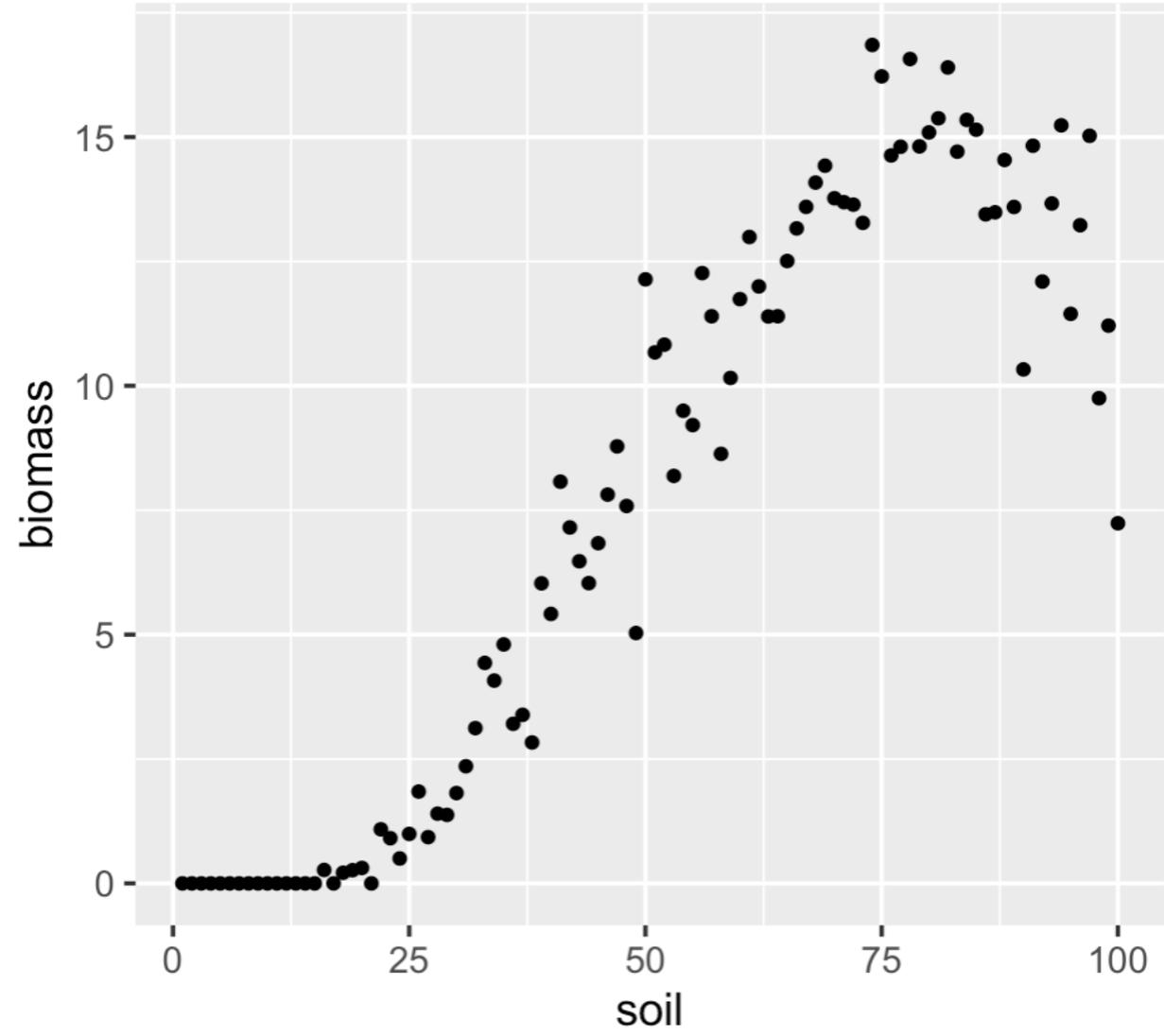
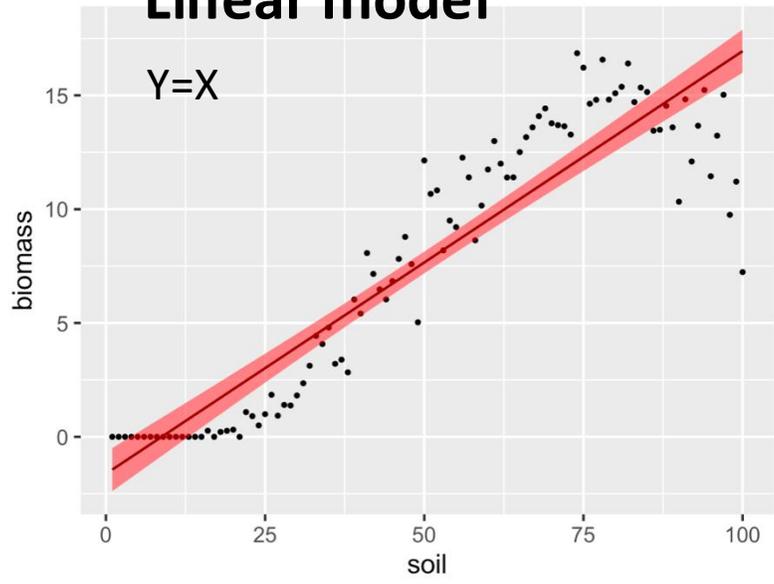


Non-Linearity

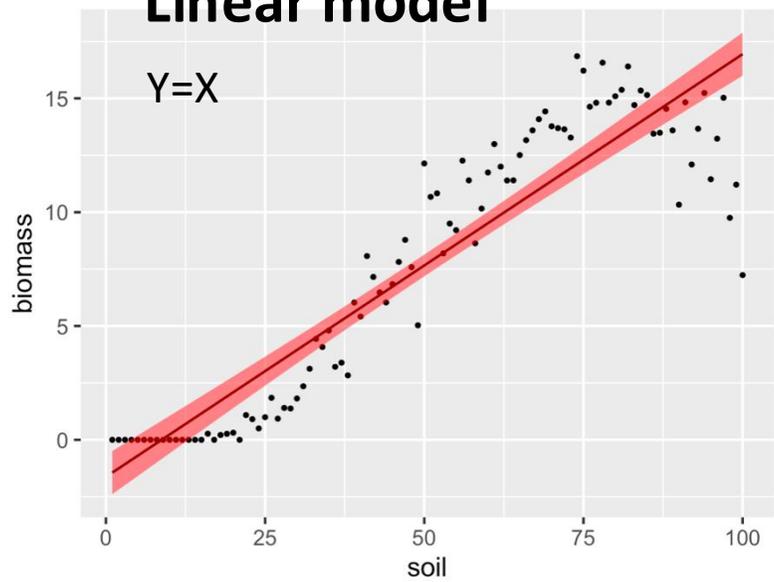
Simple polynomials
Generalized Additive Models



