# Deriving the Equations of Ecological Dynamics

## Using Symbolic Regression to replace the manual calculation of ecological functions with higher accuracy and efficiency

## ECOLOGISTS SEEK NEW METHODS FOR IDENTIFYING POPULATION DYNAMICS

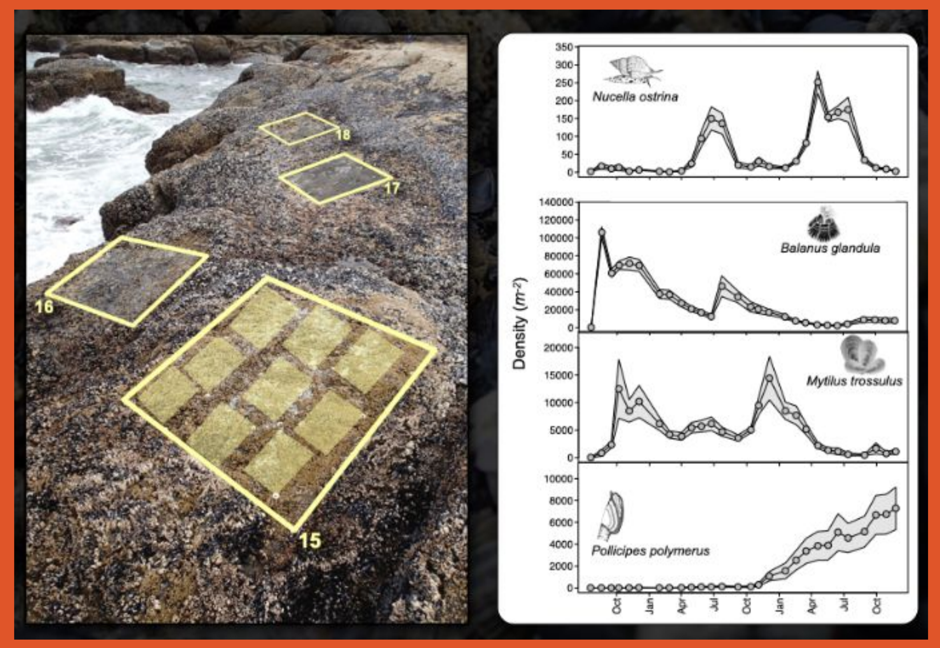

Figure 1: Data collection at the Oregon coast

- Creating models to understand species abundance is a significant area of research in the field of Ecology. Researchers must gain a deep and comprehensive understanding of population dynamics to explain interactions between predators and prey as well as predict population abundance over space and time.
- Historically, ecological researchers have used statistical methods which rely on fitting a “correct” model to data by estimating parameters that are assumed to be significant. However, symbolic regression is a promising new method for producing descriptive models because it has the potential to derive the multiple significant factors and predator-prey interactions directly from species abundance data. In addition, symbolic regression does not assume a prior model which suggests that the one it extracts is the data’s “true” model.

Oregon State University

## SYMBOLIC REGRESSION

Symbolic Regression is a form of regression analysis that seeks to maximize the accuracy of an equation without any prior assumptions. It achieves this by varying the mathematical symbols used in relationship with the input variables to match the defined output data as closely as possible.

The first generation of equations consists of randomly assigned relationships over many iterations before intelligently selecting the best performing models to mutate in successive generations. This process of mutation and selection repeats until a predetermined generation count is reached or the model reaches the ideal accuracy.

