

# PhD Lightning Round

## *Diversity-Interactions Models*

Laura Byrne, [byrnel54@tcd.ie](mailto:byrnel54@tcd.ie)



Trinity College Dublin  
Coláiste na Tríonóide, Baile Átha Cliath  
The University of Dublin





First Direction  
*Dataset Provided by*  
*Kerry Ryan, with*  
*John Finn &*  
*Fiona Brennan*  
*NUI Galway & Teagasc,*

Applying different link  
functions and  
transformations to standard  
DI models

# Experiment Set-Up

- Site in Wexford
- 6 Species from 3 Functional Groups (grass, legume, herb)
- 43 Plots with Split-Plot Design
- First Focus on Control Plots
- Nitrogen Fertiliser Treatment
  
- Focus on Microbials & Functional Assays of Soil
  - Nutrient Cycling/Storage
  - Nitrification
  - Gene Copy Numbers
  
- Simplex Design for DI Models use



### Simplex Design

Nfert  300  150    Proportions  0  0.1  0.17  0.25  0.5  0.6  1

Communities	Species						Richness
	SownLp	SownPp	SownRc	SownWc	SownCi	SownPI	
1	1	0	0	0	0	0	1
1	1						1
1		1					1
1			1				1
1				1			1
1					1		1
2	0.5	0.5					2
2			0.5	0.5			2
2					0.5	0.5	2
4	0.25	0.25	0.25	0.25			4
4	0.25	0.25			0.25	0.25	4
4			0.25	0.25	0.25	0.25	4
5	0.6		0.1	0.1	0.1	0.1	5
5	0.1	0.1	0.6		0.1	0.1	5
5	0.1	0.1	0.1	0.1	0.6		5
5	0.1	0.1	0.1	0.1		0.6	5
5	0.1	0.1		0.6	0.1	0.1	5
5		0.6	0.1	0.1	0.1	0.1	5
6	0.17	0.17	0.17	0.17	0.17	0.17	6

# Simplex Design

- Columns are Sown Proportions
- Row is Plot Design
- Monocultures & Mixtures
  
- Set up for FG Comparisons
  
- Responses are linearly related to the gradient of these proportions
  - Relationship comes with assumptions, as all do

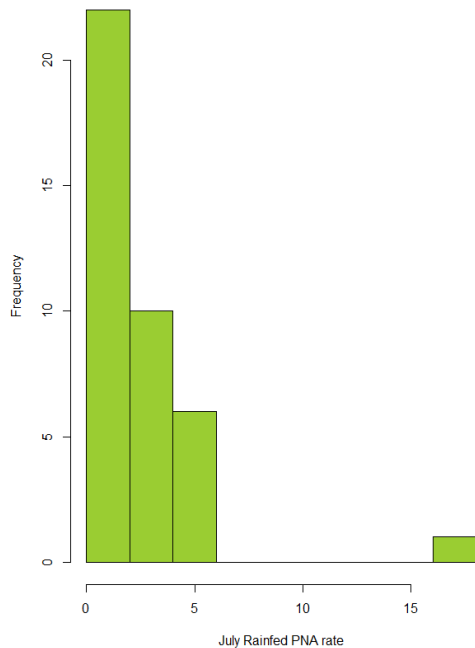
Graphic provided by Rishabh Vishwakarma

## Example Response

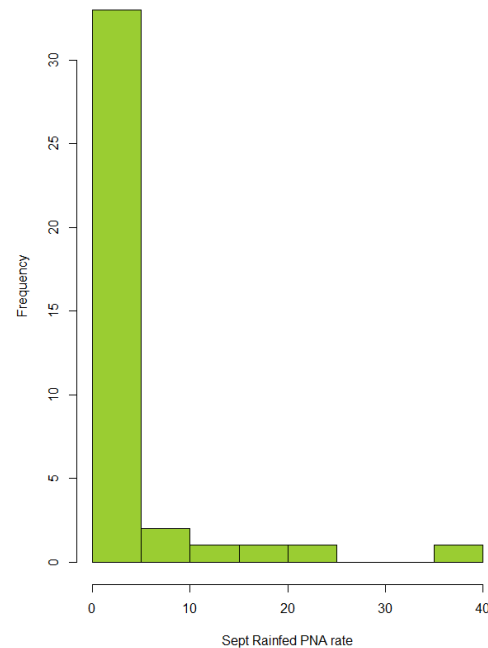
Potential Nitrification Rate of the Rainfed Plots in July/Sept, excluding high N plots

- Highly skewed
- Non-normal Residuals
- Residuals with non-constant variance
- Data is suitable for categorization
- We try binary classification
  - Either nitrification has occurred (1) or it has not (0)
- Now we can use logistic regression

Histogram of datarepJRainSPNArate



Histogram of datarepSRainSPNArate



## Logistic Regression

Model example is Functional Group:

$$\log\left(\frac{p}{1-p}\right) = \beta_G P_G + \beta_L P_L + \beta_H P_H + \\ \beta_G P_G * (\beta_L P_L + \beta_H P_H) + \beta_L P_L * (\beta_H P_H) + \\ \beta_{G1} P_{G1} * \beta_{G2} P_{G2} + \beta_{L1} P_{L1} * \beta_{L2} P_{L2} + \beta_{H1} P_{H1} * \beta_{H2} P_{H2}$$

