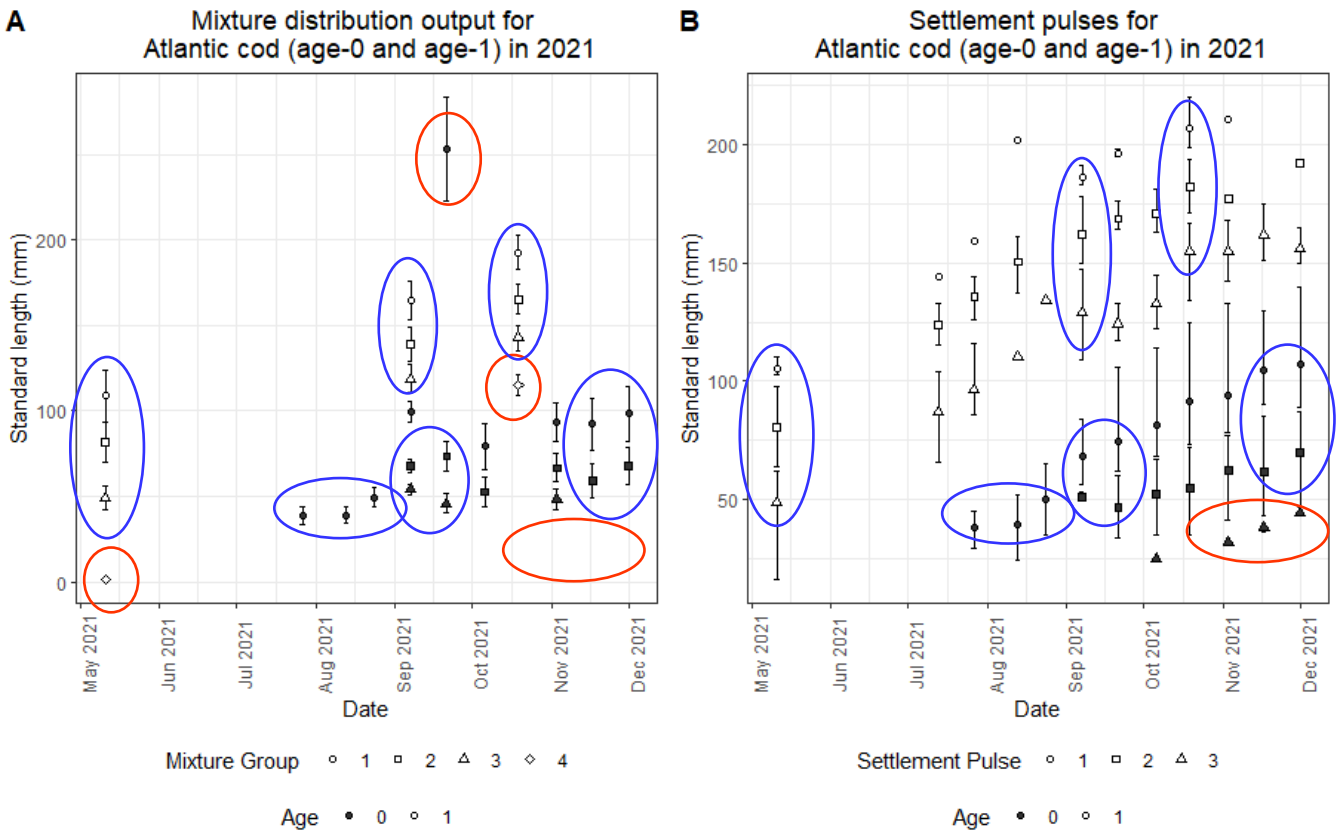

Figure S5. Finite mixture distribution model output (a) and pulse assignments (b) for age-0 (2021 cohort) and age-1 (2020 cohort) Atlantic cod, *Gadus morhua* in Newfoundland Canada from May to December 2021. Open shapes represent age-1 Atlantic cod and closed shapes represent age-0 Atlantic cod. Shape represents the estimated group from (a) mixture distribution models for each sampling interval; error bars for mixture groups are ± 1 standard deviation. Shape represents the settlement pulse (b) for each sampling error; error bars for settlement pulses represent the minimum and maximum length for each pulse. Pulse assignments are derived from finite mixture distribution models. Finite mixture distribution models were not used when there was an insufficient sample size (n<30). Pulses are guided by mixture distribution model results, but are corrected based on growth trajectories within the sampled population when the mathematical solution is outside of biological parameters. Blue circles represent instances where mixture distribution models accurately described settlement pulses, and red circles represent instances where the mixture distribution model required revision based on biological parameters.

Literature Cited

Gregory, R.S., Morris, C., Newton, B., and P. Sargent. 2017. Relative strength of the 2010-12 year classes, from nearshore surveys of demersal age 0 & 1 Atlantic Cod in Newman Sound, Bonavista Bay. Can. Sci. Advis. Sec. Res. Doc. 2016/081.

Gregory RS, Dalley K, Newton B, Sargent PS, and Cooke ELL (2019) Relative strength of three cohorts (2015-17) of Atlantic cod, from nearshore surveys of demersal age 0 and age 1 juveniles in Newman Sound, Bonavista Bay. Can. Sci. Advis. Sec. Res. Doc. 2019/012.

Macdonald P, Du J (2018) Mixdist: finite mixture distribution models. R package version 0.5-5. https://CRAN.R-project.org/package=mixdist.

Macdonald PDM and Pitcher TJ (1979) Age-groups from size frequency data: a versatile and efficent method of analyzing distribution mixtures. J Fish Res Board Can 36:987–10001.

McLachlan G, Peel D (2000) Finite mixture models. John Wiley & Sons, New York. 419 pp.

R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org/