ssdtools: An R package to fit Species Sensitivity Distributions

Joe Thorley[1] **and Carl Schwarz**[2]

1 Poisson Consulting Ltd., Nelson, Canada **2** Simon Fraser University, Vancouver, Canada

Summary

Species sensitivity distributions (SSDs) are cumulative probability distributions that are used to estimate the percent of species that are affected by a given concentration of a chemical. The concentration that affects 5% of the species is referred to as the 5% Hazard Concentration (HC). Hazard concentrations are used for the derivation of environmental quality criteria and ecological risk assessment for contaminated ecosystems (Posthuma, Suter II, & Traas, 2001).

ssdtools is an R package (R Core Team, 2018) to fit SSDs using Maximum Likelihood (Millar, 2011). The available distributions include the log-normal (lnorm), log-logistic (llog), Gompertz, log-Gumbel (lgumbel), gamma and Weibull distributions. Model selection or model averaging can be performed using Akaike Information Criterion (AIC), AIC corrected for small sample size (AICc) and Bayesian Information Criterion (BIC) (Burnham & Anderson, 2002). Confidence intervals can be calculated for the cumulative distribution or specific hazard concentrations (percentiles) using parametric bootstrap resampling.

ssdtools uses the fitdistrplus R package (Delignette-Muller & Dutang, 2015) for model fitting, calculation of information criteria and bootstrapping. ssdtools extends the ggplot2 R package (Wickham, 2016) by defining ssdfit, xribbon and hcintersect geometries to allow the user to produce custom SSD plots.

The following code fits the default distributions to the data set for boron and plots the results with the model averaged 5% HC as a dotted intersection (Figure 1).

```r
# install.packages("ssdtools")
library(ssdtools)
library(ggplot2)

dists <- ssd_fit_dists(boron_data)
hc <- ssd_hc(dists)

autoplot(dists) +
  geom_hcintersect(data = hc, aes(xintercept = est, yintercept = percent/100))
```

Acknowledgements
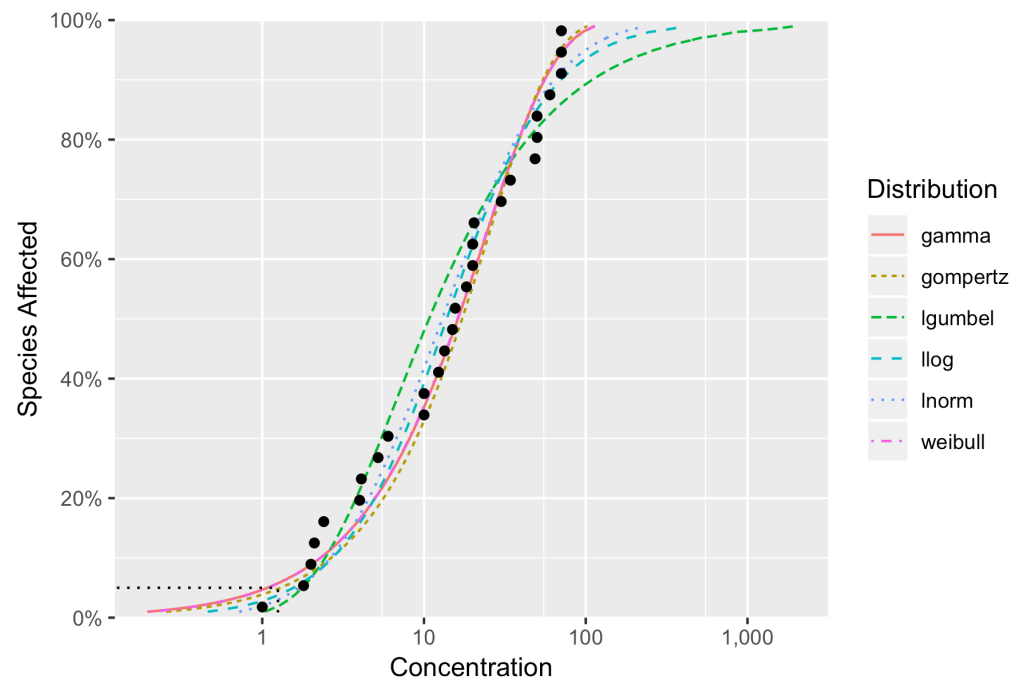
Figure 1: Species sensitivity distributions for sample species concentration values

References

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* New York, NY: Springer New York. doi:10.1007/b97636

Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software*, *64*(4), 1–34. doi:10.18637/jss.v064.i04

Millar, R. B. (2011). *Maximum likelihood estimation and inference: With examples in R, SAS, and ADMB.* Statistics in practice. Chichester, West Sussex: Wiley. doi:10.1002/9780470094846

Posthuma, L., Suter II, G. W., & Traas, T. P. (2001). *Species sensitivity distributions in ecotoxicology.* CRC press. Retrieved from https://www.crcpress.com/Species-Sensitivity-Distributions-in-Ecotoxicology/Posthuma-II-Traas/p/book/9781566705783

R Core Team. (2018). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York. doi:10.1007/978-0-387-98141-3