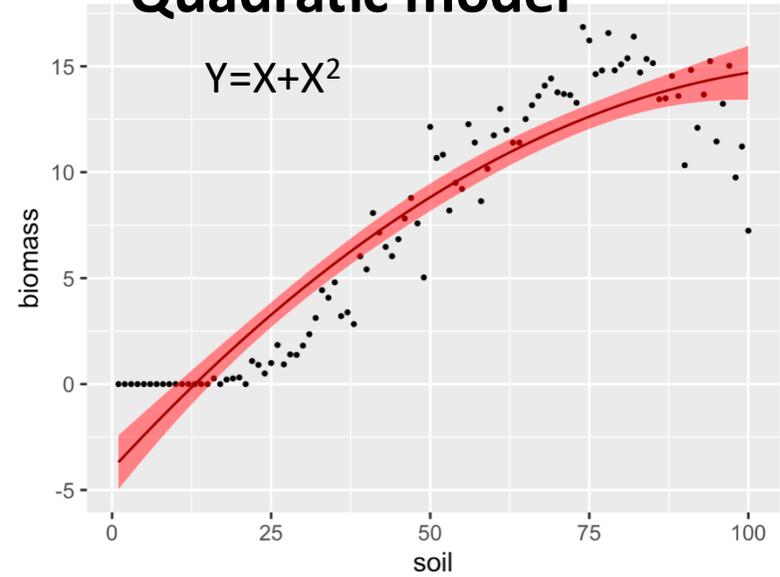
Linear model



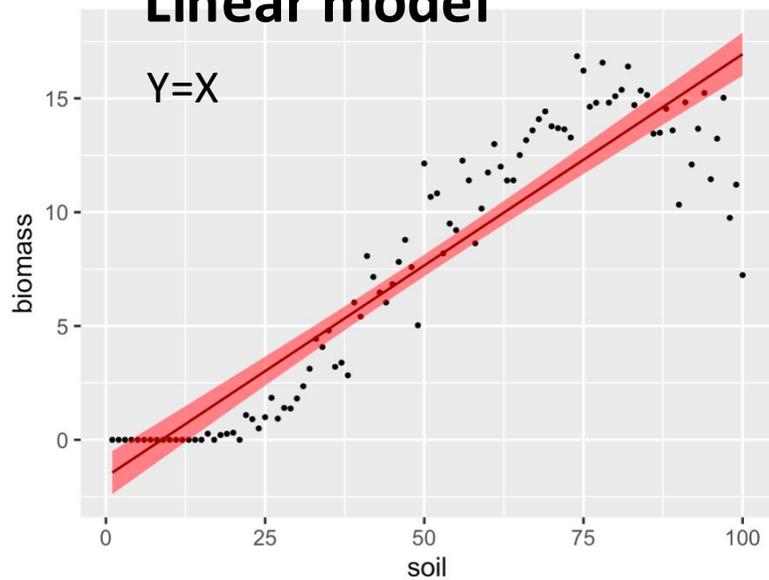
Linear model



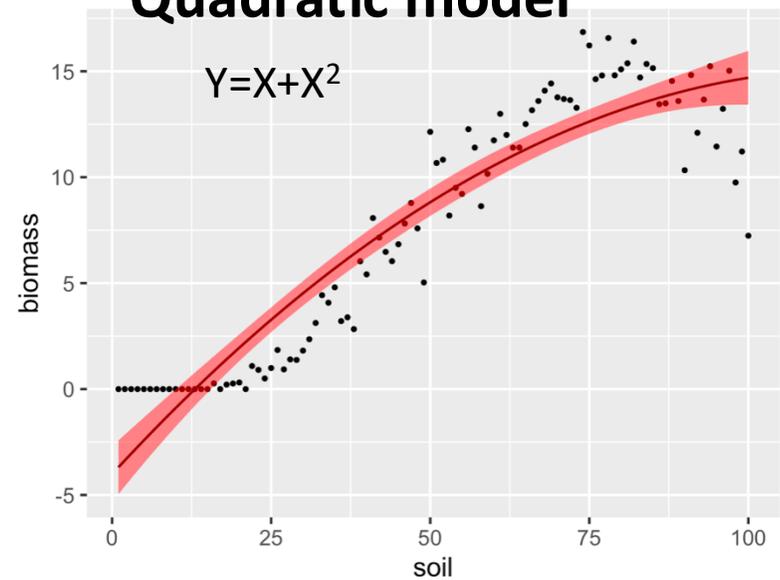
Quadratic model



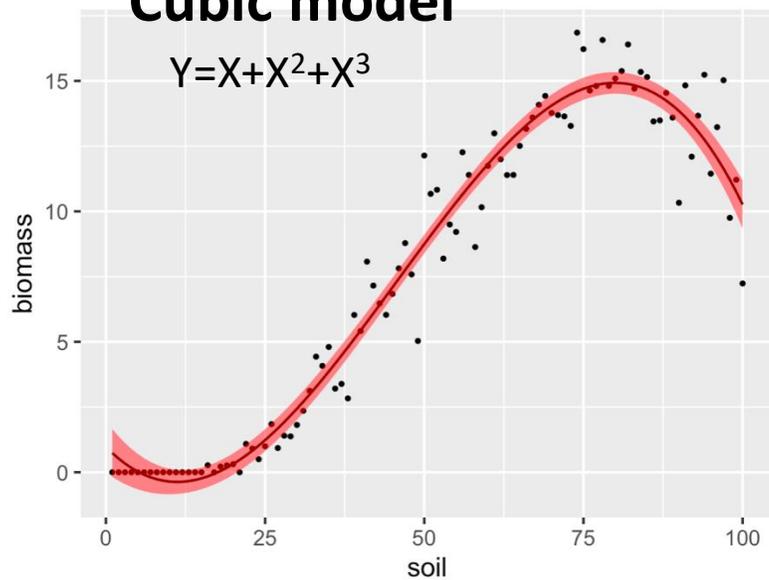
Linear model