$$\frac{Al^2(0.091No - Tp)(3Al - 2Nc + Se)}{No}$$

Figure 3: Example equation derived for Barnacle Growth Rate

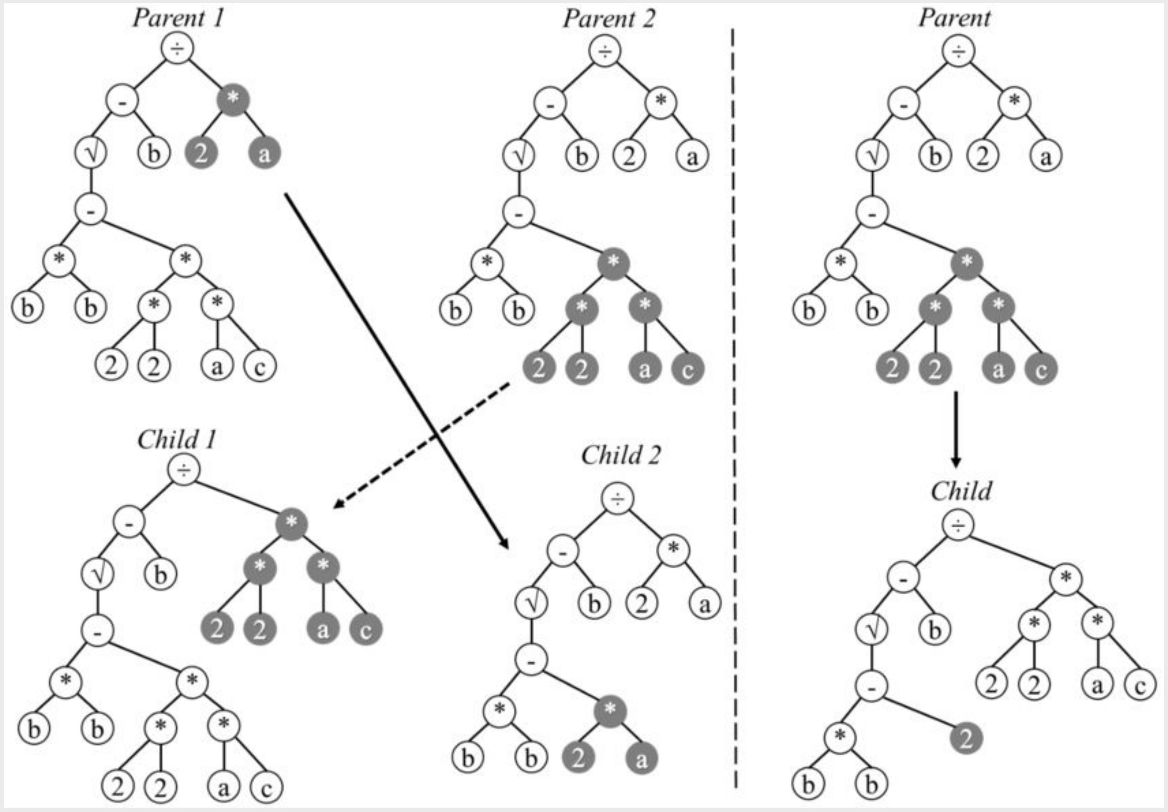


Figure 2: Symbolic Regression visualized with a genetic tree

## IMPORTANT TAKEAWAYS

- Balance - Symbolic Regression is a balancing act that requires the parameters to be adjusted in a way so that the regressor can explore an accurate equation while limiting equation size. This was achieved through trial and error.
- Parameters - Part of being able to balance is knowing how the regressor parameters behave. We discovered the regressor parameters (i.e. parsimony coefficient) are not scaled meaning their values behave differently with different data inputs.
- Transformations - Transforming data is a good way to extract a variety of equations from the same dataset. We applied log functions and removed outliers as we found the regressor was very sensitive to outliers.

## PROJECT DELIVERABLES

Over the course of this project, the development team focused primarily on two equations that define predator-prey relationships in the Oregon Marine Intertidal Region.

The first model focused on the how specie counts affected the growth rate of barnacles and mussels. One equation derived is shown in figure 3. Although a final equation was never achieved, several equations were generated with each one teaching us something new about the dataset or symbolic regression.

The second model focused on the relationship between density and temperature of the patch and how it impacted the feeding rates of the predator species. The model parameters were derived from a method called Grid Search; different combinations of input parameters were used for smaller generation amounts allowing us to further explore the most promising models generated.

## ACKNOWLEDGEMENTS


- Members and contact information:
  Markus Bauer, bauermar@oregonstate.edu
  Samson Mont, montsa@oregonstate.edu
  Nour Rahal-Arabi, rahalarn@oregonstate.edu
  Madelyn Smith, smitmad9@oregonstate.edu
  Jinshui Wang, wangji5@oregonstate.edu


- Our research was conducted in close collaboration with our project partner Mark Novak and his lab, NovakLab. NovakLab examines the interaction between marine species on the Oregon coast and the consequent effects on the structure and dynamics of their ecological communities.
- NovakLab is located in the Department of Integrative Biology at Oregon State University: https://novaklabosu.github.io/


Jinshui Wang, Madelyn Smith, Markus Bauer, Nour Rahal-Arabi, and Samson Mont