MANUSCRIPT UNDER REVIEW (April 2026)

`fishmax`: An R package for Bayesian estimation of maximum body size in animals

Author names: Freddie Heather*(1), Stephan Munch (2), Asta Audzijonyte (1, 3)

***Corresponding author:** freddie.heather@utas.edu.au (ORCiD: 0000-0002-1650-2617)

Institutions and addresses:

(1) Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Australia

(2) Department of Applied Mathematics, UC Santa Cruz, California, USA

(3) Centre for Marine Socioecology, University of Tasmania, Hobart, Australia

Acknowledgements: We would like to thank Daniel Pauly who triggered this study and suggested that maximum body lengths in fishes should be estimated using statistical methods and extreme value theory. The research presented here was made possible by funding from the Australian Research Council (ARC) Discovery Project (DP220102446) to A.A, and funding from Pew Marine Conservation fellowship associated funding to A.A. and S.M.

Author contributions: F.H, S.M, A.A contributed to the conception and design of the study, analysis and interpretations. F.H. wrote the R code. F.H and A.A drafted the article. F.H, S.M, A.A contributed to the revision of the manuscript.

Funding Declaration: This study was supported by the Australian Research Council (ARC) Discovery Project (DP220102446) to A.A, and a Marine Pew Fellowship in Marine Conservation to A.A.

Data availability statement: There are no new data presented in this study.

Conflict of Interest: The authors declare no conflict of interests.

Keywords: Maximum length, fish, Lmax, EVT, exact-finite sample, extreme value analysis

Abstract

1. Maximum body size is a key trait of an animal species, determining its population dynamics, productivity, and interactions with other species and the environment. For fish in particular, the maximum body length ($\mathrm{L_{max}}$) correlates strongly to other life history parameters and is commonly used in ecological models and population assessments.
2. Maximum size of a species is often defined as a single value, representing the single largest observed individual or an average of several of the largest individuals. This approach, however, does not explicitly account for sampling effort, where greater sampling intensity generally leads to a higher probability of finding a larger individual.
3. Here we present an R package for Bayesian estimation of maximum animal body size and its uncertainty using the largest observed individuals from multiple samples. Although developed for maximum body length, the method could be applied to estimate other extremes (e.g., greatest body weight, tallest height) from other taxa that grow indeterminately.
4. The package uses three alternative methods – extreme value theory and two versions of exact finite sample approach. The latter two are suitable in cases where the underlying population body size distribution can be approximated by a truncated normal distribution.
5. To demonstrate the application of the method we used a hypothetical fishing competition data, but any other records of maximum observed individuals across multiple comparable samples (scientific surveys, trophy hunting, underwater surveys, historical accounts) are also suitable.

1 BACKGROUND

For species that grow indeterminately body size is a strong predictor of many life history parameters including energy requirements, growth rate, predation risk, and trophic position (Blueweiss et al., 1978; Jennings et al., 2001; Peters, 1986; Schmidt-Nielsen, 1984). For fishes, which constitute the majority of vertebrate species, the maximum observed body length (L_{max}) is among the most commonly used life history parameter to inform stock assessment models (e.g., Brey and Pauly, 1986; Hordyk, 2021), or as a correlate for ecological analyses (Peters, 1986). Body size, and specifically body length, is also an easily obtainable metric, which can be estimated or measured using non-destructive methods.

L_{max} is commonly reported as a single species-specific value representing the largest individual ever observed, as compiled in FishBase (Froese and Pauly, 2025) - the largest and main repository of fish related information. Such values are widely used as key traits in data-poor fisheries and cross-species analyses (Östman et al., 2023). However, treating L_{max} as a fixed species-level parameter is problematic because the likelihood of observing extremely large individuals depends strongly on sample size, fishing pressure (Barnett et al., 2017), and temperature-driven variation in growth and longevity (Atkinson, 1994; Brooks et al., 2026). Moreover, estimating a species trait from a single observed individual risks biased inference. Ideally, L_{max} estimates should include uncertainty reflecting the underlying sampling effort, but no formal framework for doing so currently exists.

