**COURSE: Function-Valued Trait Modeling**

Instructor: Rob Baker

TAs: Deannah Neupert and McKenna Burns

**COURSE OVERVIEW:** This course assumes some knowledge of R. The analyzing data collected over the course of time (longitudinal data). In biology, this is a very common practice. For instance, you might be interested in how organisms grow in two different environments. However, analyzing these data properly can be remarkably tricky. Here we will learn about one method that uses mathematical modeling to examine trends in data. We will focus on plant development over time, but the same technique could be applied to animals, bacteria, or community composition of mutualist partners, or gradients of biotic or abiotic stress such as pathogen load.

**Course Materials:**

Workshop Website: <https://compscibio.com/courses/>

Module materials: <https://github.com/ComputerScienceinModernBiology?tab=repositories>

Alternative githup: <https://github.com/rlbaker5/AJB_FVTmodeling2021>

Citation (background reading & links to all materials):

**Baker, R. L.** and D. R. Wang. 2021*.* “Working with longitudinal data: quantifying developmental processes using function-valued trait modeling”. *American Journal of Botany.* <https://doi.org/10.1002/ajb2.1677>

**LOGISTICS:**

Software Needed:

1. R and RStudio (will be teaching/working in RStudio, which requires installation of R)

* Download R: <https://cloud.r-project.org/>
* Download RStudio: <https://rstudio.com/products/rstudio/download/>

1. Packages Day 1: “plyr” and “lubridate”, “car”, “tidyverse”

Day 2: “tidyverse”, “readxl”, “filesstrings”

* + Not vital to install before workshop, as discussing packages is part of the lesson plan
  + If not installed prior, however, participant will need internet connection to download during lesson (assume will have internet to connect over Zoom?)

Preparation Needed:

1. Install R, RStudio, and possibly packages on personal computers
2. Download tutorial .html files. These will need to be downloaded and then opened locally on your computer; you will not be able to read them if you try to access them directly from github:

Day 1: “S2\_FVTmodeling.html”

Day2: “S3\_areaundercurve.html”

1. Download and know where you have saved the example data sets. The dataset for Day 1 is on github; the dataset from Day 2 is from a peer reviewed publication:

Day 1: “CEPF\_TopSurface2020.csv

Day2: Data is from:

“Epistatic Transcription Factor Networks Differentially Modulate Arabidopsis Growth and Defense” Genetics vol. 214 no. 2 529-541; [https://doi.org/10.1534/genetics.119.302996.](https://doi.org/10.1534/genetics.119.302996)

The data are available at:

<https://gsajournals.figshare.com/articles/dataset/Supplemental_Material_for_Li_et_al_2020/11388195>

Download the and unzip the file “S1.zip” (2.92 MB) (the link is near the bottom of the screen).

**LESSON PLAN:**

Day 1:

1. **Introductions: to the Workshop and to the Instructors and TAs**
   1. **The workshop**
   2. **The instructors**
      1. **Rob Baker**
      2. **Deannah Neupert**
      3. **McKenna Burns**
   3. **Survey**
2. **Function Valued Traits: What are they and why should you care?**
3. **Overview of course objectives**
4. **FVT modeling: logistic growth curves**
   1. Prep:
      1. Install R, R-studio
      2. Download access tutorial files (.html)
      3. Download example data

**BREAK: 15 Minutes**

* 1. Setting up the analysis:
     1. Download required packages
     2. Load and explore data
     3. Reformat data for analysis
  2. FVT modeling: a single individual
     1. Visual data inspection
     2. Fitting logistic growth curves
     3. Plotting growth cruves
     4. Extracting parameters of the logistic growth function

**BREAK: 15 MINUTES**

* 1. FVT modeling: the entire dataset
     1. Setting up for-loops and if-statements
     2. Set up a dataframe to store extracted parameters
     3. Plot all the data
  2. Downstream analyses:
     1. Linear models for all paremeters
     2. Assess outcomes via p-values
     3. Data interpretation

1. **Interpretation and wrap up**
2. **Workshop individual datasets as necessary/available**

Day 2:

1. **Quick review from Day 1: 15 min**
2. **Pitfalls and complications for FVT modeling**
3. **Example: genotypes fit to two different functions**
   1. Prep:
      1. Download and access tutorial files (.html)
      2. Download and unzip example data files (from figshare)
   2. Setting up the analysis:
      1. Download required packages
      2. Load and explore data
      3. Reformat data for analysis
   3. Test case with graphical interpretation

**BREAK: 15 MINUTES**

* 1. Apply the functions to the entire dataset
     1. For-loops
     2. Extracting parameters and storing them
  2. Compare “area under the curve” approaches to other approaches
     1. Correlations to individual values within the dataset
     2. Compare with slopes calculated in Li et al, 2020

**BREAK: 15 MINUTES**

1. Workshop individual datasets as necessary/available
2. Wrap up, etc.