

dia: An R package for the National Oceanic and Atmospheric Administration dam impact analysis

Daniel S. Stich¹✉, Julie L. Nieland², and Timothy F. Sheehan²

¹ Biology Department and Biological Field Station, State University of New York at Oneonta, NY 13280 USA ² National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA 02543 USA ✉ Corresponding author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- Review ✉
- Repository ✉
- Archive ✉

Editor: [Open Journals](#) ✉

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

Populations of anadromous (sea-run) fishes such as Atlantic salmon *Salmo salar* have experienced severe global declines as a result of pollution, overfishing, and construction of dams (Limburg & Waldman, 2009). Life-history-based simulation models are commonly used for planning and implementing fisheries recovery activities for many diadromous species (Barber et al., 2018; e.g., Nieland et al., 2013; Stich et al., 2019; Zydlewski et al., 2021). Prior to the expansion of open-source development, many institutionalized decision-support tools utilizing such models relied on closed-source or paid software. For example, the National Oceanic and Atmospheric Administration (NOAA) Dam Impact Analysis (DIA) was originally created as a stochastic life cycle model for Atlantic salmon in the @RISK add-in within Microsoft Excel (Nieland et al., 2013). This tool differs from those hosted in similar R packages for non-salmonid species (e.g., Stich et al., 2019; Zydlewski et al., 2021) in that it employs homing and straying probabilities to natal streams for adult spawning and integrates results of physical modeling to inform population dynamics (Nieland et al., 2013; Nieland & Sheehan, 2020). This class of tools, in general, provides advantages for decision making related to anadromous species because it allows integration of geographically and temporally explicit stock dynamics (e.g., influences of dams) that are not readily implemented in classical fisheries stock assessment tools such as those available in existing R packages (Erickson et al., 2022; e.g., Kell et al., 2007; Ogle et al., 2022). We created the dia package (Stich et al., 2021) for the R programming language (R Core Team, 2024) as a freely accessible, open-source implementation of these tools that will promote transparency in planning and decision making and improve the availability and quality of these tools over time.

Statement of need

dia is an R-based implementation of a previously closed-source life cycle model of Atlantic salmon population dynamics that is used to understand the sensitivity of species recovery trajectories to natural and anthropogenic marine and freshwater influences. It was developed to assess the sensitivity of restoration outcomes to uncertainty in life-history inputs alongside the impacts of dams and restoration activities. The DIA model uses empirical life-history estimates (e.g., survival), predictive flow and resulting flow-specific dam survival modeling, and other empirical data in freshwater and marine environments to simulate consecutive generations of Atlantic salmon in the Penobscot River (Maine, USA) under varying environmental conditions or management decisions. As the largest remaining population of this critically endangered species in the USA, the population is intensively managed. Management decisions include fish passage rates at dams, fishery harvest rates, and numbers and locations for hatchery stocking of fish (Nieland et al., 2013). Since development, it has been used for mechanistic exploration of key life-history uncertainties within the context of species recovery (Nieland et al., 2015).

43 and to support decision-making at federally regulated hydropower dams on the Penobscot
44 River (e.g., [National Marine Fisheries Service, 2013](#); [Nieland & Sheehan, 2020](#)).

45 We created dia for use by fisheries researchers, managers, and practitioners interested in
46 understanding population dynamics of intensively managed endangered Atlantic salmon in the
47 USA. The R package maintains the core routines from the original closed-source version of
48 the model by replicating spreadsheet-based calculations, and incorporates original data and
49 parameter sets as built-in objects that serve as default values for arguments of the primary
50 user-facing functions. However, it also allows exploration of future restoration scenarios through
51 a variety of user-facing options and allows exploration of uncertainty associated with those
52 parameters.

53 The two primary user-facing functions within the dia package are the `run_dia()` and
54 `run_dia_shiny()`, which provide redundant interfaces for using Dam Impact Analysis models
55 in different ways. The `run_dia()` function provides an extensible interface to DIA that can
56 be used for long-run simulation or decision-optimization studies, and allows incorporation
57 of user-specified data sets such as flow-correlated survival probabilities at dams and in
58 free-flowing river reaches, marine survival and other life-history inputs, or fish-stocking data.
59 The `run_dia_shiny()` function deploys a graphical user interface using the shiny package
60 ([Chang et al. \(2022\)](#)) that is less extensible but more easily used by fishery managers and
61 practitioners who may be less familiar with programming. `run_dia_shiny()` also includes
62 exportable results from simulation models including .csv or other flat-file formats and default
63 plots through the ggplot2 R package [[Wickham \(2016\)](#); [Wickham et al. \(2019\)](#); Figure 1].
64 Both can be deployed on networked servers as any other R or shiny application to improve
65 accessibility or facilitate use on high performance computers for large simulations. The GitHub
66 repository ([Stich et al., 2021](#)) includes additional instructions for installation and a variety
67 of potential uses of `run_dia()` and `run_dia_shiny()`, with shorter examples included in the
68 package help files. While implementation is currently limited to the Penobscot River as a
69 priority conservation water in the USA, the package serves as one example to help generalize
70 these modeling approaches to Atlantic salmon and other sea-run fish in watersheds globally.
71 Specifically, while many of the built-in datasets and helper functions in dia are generalized or
72 could be used to simulate life-history information for other systems and species, generalizing
73 the geographic component (i.e., structural river system) represents an important priority for
74 future development.