Assessing uncertainty around extreme observations (here, the largest observed individual) is not unique to body size data. Extreme value theory (EVT) is a statistical framework developed to model rare extremes, such as annual maximum rainfall or temperature (Nerantzaki and Papalexiou, 2022), and is conceptually analogous to the central limit theorem (CLT), but for extremes rather than means (see also Gaines and Denny, 1993 for ecological applications). While the CLT states that sample means converge to a normal distribution regardless of the underlying population, EVT posits that sample maxima follow a known family of distributions, the generalised extreme value (GEV) distribution (Gumbel, 1958). Thus, for repeated samples of a species, the largest individual in each sample ('sample maxima') is expected to follow a GEV distribution. This allows estimation of the probability of observing individuals of a given body size and associated uncertainty. Importantly, the GEV is robust to the form of the underlying population distribution, converging to the Gumbel distribution when that distribution is exponential-like (e.g. lognormal, normal, or exponential).

Heather *et al.* (2026) showed how EVT, using either the generalised extreme value (GEV) or Gumbel distributions, can be applied within a Bayesian framework to formally estimate maximum length ($\mathrm{L}_{\mathrm{max}}$) for fish species or populations. They also introduced an alternative 'exact finite sample' (EFS) method, which incorporates prior knowledge of the underlying body-size distribution from which observations are drawn (see Heather et al., 2025 for justification of why a general shape could be assumed for most fish species). The EFS approach can be based on either a single sample maximum or multiple largest individuals per sample ('exact finite sample multiple maxima' or EFSMM), although using multiple maxima provides little additional gain in estimation precision or accuracy (Heather et al., 2026).

Here we present `fishmax`, an R package that implements methods described in Heather et al. (2026) and provides a user-friendly Bayesian framework for estimating maximum size ($\mathrm{L}_{\mathrm{max}}$) and its uncertainty at the species or population level. We illustrate the workflow using data from a fishing competition to estimate $\mathrm{L}_{\mathrm{max}}$ for a fish population. The method is broadly applicable to any sampling context that yields sample maxima, including scientific trawls, visual censuses, or historical records of the largest fish caught. In the example competition, five fishers (k = 5) each catch multiple fish ($n_1, n_2, \ldots, n_k$) but report only their largest individual (M_k). The `fishmax` package uses these maxima records ($M_1, M_2, \ldots, M_k$) to estimate the probability density function of $\mathrm{L}_{\mathrm{max}}$ using both EVT- and EFS-based approaches.

For the EVT approach, `fishmax` fits either a generalised extreme value (GEV) or Gumbel distribution to the sample maxima, representing the probability of observing individuals of a given size as a function of sampling effort. For example, the 95th percentile corresponds to the expected largest individual from 20 samples ("20-sample $\mathrm{L}_{\mathrm{max}}$") of similar size to those used for model fitting (here, five n_k values). Bayesian posteriors provide credible intervals around these estimates (default: 80%). Users may choose between the GEV or Gumbel distributions (see section *3.2: Choosing the model*). EVT predicts convergence of sample maxima to the GEV or Gumbel as sample size increases, meaning no strict minimum sample size is required; however, in practice, small samples lead to high uncertainty and potentially unreliable parameter estimates (see Heather et al., 2026).

`fishmax` also implements the exact finite sample (EFS) method, which assumes that the underlying body-size distribution can be approximated by a positive-normal distribution, an assumption justified for coastal fish species body lengths (Heather et al., 2025). Under EFS, the mean (μ) and standard deviation (σ) of this distribution are estimated together with a parameter λ describing sample sizes (n_k) that generated the observed maxima. These sample sizes are assumed to follow a Poisson distribution, where λ is the shape parameter. Thus, EFS jointly estimates μ, σ, and λ. Like the EVT approach, EFS yields a "20-sample L_{max}". Full methodological details are provided in Heather *et al.* (2026).