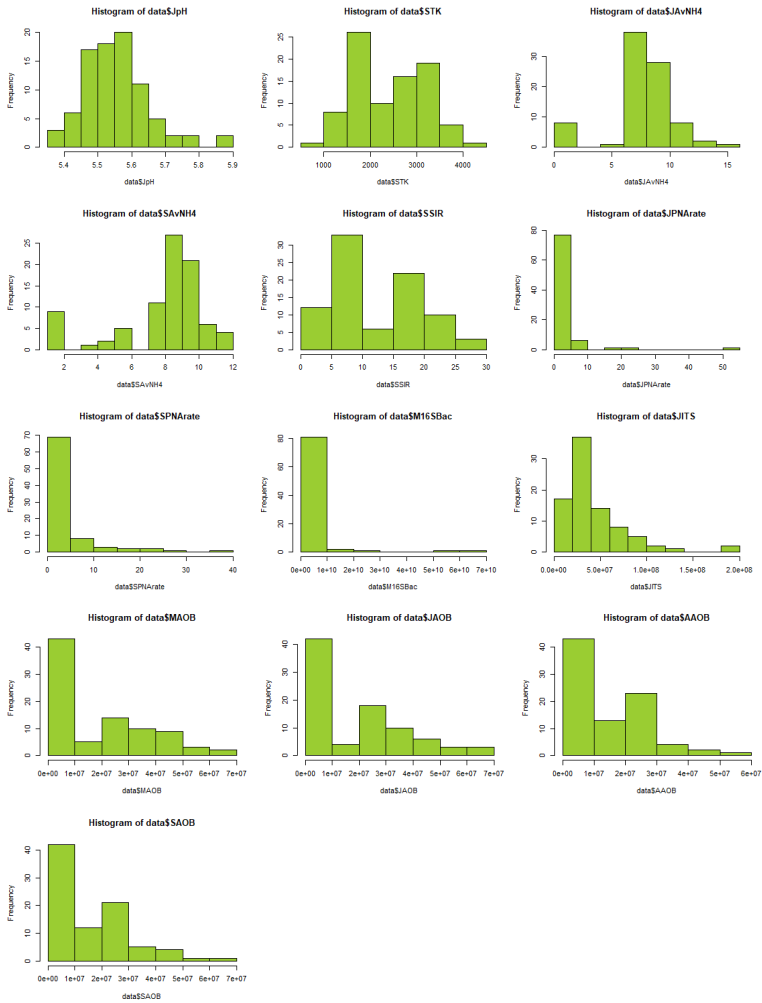
$E[Y] = p$ , where  $Y$  is the binary response variable (PNA rate)

$$\eta = \log\left(\frac{p}{1-p}\right) \quad \& \quad p = \frac{e^\eta}{1+e^\eta}$$

- If the linear predictor estimate is positive, then  $\hat{p} > 0.5$ 
  - where  $p$  is the probability of a non-zero potential nitrification rate



# Other Responses



- July pH, Sept Total K, July/Sept NH<sub>4</sub>, Sept SIR, May 16S bact, July ITS, May/July/Aug/Sept AOB
- Multi-modal and heavily skewed response distributions
- Heavy tailed residuals



## Possible Solutions

- Weighted Least Squares
  - Can be used in cases where some predictors are causing the problem
- Box-Cox
  - Normalises response
  - Fails more often when values are not strictly positive
- Converting responses to binary or ordinal bins
  - Logistic Regression doesn't have the same assumptions
  - Consistency can help inference
  - Loss of information





# Second Direction

*Expansion of this  
idea*

Multifunctional DI models  
and the experimental design  
associated with them



## Idea and Vision

Multifunctionality is the ability of an ecosystem to perform and maintain many functions and services simultaneously

(Nutrient Cycling and Storage, Primary Production, etc.)

Multivariate Diversity-Interactions models have been developed, but jointly modelling binary and continuous responses requires work

### Design requirements:

- » Monocultures & mixtures
- » Balanced/imbanced plots
- » Non-linear regression requires more info
  - More plots necessary
- » Repetitions
  - Elimination or note of environmental factors





# Third Direction *Package Development*

DImodels package extension

## Plans & Progress

For use with

de Andrade Moral R, J Connolly, and C Brophy. (2020)  
DImodels: Diversity-Interactions (DI) Models. R package  
version 1.0.

[https://CRAN.R-project.org/package=DImodels.](https://CRAN.R-project.org/package=DImodels)

Facilitate the use of:

- Multivariate regression models
- Data with repeated measures
- A combination of both

In very early stages of development~



## Underlying Tech


- `lm()` can handle multivariate data
- `lme4` package easily works with repeated measures and can be tricked into working with paired data
- But neither can do both as required
  
- `glmmTMB` package builds on `lme4`
  - Adding design matrix capabilities
  
- `mcglmm` package
  - Built for this purpose, very flexible/customisable
  - But recently removed from CRAN





Trinity College Dublin  
Coláiste na Tríonóide, Baile Átha Cliath  
The University of Dublin

*Thank You*

 Laura Byrne

 [byrne154@tcd.ie](mailto:byrne154@tcd.ie)