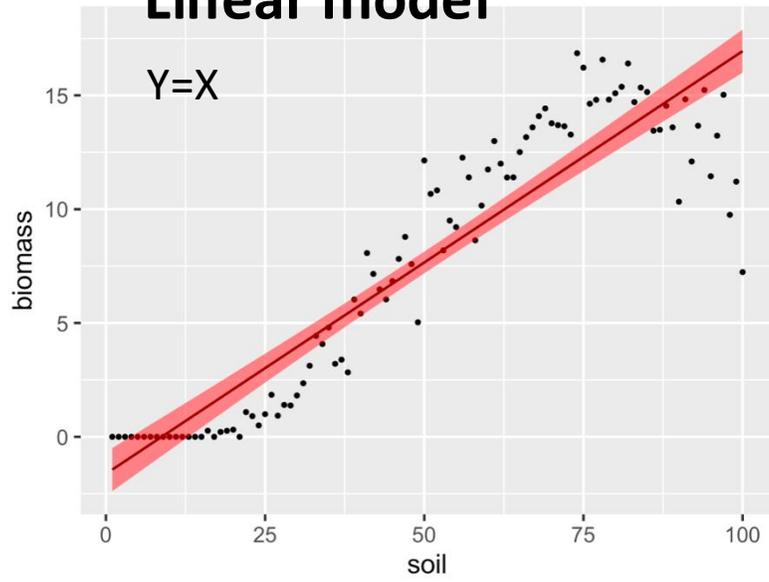
Quadratic model



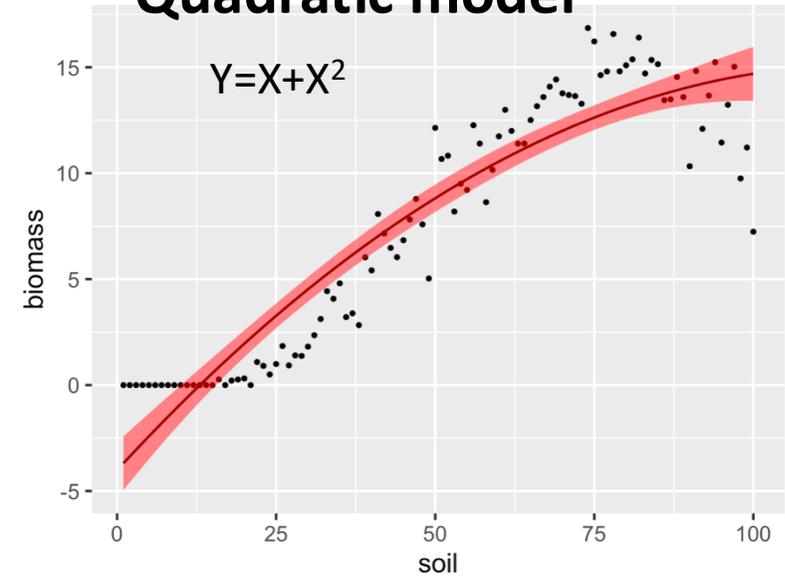
Cubic model



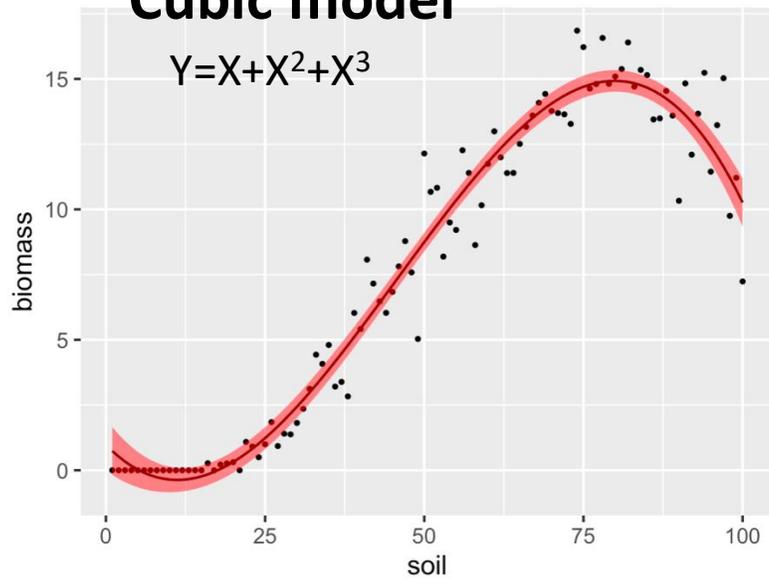
Linear model



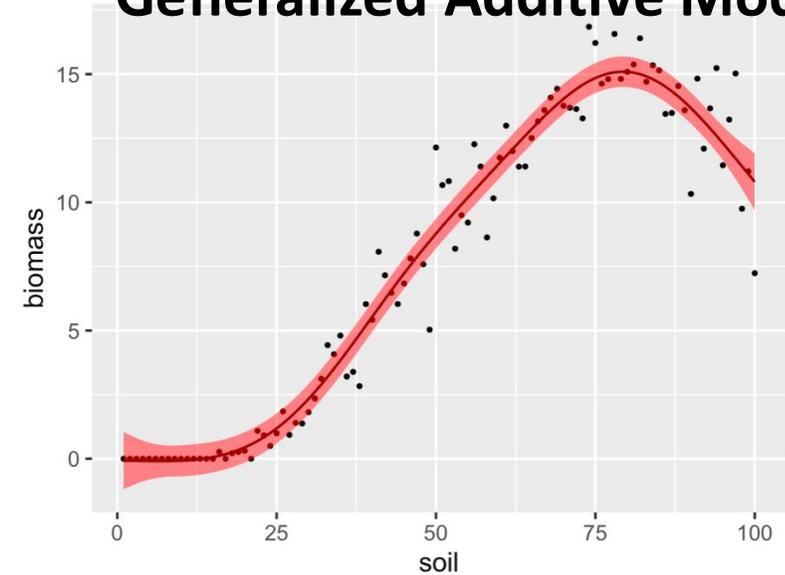
Quadratic model



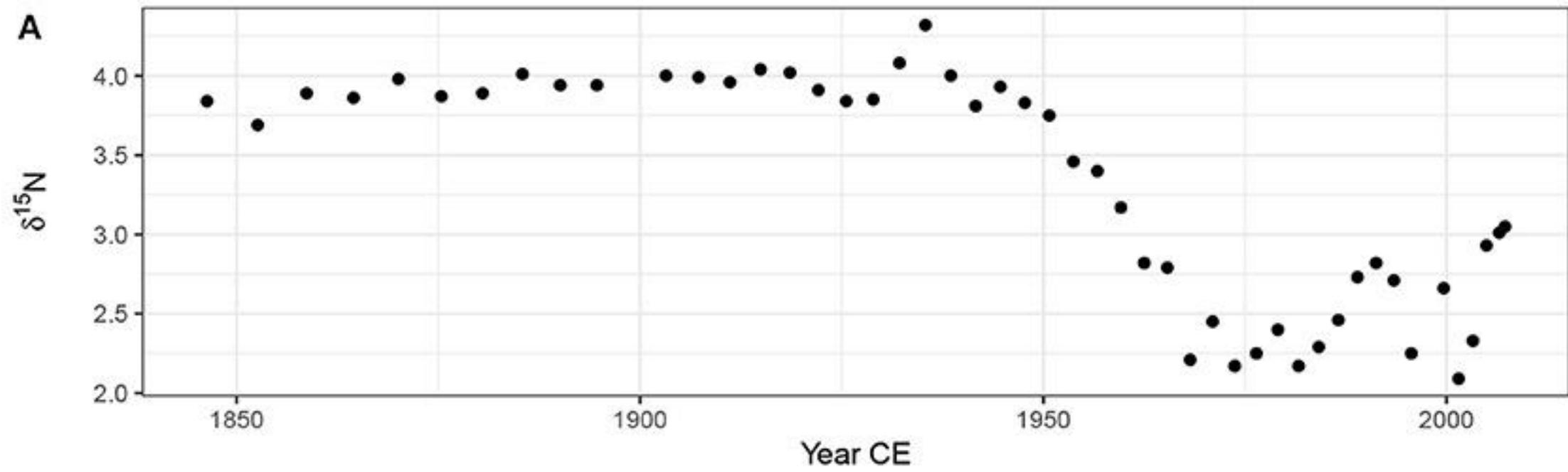
