

Environmental Data from the common garden experiments run as part of CviMVP project from 2023-2025

Website: <https://www.bco-dmo.org/dataset/1007572>

Data Type: Other Field Results, experimental

Version: 1

Version Date: 2026-09-10

Project

» [CAREER: Evaluation of machine learning algorithms for understanding and predicting adaptation to multivariate environments with a Model Validation Program \(MVP\)](#) (Model Validation Program)

Contributors	Affiliation	Role
Lotterhos, Katie	Northeastern University	Principal Investigator
Small, Jessica	The College of William & Mary	Co-Principal Investigator
Carnegie, Ryan	Virginia Institute of Marine Science (VIMS)	Scientist
Eppley, Madeline	Northeastern University	Scientist
Rumberger, Camille	Northeastern University	Scientist
Bajaj, Kiran	Northeastern University	Student
Mongillo, Nicole	Northeastern University	Student
Segnitz, Zea	Northeastern University	Student
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Table of Contents

- [Dataset Description](#)
 - [BCO-DMO Curation Notes](#)
- [Related Datasets](#)
- [Parameters](#)
- [Project Information](#)
- [Funding](#)

Dataset Description

Eastern oyster, *Crassostrea virginica*, urn:lsid:marinespecies.org:taxname:140657

BCO-DMO Curation Notes

Curation Actions Performed on Data

* None - the data submitter is revising the submission.

Curation Actions Performed on Metadata

* None - the data submitter is revising the submission.

Known Issues [Potentially Impacting Reuse]

* N/A

[[table of contents](#) | [back to top](#)]

Related Datasets

IsRelatedTo

Rumberger, C., Lotterhos, K., Eppley, M., Small, J., Katsuki, S., Mongillo, N., Bajaj, K., Carnegie, R., Segnitz, Z. (2026) **Eastern oyster (*Crassostrea virginica*) mortality, growth, and tissue sample metadata from the common garden experiments run as part of CviMVP project from 2023-2025 in the Chesapeake Bay**. Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2026-09-10 <http://lod.bco-dmo.org/id/dataset/1007522> [[view at BCO-DMO](#)]
Relationship Description: Data from the same common garden experiments.

[[table of contents](#) | [back to top](#)]

Parameters

Parameters for this dataset have not yet been identified

[[table of contents](#) | [back to top](#)]

Project Information

CAREER: Evaluation of machine learning algorithms for understanding and predicting adaptation to multivariate environments with a Model Validation Program (MVP) (Model Validation Program)

Coverage: East coast of North America

NSF Award Abstract:

Environmental change can be rapid and involve multiple aspects of the environment changing at the same time, such as warming and increased disease pressure. Rapid environmental change threatens the productivity of aquaculture and crops on which humans depend. Predicting organisms' vulnerabilities to rapid and multifactor environmental change, however, is a major scientific challenge. A hurdle to addressing this challenge arises from the complex and non-intuitive ways that organisms adapt, through changes at the level of the DNA sequence, to many environmental stresses at the same time. Thus, there is a need for new approaches to understand and predict adaptation in multivariate environments. To address this need, this project integrates research and education with a Model Validation Program (MVP). The research is developing and evaluating Machine Learning Algorithms (MLAs) for understanding and predicting adaptation of organisms to multivariate environments from their DNA sequences. To evaluate MLAs, this research combines both data simulation and an empirical test in the field with the Eastern Oyster, which provide important ecosystem services and support a multi-million dollar industry. For oysters, this research is studying how temperature, disease pressure, and salinity interact with evolutionary history to determine fitness in the field. This research advances efforts toward addressing the major scientific challenge of predicting adaptation in complex environments by integrating concepts across the frontiers of marine, evolutionary, and statistical sciences in a new way. Machine learning and model validation are not traditionally taught in the marine and environmental sciences, but are becoming increasingly relevant to these fields. As part of a broader education program, this research is developing MVP Learning Modules for high school students and undergraduates, which help students build the foundational knowledge they need to critically evaluate and apply models. Modules are being disseminated to hundreds of students in the greater Boston area and are being made available online for widespread use. The MVP mentoring program is training graduate students, undergraduates, and high school students in marine evolutionary ecology, statistical genomics, and machine learning. This research addresses a pressing societal need to more informatively match genotypes to environments for restoration, farming, and assisted gene flow efforts. Results are being disseminated to stakeholders in the oyster industry.

The goal of this research is to evaluate if MLAs, which can model non-linearities, can be used to understand and predict adaptation to multivariate environments under a wide range of scenarios. In Objective 1, the Principal Investigator (PI) is creating simulated datasets with different aspects of realism, and using them to evaluate and refine the MLAs. This novel set of simulations is studying genome evolution under high gene flow in complex, multivariate environments. In Objective 2, the PI is building on their expertise with the Eastern oyster to evaluate the MLAs in a field setting. The PI is first developing a comprehensive seascape genomic dataset and using it to train MLAs to predict an individual's multivariate environment based on a single nucleotide polymorphism genotype. Then, the PI is testing if the MLA prediction can predict the fitness of

different genotypes from across the species range when raised in common garden field conditions. In Objective 3, the PI is integrating research and education by using the data obtained from Objs. 1 and 2 to develop a series of original "MVP Learning Modules" with interactive web apps for persons at different levels of understanding, using the relatable example of an oyster restoration project. This research lays the foundation for future studies by producing datasets that could become classical examples for developing and benchmarking innovative modeling approaches.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

[[table of contents](#) | [back to top](#)]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-2043905

[[table of contents](#) | [back to top](#)]