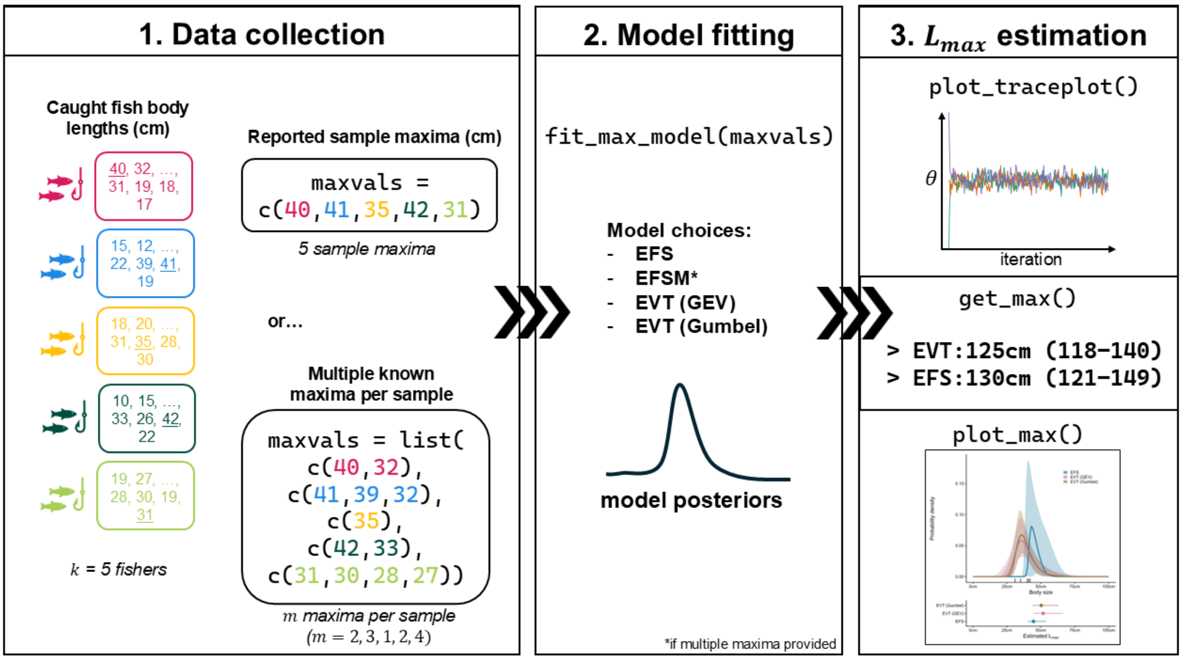

Figure 1. Illustration of the workflow for estimating maximum body size ($\mathrm{L_{max}}$) from a set of sample maxima using the *fishmax* package. Sample maxima are represented by the lengths of the largest fish observed in each of five samples with unknown sample sizes. Users can estimate $L_{\max}$ using either an extreme value theory (EVT) approach, fitting GEV or Gumbel distributions, or the exact finite sample (EFS) method using single or, when available, multiple maxima per sample (EFSMM). After model fitting, users can assess model convergence and estimate the expected maximum body size and associated uncertainty for a specified number of samples (default - 20).

2 IMPLEMENTATION

2.1 Installation

`fishmax` can be run in R (v. 4.1.0 or above; R Core Team, 2025). The `fishmax` package can be installed from the GitHub repository using the `remotes` package (Csárdi et al., 2024). The `fishmax` package depends on the following CRAN packages: `dplyr` (Wickham et al., 2023), `ggplot2` (Wickham, 2016), `tidyr` (Wickham et al., 2025), `purrr` (Wickham and Henry, 2025), `tibble` (Müller and Wickham, 2026), `methods` (R Core Team, 2025), `glue` (Hester and Bryan, 2024), `posterior` (Bürkner et al., 2023), `patchwork` (Pedersen, 2025), `truncnorm` (Mersmann et al., 2023) and `evd` (Stephenson, 2002). Detailed installation instructions can be found at [GitHub link - anonymous until publication].

To estimate maximum body size, `fishmax` uses a Bayesian framework, with parameter estimation performed using CmdStan via the cmdstanr interface (Gabry et al., 2025), and the C++ backend *cmdstan*. Both `cmdstanr` and CmdStan should be installed prior to the installation of `fishmax`. More information regarding the `cmdstanr` package and *cmdstan* can be found at https://mc-stan.org/cmdstanr.