Cubic model



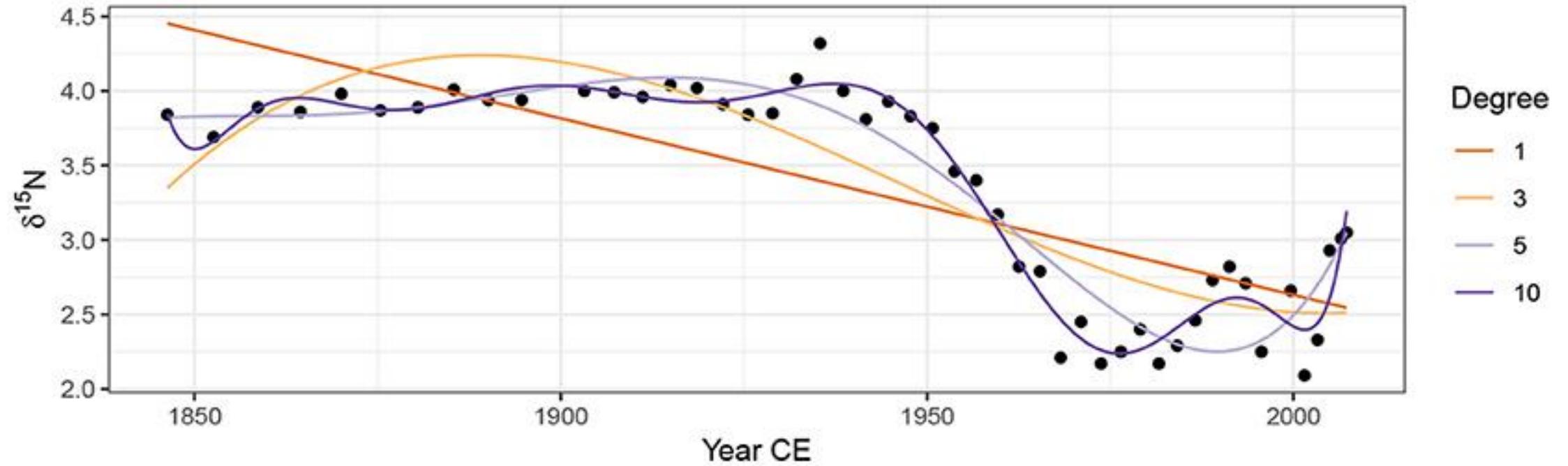
Generalized Additive Model



```
model<-gam(biomass~s(soil), data=species1, method="REML")
```

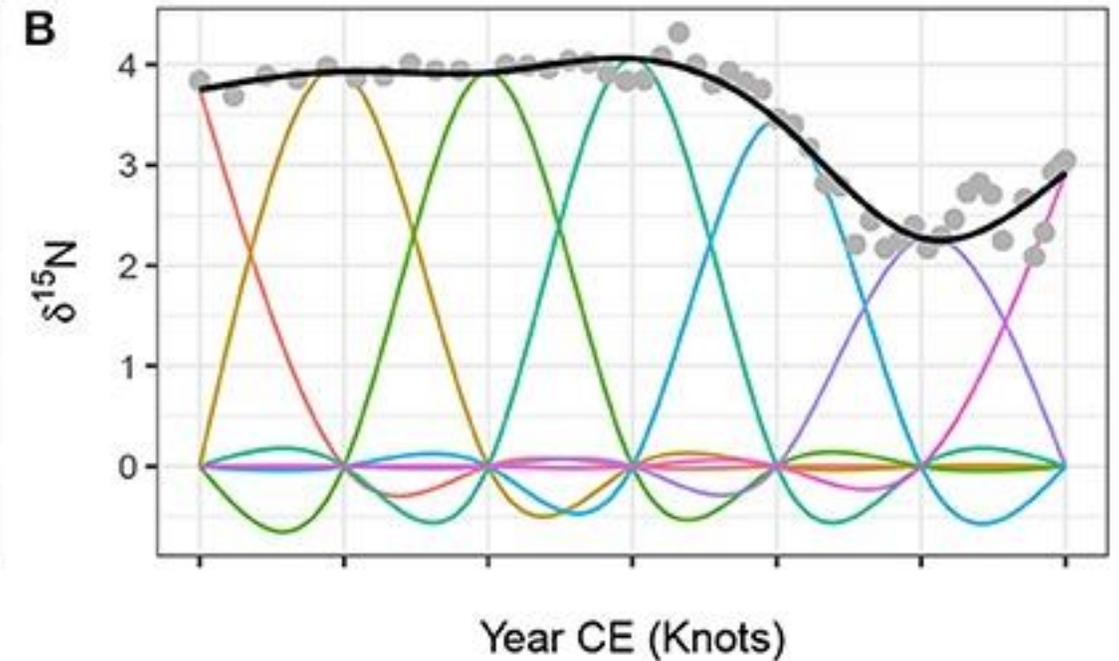
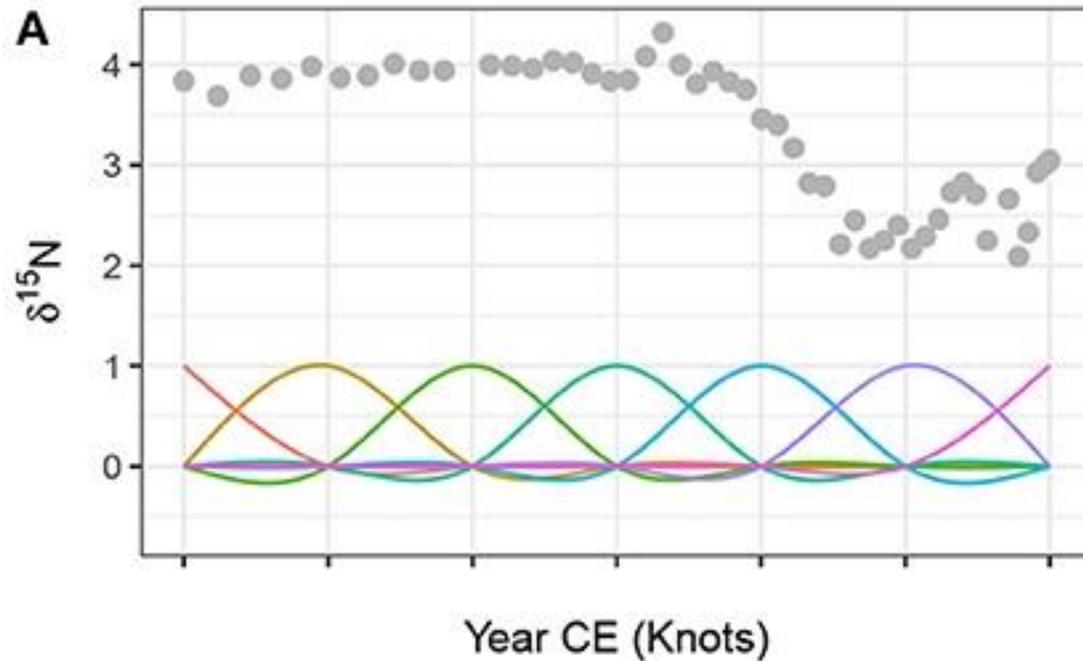