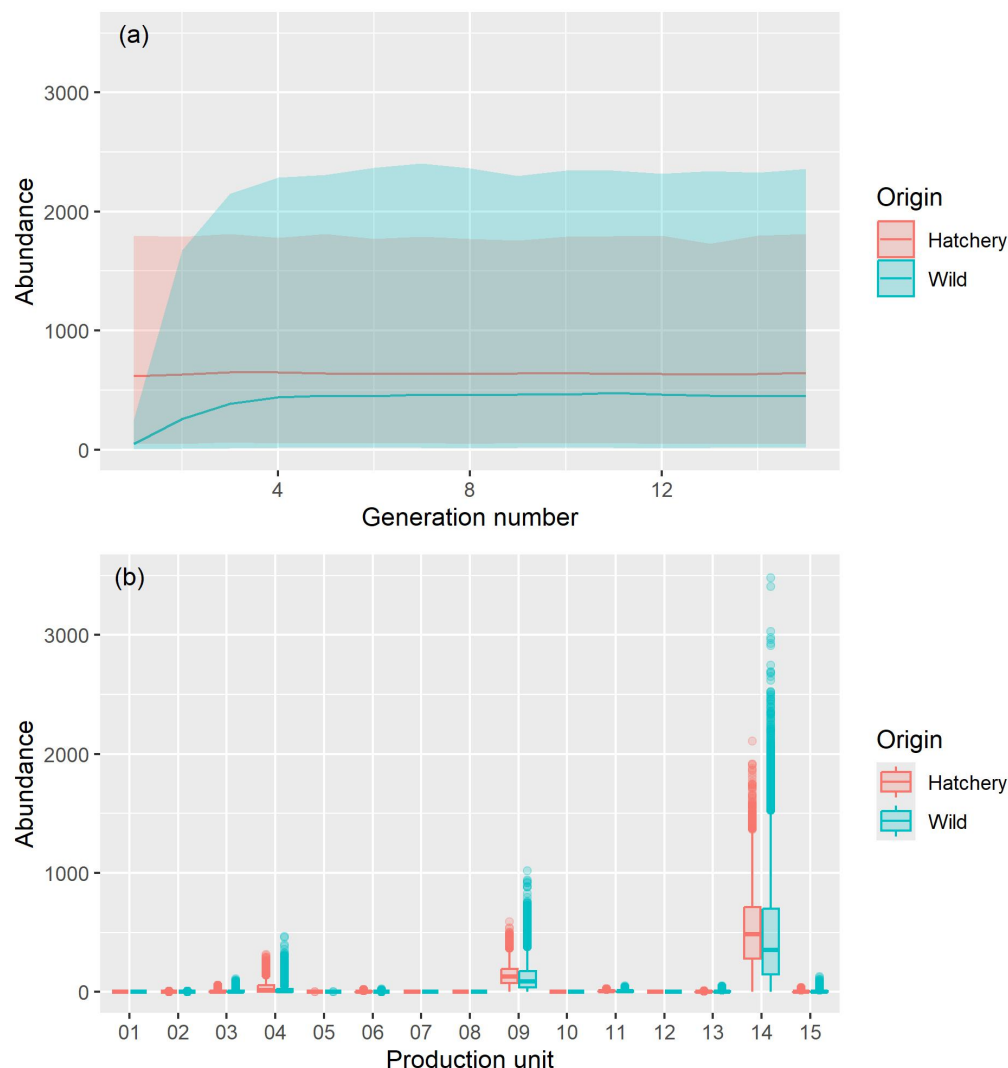


Figure 1: Example graphical outputs using the default argument values (Nieland & Sheehan, 2020) for `run_dia()` to run 10,000 simulations, showing (a) the number of two-sea-winter adult females of hatchery or wild origin returning to the watershed during each generation to spawn, and (b) the number of those fish returning to each production unit within the watershed after 15 generations.