Steps for the installation of `cmdstanr` and CmdStan:

```
install.packages(
  "cmdstanr",
  repos = c("https://stan-dev.r-universe.dev", getOption("repos"))
)
cmdstanr::install_cmdstan()
```

followed by `fishmax` installation:

```
install.packages("remotes")
remotes::install_github("[ANONYMOUS]/fishmax")
```

2.2 Example data

The data provided for the maximum size estimation represent a set of observations of maximum body sizes observed in k samples of a studied population. The unit of measurement does not matter, but in the example below we are using fish lengths in centimetres. For each of the k samples, the dataset has m largest body size observations, body lengths, were m could be just a single largest individual (single maxima approach) or several largest individuals (multiple maxima approach) (left panel in Figure 1). Below we run through an example where only the single largest individual is reported per sample. For example, we have a fishing competition where five fishers report their single largest caught fish, resulting in five values in total: 40, 41, 35, 42, and 31 cm.

```
sample_maxima <- c(40, 41, 35, 42, 31) # single maximum per sample
```

2.3 Functions

The `fishmax` package contains five primary functions: `fit_max_model`, `get_max`, `plot_max`, `max_posterior`, and `plot_traceplot` (Table 1). Detailed descriptions of the functions along with examples are available through the help function using the command (e.g., `?fit_max_model`). A user should start by loading the `fishmax` package:

```
library(fishmax)
```

Table 1. Main functions in the `fishmax` package.

Function	Description
`fit_max_model()`	Takes the sample maxima as its primary argument, either in the form of a vector (single maxima approach) or as a list of vectors (multiple maxima approach). Fits the Bayesian models, selected by the user, to the sample maxima values. The user can select EVT (GEV), EVT (Gumbel) or EFS/EFSMM depending on the dataset (see `model_type` argument; default fits all relevant models). `mod_fit <- fit_max_model(sample_maxima)`
`get_max()`	Takes the fitted model object as its primary argument and outputs the summary statistics of the posterior – the estimate of maximum size and credible intervals. The default is to estimate the maximum size expected from 20 samples (argument `k`) and 80% Bayesian credible intervals (argument `ci`).

	`get_max(fit = mod_fit, ci = 0.8, k = 20)`
`plot_max()`	Takes the fitted model object as its primary argument and outputs a custom plot using the *ggplot2* package creating a publication-ready visual summary of the 20-sample (or otherwise user-defined) maximum size estimate ($\mathrm{L_{max}}$ in the example case). `plot_max(fit = mod_fit)`
`max_posterior()`	Takes the fitted model object as its primary argument and returns Markov Chain Monte Carlo (MCMC) posterior samples, useful for potential further analyses that account for uncertainty. `max_posterior(fit = mod_fit)`
`plot_traceplot()`	Takes the fitted model object as its primary argument and outputs MCMC trace plots to visualise potential convergence issues. `plot_traceplot(fit = mod_fit)`

2.4 Getting L_{max} estimates

The first step in the procedure of estimating L_{max} requires the user to fit the desired model(s) using the `fit_max_model` function:

```
mod_fit <- fit_max_model(sample_maxima)
```

From the fitted model(s) we can extract the estimates of L_{max}, using the default 20-sample L_{max}, and 80% confidence intervals, using:

```
get_max(fit = mod_fit, ci = 0.8, k = 20)
```

The default value of the k argument in the `get_max()` function is 20. This value represents the 20-sample L_{max} that we recommend reporting to maintain consistency between studies (see *Interpretation and Guidance* for more details). Note that this k value in the `get_max()`

function does not refer to the number of samples or sample maxima used for fitting, but the number of hypothetical samples for which to report $\mathrm{L_{max}}$.

In the example dataset presented here, the 20-sample L_{max} based on EFS was slightly lower, at 44.5 cm (40.3 – 53.8 cm 80% credible interval), compared to the estimate from EFS using GEV and Gumbel distributions, at 51.7 cm (44.5 – 66.3 cm) and 50.3 cm (43.7 – 62.9 cm), respectively (Table 2, Figure 2).

Table 2. Estimated 20-sample L_{max} with 80% and 95% credible intervals as provided - output from the `get_max()` function.