Polynomials

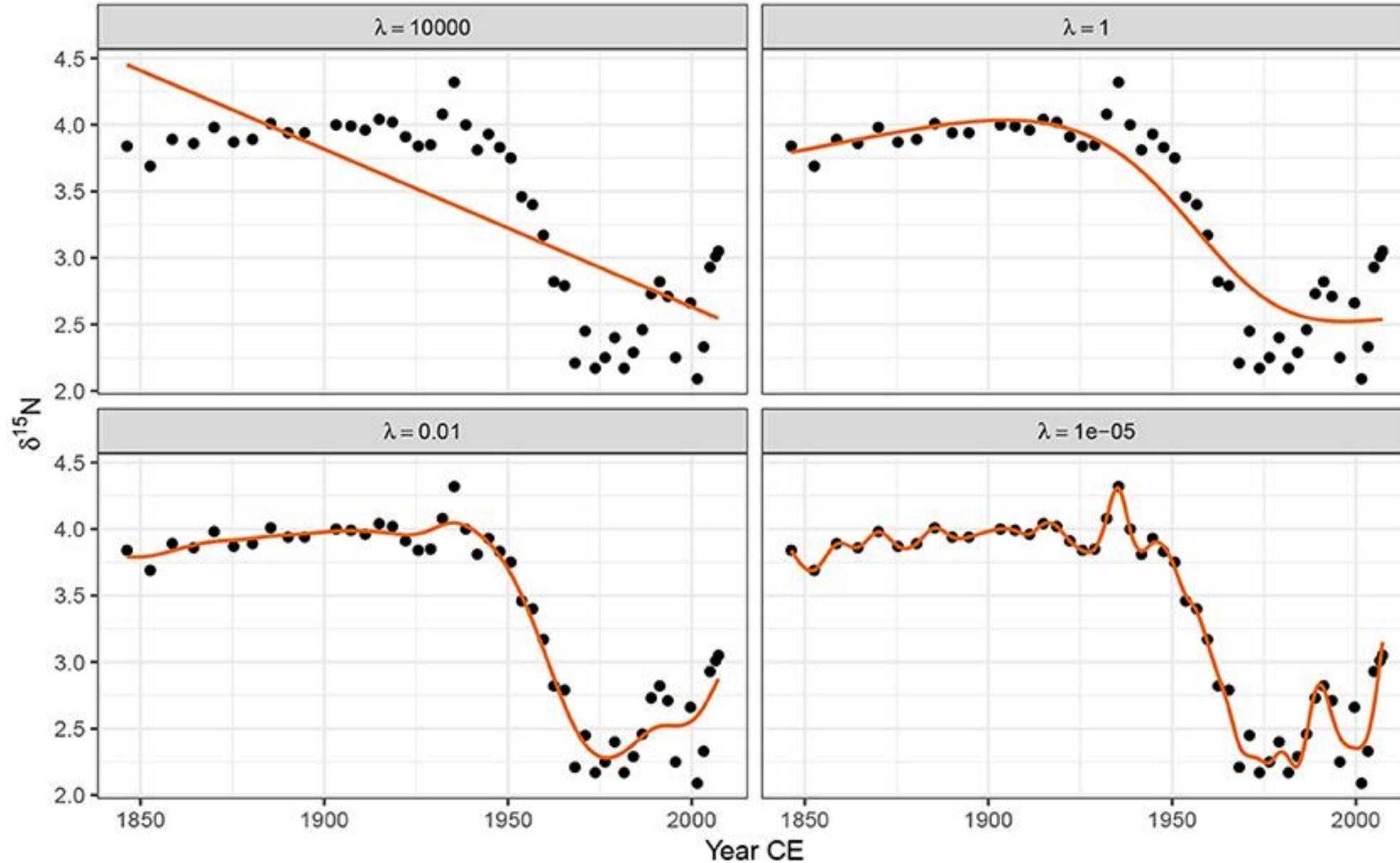


Generalized Additive Model

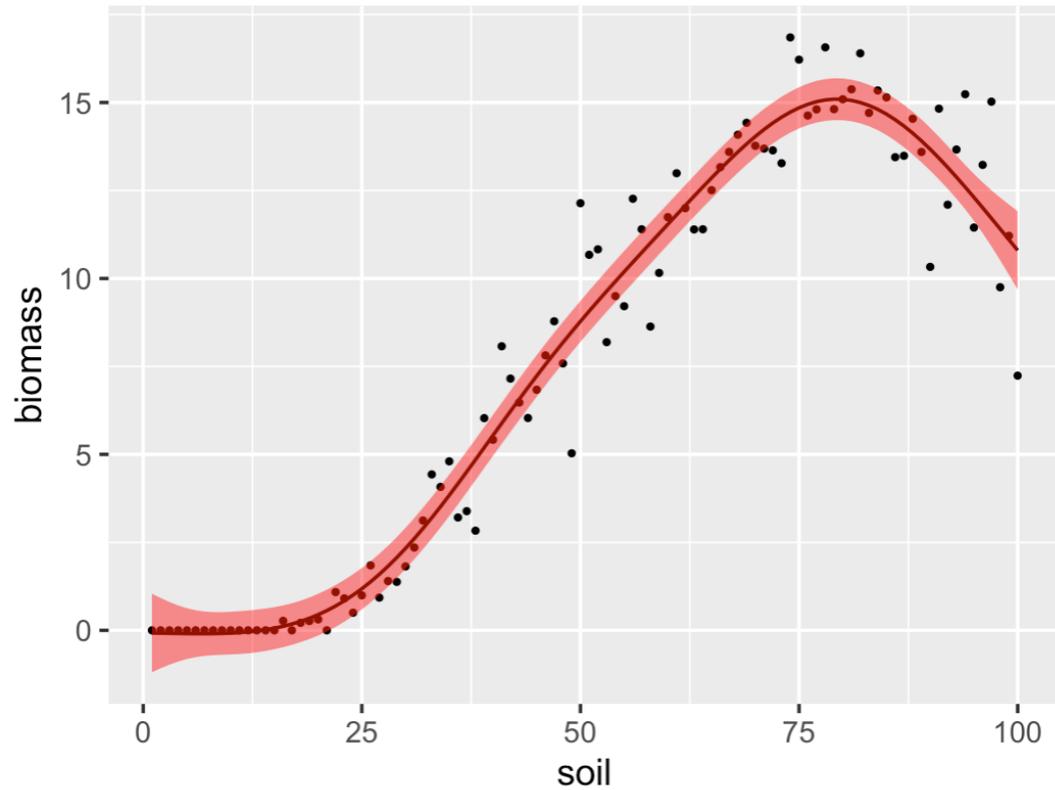
Uses multiple “basis” functions at evenly dispersed “Knots”
Each Knot gets a weight which yields the predicted curve