Acknowledgements

We acknowledge contributions from Rory Saunders, Tara Trinko-Lake, Jeffrey Murphy, and Justin Stevens in the original development of the NOAA Dam Impact Analysis model. This work was funded by the NOAA National Marine Fisheries Service Northeast Fisheries Science Center, with additional support from the SUNY Oneonta Biological Field Station. We thank the reviewers for helpful guidance and comments that greatly improved the quality of this paper and the open-source software package. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Barber, B. L., Gibson, A. J., O'Malley, A. J., & Zydlewski, J. (2018). Does what goes up also come down? Using a recruitment model to balance alewife nutrient import and export.

- 86 *Marine and Coastal Fisheries*, 10(2), 236–254. <https://doi.org/10.1002/mcf2.10021>
- 87 Chang, W., Cheng, J., Allaire, J. J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson,
88 J., Dipert, A., & Borges, B. (2022). *Shiny: Web application framework for R*. <https://doi.org/10.32614/CRAN.package.shiny>
- 89
- 90 Erickson, R. A., Stich, D. S., & Hebert, J. L. (2022). fishStan: Hierarchical bayesian models
91 for fisheries. *Journal of Open Source Software*, 7(71), 3444. [https://doi.org/10.21105/](https://doi.org/10.21105/joss.03444)
92 [joss.03444](https://doi.org/10.21105/joss.03444)
- 93 Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., Jardim,
94 E., Mardle, S., Pastoors, M. A., Poos, J. J., Scott, F., & Scott, R. D. (2007). FLR: An
95 open-source framework for the evaluation and development of management strategies.
96 *ICES Journal of Marine Science*, 64, 640–646.
- 97 Limburg, K. E., & Waldman, J. R. (2009). Dramatic declines in North Atlantic diadromous
98 fishes. *BioScience*, 59(11), 955–965. <https://doi.org/10.1525/bio.2009.59.11.7>
- 99 National Marine Fisheries Service. (2013). *National Marine Fisheries Service Endan-*
100 *gered Species Act Biological Opinion Amendment of License for the Mattaceunk*
101 *Project f/NER/2013/9640*. National Oceanic and Atmospheric Administration.
102 <https://repository.library.noaa.gov/view/noaa/55737>
- 103 Nieland, J. L., & Sheehan, T. F. (2020). Quantifying the effects of dams on Atlantic salmon in
104 the Penobscot River watershed, with a focus on Weldon Dam. In *Northeast Fisheries Science*
105 *Center Reference Document 19-16*. National Oceanic and Atmospheric Administration.
106 <https://doi.org/10.25923/v67x-kk62>
- 107 Nieland, J. L., Sheehan, T. F., & Saunders, R. (2015). Assessing demographic effects of
108 dams on diadromous fish: a case study for Atlantic salmon in the Penobscot River, Maine.
109 *International Council for the Exploration of the Sea Journal of Marine Science*, 72(8),
110 2423–2437. <https://doi.org/10.1093/icesjms/fsv083>
- 111 Nieland, J. L., Sheehan, T. F., Saunders, R., Murphy, J. S., Trinko Lake, T., & Stevens, J. R.
112 (2013). Dam impact analysis model for Atlantic salmon in the Penobscot River, Maine.
113 In *Northeast Fisheries Science Center Reference Document 13-09*. National Oceanic;
114 Atmospheric Administration. <https://repository.library.noaa.gov/view/noaa/4559>
- 115 Ogle, D. H., Doll, J. C., Wheeler, P., & Dinno, A. (2022). *FSA: Fisheries stock analysis*.
116 <https://github.com/fishR-Core-Team/FSA>
- 117 R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation
118 for Statistical Computing. <https://www.R-project.org/>
- 119 Stich, D. S., Nieland, J. L., & Sheehan, T. F. (2021). dia: Atlantic salmon dam impact analysis
120 (DIA) for R. In *GitHub repository*. GitHub. <https://doi.org/10.5281/zenodo.13376045>
- 121 Stich, D. S., Sheehan, T. F., & Zydlewski, J. D. (2019). A dam passage performance standard
122 model for American shad. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(5),
123 762–779. <https://doi.org/10.1139/cjfas-2018-0008>
- 124 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.
125 <https://doi.org/10.32614/CRAN.package.ggplot2>
- 126 Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R.,
127 Grolemon, G., Haye, A., Henr, L., Heste, J., Kuh, M., Lin Pederse, T., Mille, E., Bach, S.
128 M., Müll, K., Jeroen Oo, D., Seid, D. P., V., ... Yutani, H. (2019). Welcome to the tidyverse.
129 *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- 130 Zydlewski, J., Stich, D. S., Roy, S., Bailey, M., Sheehan, T., & Sprankle, K. (2021). What
131 have we lost? Modeling dam impacts on American shad populations through their native
132 range. *Frontiers in Marine Science*, 8. <https://doi.org/10.3389/fmars.2021.734213>