Model	L_{max} estimate (cm)	80% CI (cm)	95% CI (cm)
EFS	44.5	40.3 – 53.8	39.1 – 61.8
EVT (GEV)	51.7	44.5 – 66.3	42.0 – 77.4
EVT (Gumbel)	50.3	43.7 – 62.9	41.2 – 75.4

The posterior samples of the parameter estimates and posterior samples of L_{max} can be extracted using the `max_posterior(mod_fit)`. These posterior samples might be useful in the case of propagating error in the estimates of L_{max} in further analysis such as the effect of temperature or fishing pressure on L_{max}.

2.5 Convergence and visualisation

It is important to check that the models have converged properly before interpreting the results. The first step in checking model convergence is to look at the summary statistics of the model fits in the fitted model object (`mod_fit` in the example above), specifically the diagnostic statistics of *rhat* reported for each of the estimated parameters. The general advice is to use four chains and to only trust the fitting if *rhat* is 1.01 or lower. *rhat* values above 1.05 may be problematic (Gabry et al., 2025).

The user can modify the number of MCMC chains (default = 4 chains) for model fitting using the `chains` argument inside `fit_max_model(sample_maxima, chains = 4)`. If the models are not converting, the number of iterations in the MCMC process should be increased (default = 2000 warmup iterations and 1000 sampling iterations). The MCMC traceplots can be visualised for each of the estimated parameters using `plot_traceplot()` function:

```
plot_traceplot(fit) # visualise all models
plot_traceplot(fit["efs"]) # visualise a specific model
```

where '`efs`' to '`evt`', or '`evtg`' refers to the estimates from exact finite sample, extreme value theory GEV distribution or extreme value theory Gumbel distribution, respectively.

The estimate of L_{max} can be visualised using `plot_max(mod_fit)` function (Figure 2). This function takes the fitted model (or models) as its primary argument, i.e., the output from the `fit_max_model()` function. For EFS methods the `plot_max()` function visualises the PDF of the 20-sample (or other user-defined k value) $\mathrm{L_{max}}$ posterior – where the mean of this distribution represents the expected 20-sample L_{max} (dashed pink vertical line in Figure 2). For the EVT methods, the `plot_max()` function visualises the PDF of the sample maxima – where the mean of the distribution represents the expected sample maxima (i.e., equivalent to a "one-sample L_{max}"). To estimate the 20-sample (or other user-defined k value) L_{max}, one would extract the 95% percentile of this distribution. The given percentile is defined by the user-defined k value as $1 - (1/k)$.