GAM automatically chooses optimal wiggleness by penalizing lambda



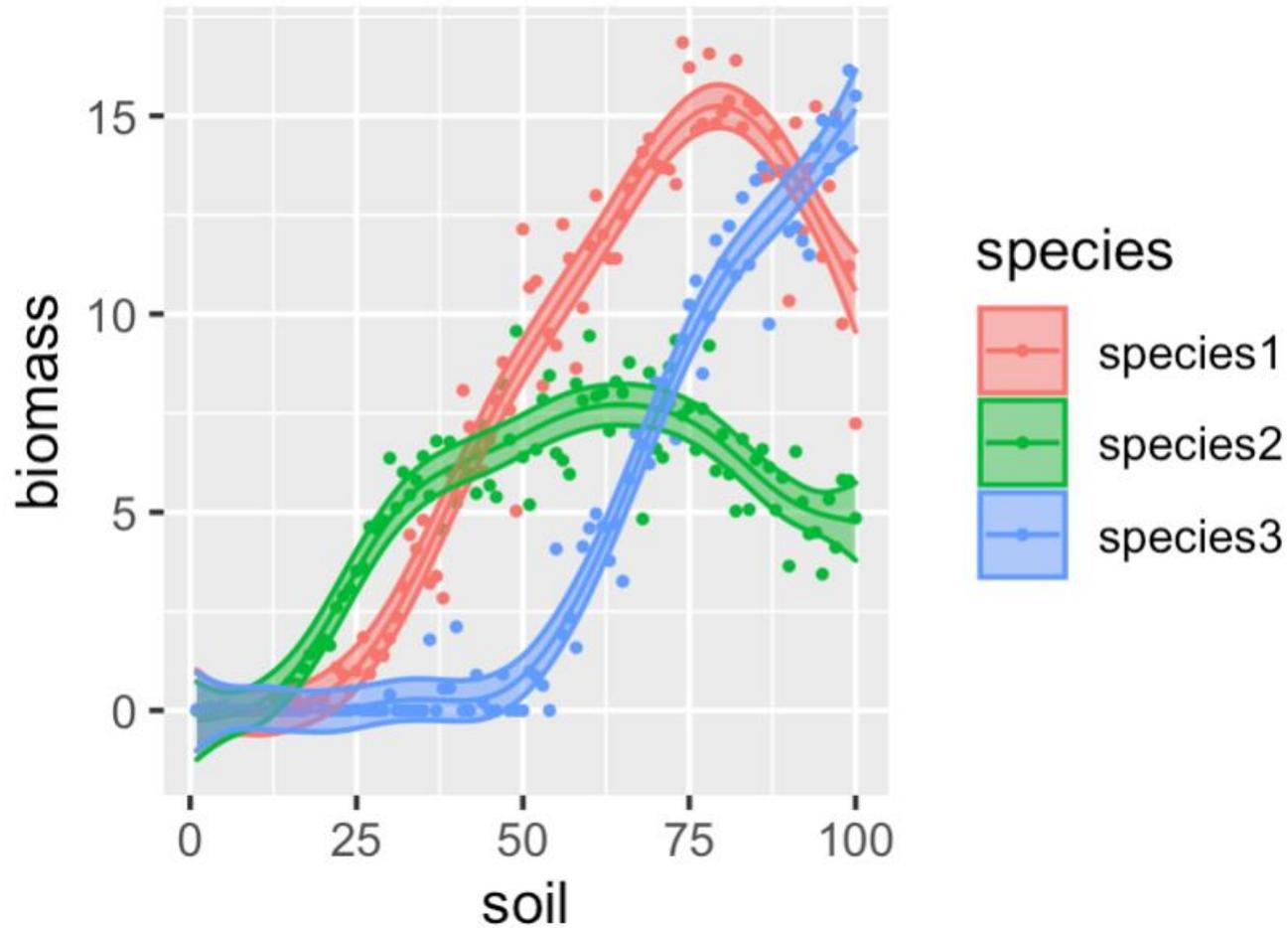
Water response in 1 species



**Generalized Additive Model fit
any arbitrary function**

```
model<-gam(biomass~s(soil), data=species1, method="REML")
```

Water response in 3 species



GAM can model with categorical variables

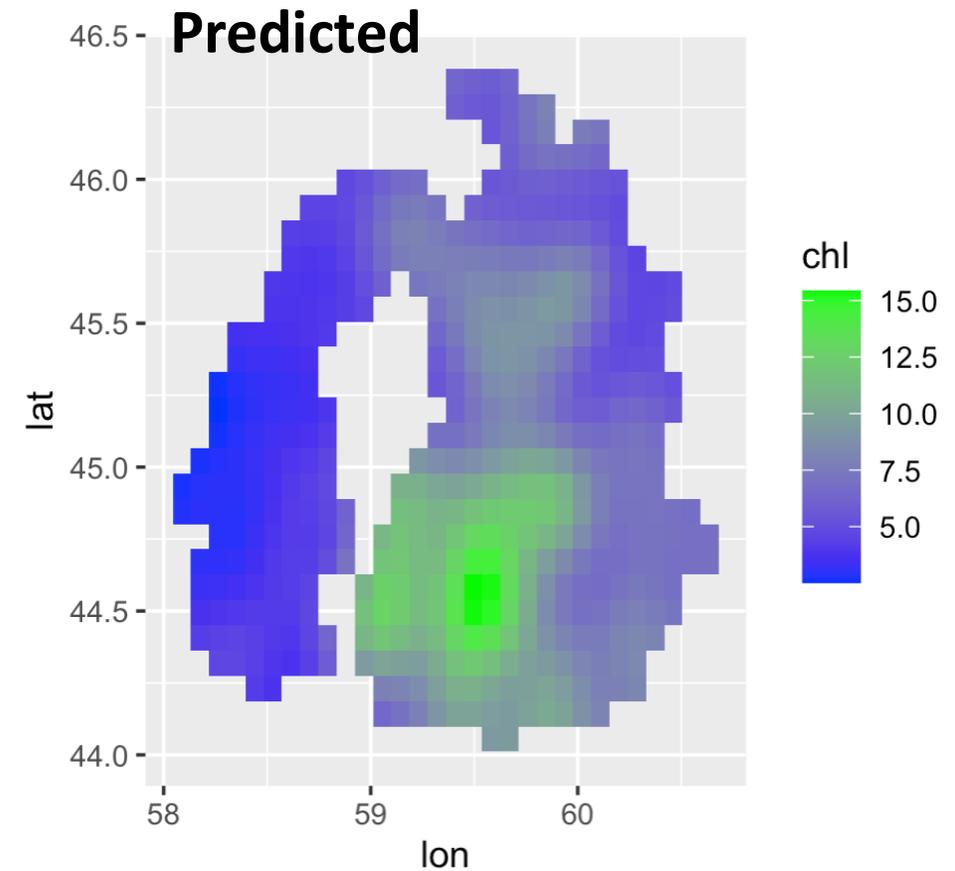
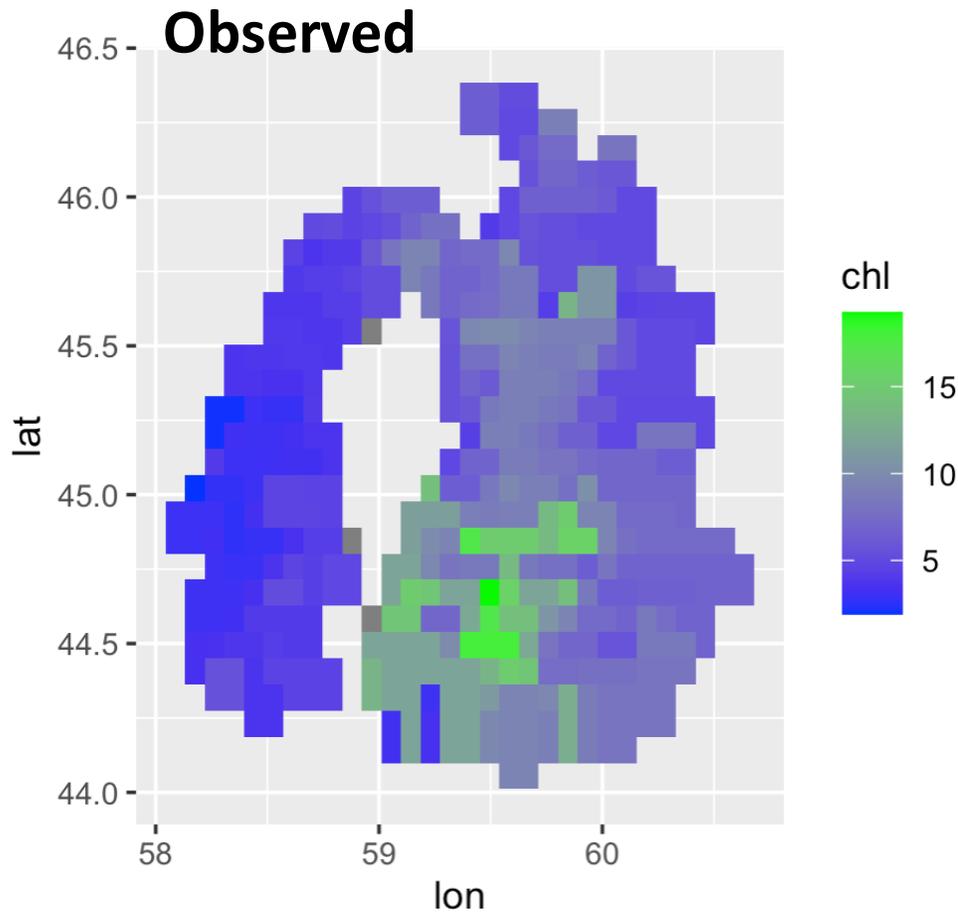
Also different groups can get different functions (interaction with "by =")

```
model<-gam(biomass~species+s(soil, by=species), data=drought, method="REML")
```

GAM can model interactions between quantitative variables

```
model <- gam(chl ~ s(lat, lon, k=100), data=ara1, method = "REML")
```

Modelling
chlorophyll in Aral
Sea from space

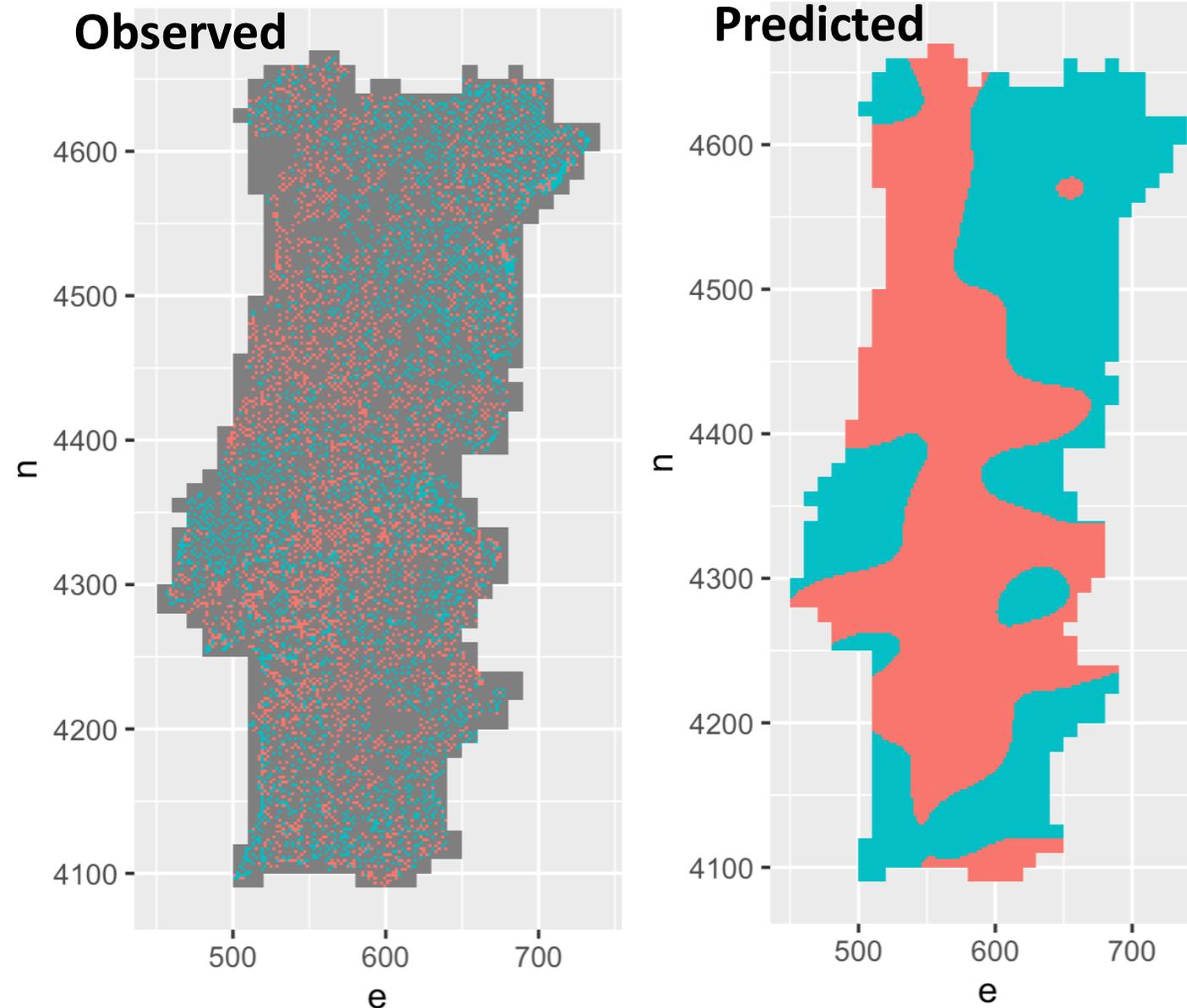




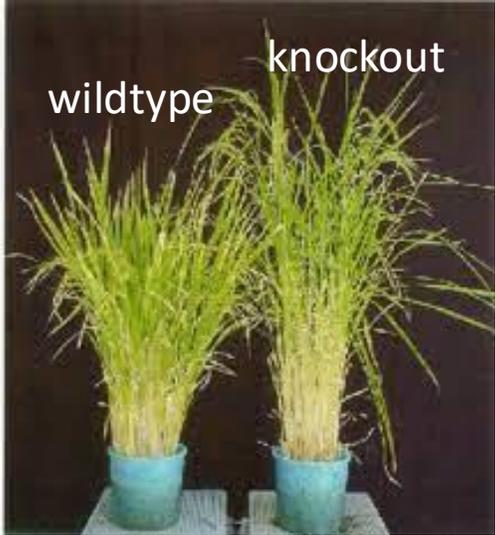
Linnet

GAM can do classification!