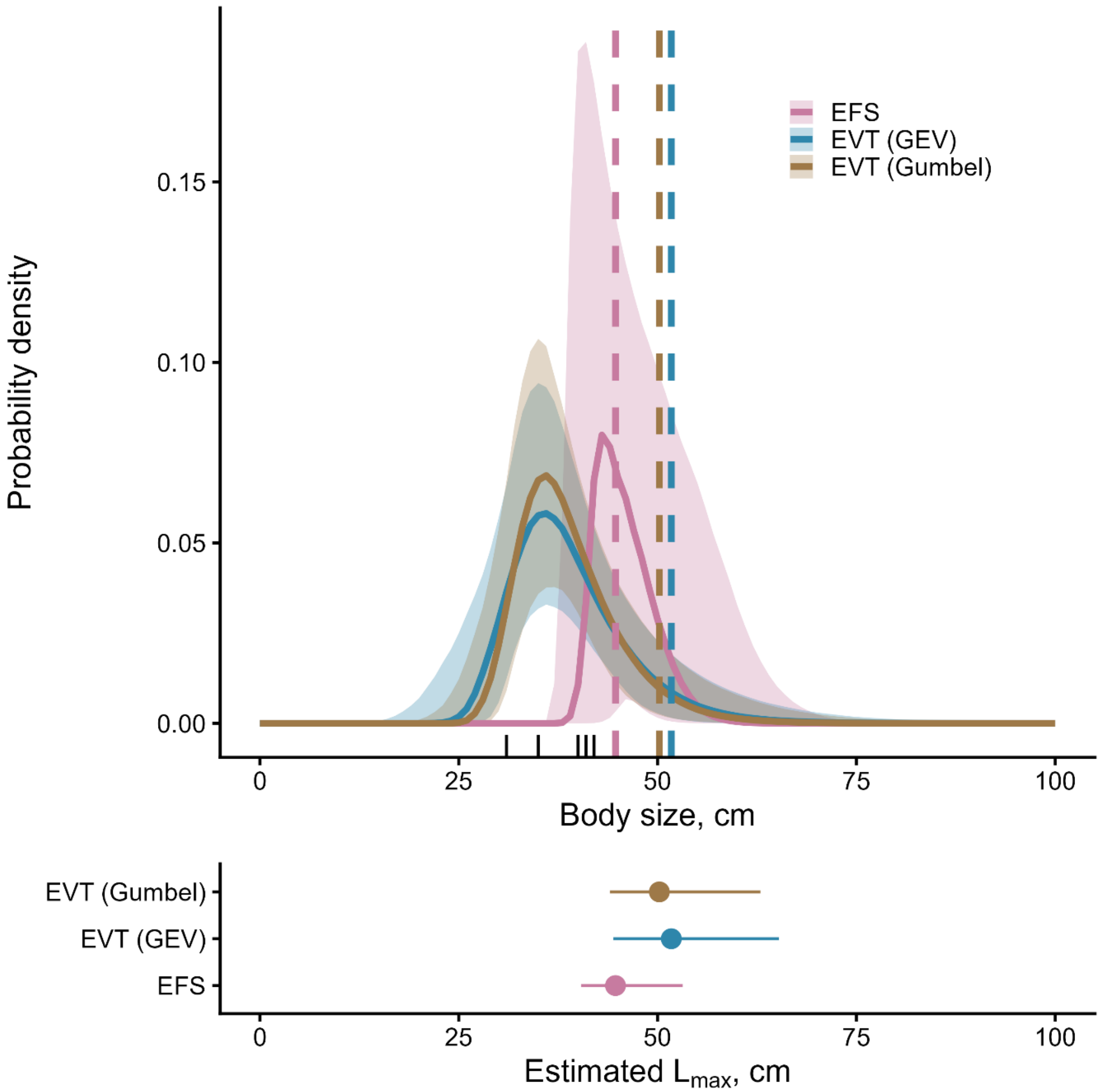

Figure 2. Estimated 20-sample L_{max} based on five sample maxima, shown using output from the `plot_max()` function. The upper panel displays the probability density functions of the sample maxima from two EVT models (where the 20-sample L_{max} corresponds to the 95th percentile) and the probability density function of the 20-sample L_{max} from the EFS method (where the estimate corresponds to the median). The lower panel compares the estimated 20-sample L_{max} across the three approaches, with associated 80% credible intervals.

2.6 Multiple maxima per sample

In cases where users have information on several largest individuals per sample (left panel in Figure 1), maximum body size can be estimated using the exact finite sample multiple maxima (EFSMM) method. For the fishing competition example case, the data available may look like the list below:

```
# if more than the single maximum is known…
maxima_list <- list(c(40, 39), 41, c(33, 34, 35), c(42, 40, 39), 31)
```

Note that if the input data is provided as a list of vectors, where at least one element of the list is of length greater than 1 (i.e., m >1), `fishmax` by default will fit the EFSMM model as well as the EFS (single maxima – only the maxima are considered, other values ignored). Note that the EVT methods cannot utilise more than the single maximum per sample, therefore the second, third (etc.) maximum per sample are ignored.

```
fit_multiple <- fit_max_model(maxima_list) # multiple maxima
```

All other functions that take the fitted model object are the same and use `fit_multiple` object as an argument (e.g., `get_max(fit_multiple)`).

3 INTERPRETATION AND GUIDANCE

To our knowledge, EVT has been formally used to estimate uncertainty in fish maximum length (L_{max}) in only two published cases. Formacion (1991) applied the Gumbel distribution to sample maxima from purse-seine and trawl surveys to estimate L_{max} of short mackerel (*Rastrelliger brachysoma*) in the central Philippines, an approach later incorporated into the FiSAT II stock-assessment software (Gayanilo Jr et al., 1996). Only one subsequent study appears to have used this feature, applying it to four samples to estimate the maximum length of the invasive toadfish (*Lagocephalus sceleratus*) using media-reported trophy catches (Ulman et al., 2022). However, FiSAT is now rarely used, outdated, and relies on distributional approximations developed prior to modern computing. We hope that `fishmax` will help enable a new generation of statistically robust maximum body-size estimates, allowing more rigorous comparisons of changes driven by human and environmental pressures.

3.1 Why the 20-sample L_{max}?

A natural question is why `fishmax` uses a 20-sample L_{max} by default. Under both EFS (assuming positive-normal underlying body size distribution) and EVT (GEV or Gumbel) frameworks, the right tail of the fitted distribution is unbounded, implying no theoretical upper limit to observable body size – with sufficient sampling, increasingly larger individuals can always be observed. While finite populations do have a true maximum size, this value is rarely observable in practice unless every individual is measured. Instead, these methods estimate the probability of observing a given body size conditional on sampling effort. The default 20-sample L_{max} therefore represents the expected largest individual from 20 samples of similar size to those used to fit the model. For EVT, the exact number of individuals per sample is not required, provided sample sizes are broadly comparable; for EFS, sample size is estimated directly. Both methods assume approximately similar sample sizes across samples. Sensitivity analyses in Heather et al. (2026) show improved EVT performance with larger within-sample sizes (e.g. ≥100 individuals) given sufficient replication (e.g. ≥3 samples).

3.2 Choosing the model - EFS, EVT (GEV) or EVT (Gumbel)

A related question concerns which method should be used to estimate L_{max}. This issue is explored in detail by Heather et al. (2026), with key assumptions summarised in Table 3. As a general default for body-size data, we recommend the EVT approach with the GEV distribution, as it is the most flexible and least assumption-dependent, although it typically yields the widest uncertainty intervals. The Gumbel distribution is a special case of the GEV family and is appropriate when the underlying size distribution has an exponential tail (e.g. normal, lognormal, or exponential). If the underlying distribution can instead be approximated by a positive-normal distribution, as shown for hundreds of fish species (Heather et al., 2025), the EFS method provides more precise estimates. In our example, the EFS method produced slightly smaller 20-sample L_{max} estimates and narrower uncertainty intervals than either EVT approach.

Table 3. Assumptions about the shape of the underlying distribution of each of the three models, in increasing order of assumptions: extreme value theory using the generalised extreme value distribution (EVT – GEV), extreme value theory using the Gumbel distribution (EVT- Gumbel), and the exact finite sample (EFS).

Model	Notation	Underlying body length distribution shape assumptions
EVT - GEV	`evt`	No assumptions about the shape of the underlying body size distribution.

EVT- Gumbel	`evtg`	Assumes the underlying distribution follows an exponential-like tail (e.g., normal, lognormal, exponential), which means that abundance quickly declines with increasing body length.
EFS	`efs` or `efsmm`	Assumes the underlying distribution follows a positive-normal shape, and the sampling process to generate sample sizes follows a Poisson distribution.

References

Atkinson, D., 1994. Temperature and organism size-a biological law for ectotherms? Adv. Ecol. Res. 25, 1–58.

Barnett, L.A.K., Branch, T.A., Ranasinghe, R.A., Essington, T.E., 2017. Old-Growth Fishes Become Scarce under Fishing. Current Biology 27, 2843-2848.e2. https://doi.org/10.1016/j.cub.2017.07.069

Blueweiss, L., Fox, H., Kudzma, V., Nakashima, D., Peters, R., Sams, S., 1978. Relationships between body size and some life history parameters. Oecologia 37, 257–272. https://doi.org/10.1007/bf00344996

Brey, T., Pauly, D., 1986. Electronic length frequency analysis: a revised and expanded user’s guide to ELEFAN 0, 1 and 2.

Brooks, G.C., Frater, P.N., Jensen, O.P., Hansen, G.J.A., Paukert, C., Verhoeven, M., Wszola, L., Xu, L., Feiner, Z.S., 2026. Mixed Support for the Temperature-Size Rule in Wild Freshwater Fishes. Ecology Letters 29, e70344. https://doi.org/10.1111/ele.70344

Bürkner, P.-C., Gabry, J., Kay, M., Vehtari, A., 2023. posterior: Tools for working with posterior distributions. R package version 1.

Csárdi, G., Hester, J., Wickham, H., Chang, W., Morgan, M., Tenenbaum, D., 2024. remotes: R Package Installation from Remote Repositories, Including “GitHub.” https://doi.org/10.32614/CRAN.package.remotes

Formacion, S.P., 1991. Extreme Value Theory Applied to the Statistical Distribution of the Largest Lengths of Fish. AFS 4. https://doi.org/10.33997/j.afs.1991.4.2.001

Froese, R., Pauly, D., 2025. FishBase. World Wide Web electronic publication. URL www.fishbase.org

Gabry, J., Češnovar, R., Johnson, A., Bronder, S., 2025. cmdstanr: R Interface to “CmdStan.”

Gaines, S.D., Denny, M.W., 1993. The Largest, Smallest, Highest, Lowest, Longest, and Shortest: Extremes in Ecology. Ecology 74, 1677–1692. https://doi.org/10.2307/1939926

Gayanilo Jr, F., Sparre, P., Pauly, D., 1996. The FAO-ICLARM stock assessment tools (FiSAT) user’s guide. FAO computerized information series (Fisheries) 8, 1–126.

Gumbel, E., 1958. Statistics of Extremes. Columbia University Press.

Heather, F., Munch, S., Audzijonyte, A., 2026. Beyond the single biggest fish: robust inference of species maximum length (Lmax) from sample maxima. Rev Fish Biol Fisheries 36, 46. https://doi.org/10.1007/s11160-026-10047-2

Heather, F., Richards, S., Krueck, N., Stuart-Smith, R., Brandl, S., Casey, J., Edgar, G., Barrett, N., Parravicini, V., Audzijonyte, A., 2025. Consistent Unimodal Body Length

Distributions in Hundreds of Reef Fishes Across Diverse Life Histories. Fish and Fisheries 26, 533–545.

Hester, J., Bryan, J., 2024. glue: Interpreted String Literals. https://doi.org/10.32614/CRAN.package.glue

Hordyk, A., 2021. LBSPR: Length-Based Spawning Potential Ratio.

Jennings, S., Pinnegar, J.K., Polunin, N.V.C., Boon, T.W., 2001. Weak cross-species relationships between body size and trophic level belie powerful size-based trophic structuring in fish communities 11.

Mersmann, O., Trautmann, H., Steuer, D., Bornkamp, B., 2023. truncnorm: Truncated Normal Distribution. https://doi.org/10.32614/CRAN.package.truncnorm

Müller, K., Wickham, H., 2026. tibble: Simple Data Frames. https://doi.org/10.32614/CRAN.package.tibble

Nerantzaki, S.D., Papalexiou, S.M., 2022. Assessing extremes in hydroclimatology: A review on probabilistic methods. Journal of Hydrology 605, 127302. https://doi.org/10.1016/j.jhydrol.2021.127302

Östman, Ö., Hommik, K., Bolund, E., Heikinheimo, O., Olin, M., Lejk, A.M., Svirgsden, R., Smoliński, S., Olsson, J., 2023. Size-based indicators for assessments of ecological status of coastal fish communities. ICES Journal of Marine Science 80, 2478–2489. https://doi.org/10.1093/icesjms/fsad158

Pedersen, T.L., 2025. patchwork: The Composer of Plots. https://doi.org/10.32614/CRAN.package.patchwork

Peters, R.H., 1986. The ecological implications of body size. Cambridge university press.

R Core Team, 2025. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Schmidt-Nielsen, K., 1984. Scaling: why is animal size so important? Cambridge university press.

Stephenson, A.G., 2002. evd: Extreme Value Distributions. R News 2, 31–32.

Ulman, A., Kalogirou, S., Pauly, D., 2022. The Dynamics of Maximum Lengths for the Invasive Silver-Cheeked Toadfish (Lagocephalus sceleratus) in the Eastern Mediterranean Sea. JMSE 10, 387. https://doi.org/10.3390/jmse10030387

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham, H., François, R., Henry, L., Müller, K., Vaughan, D., 2023. dplyr: A Grammar of Data Manipulation. https://doi.org/10.32614/CRAN.package.dplyr

Wickham, H., Henry, L., 2025. purrr: Functional Programming Tools. https://doi.org/10.32614/CRAN.package.purrr

Wickham, H., Vaughan, D., Girlich, M., 2025. tidyr: Tidy Messy Data. https://doi.org/10.32614/CRAN.package.tidyr