Modelling spatial distribution of bird observation across a landscape: binomial regression (i.e. logistic regression with 0/1 data)

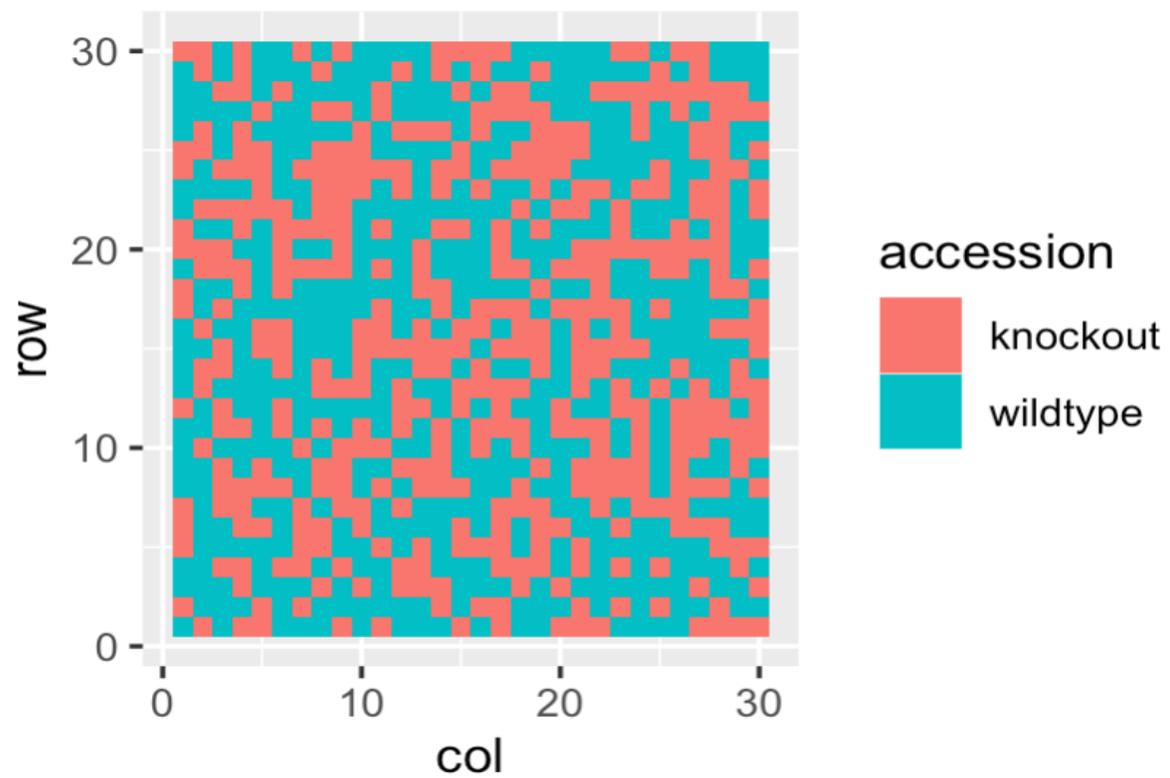


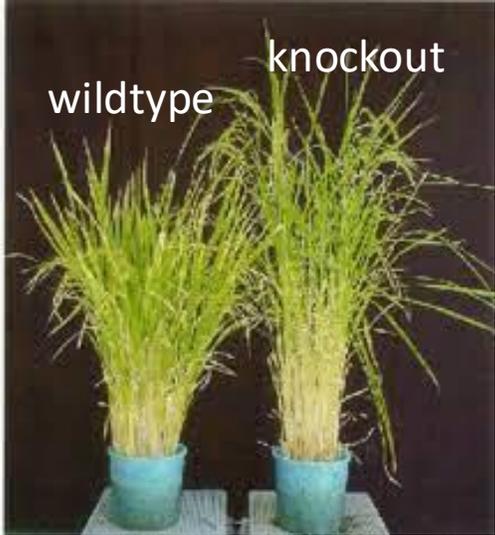
```
model<-gam(linnet~s(e,n, k=200), data=bird, family=binomial, method="REML")
```



Experiment: Do genetic knockout lines of rice perform better than wildtype lines of rice?

Field design: randomized

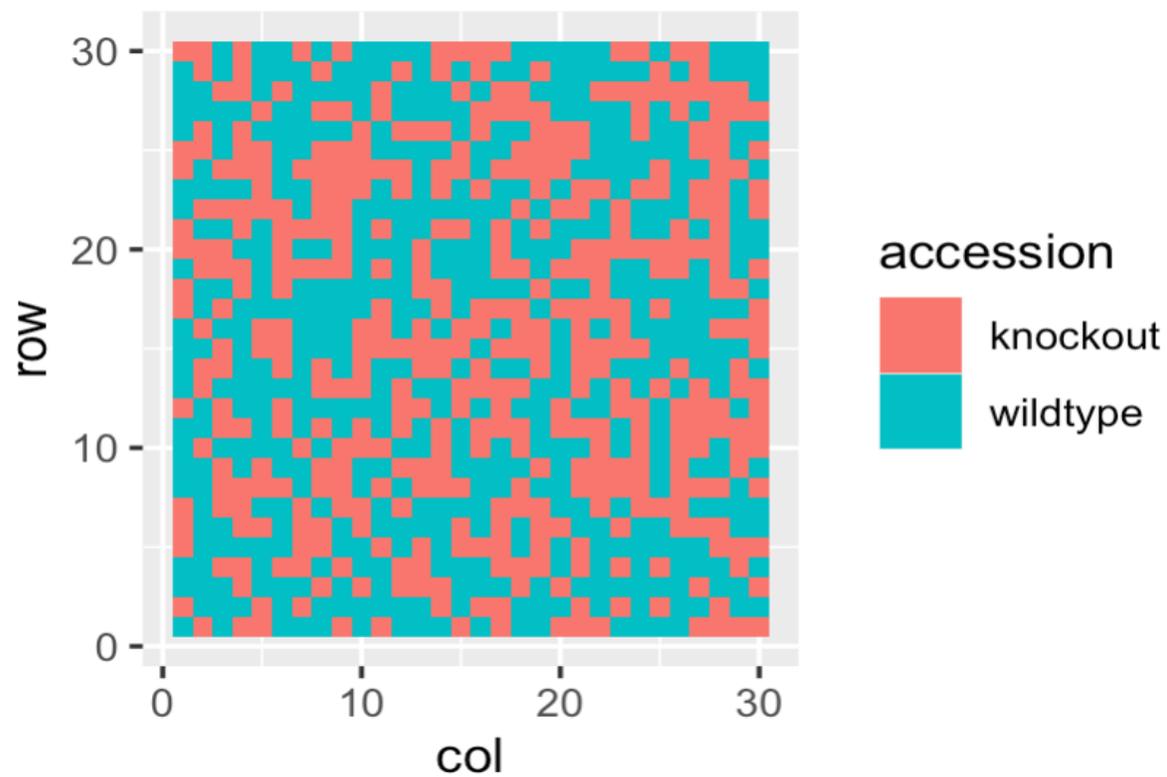




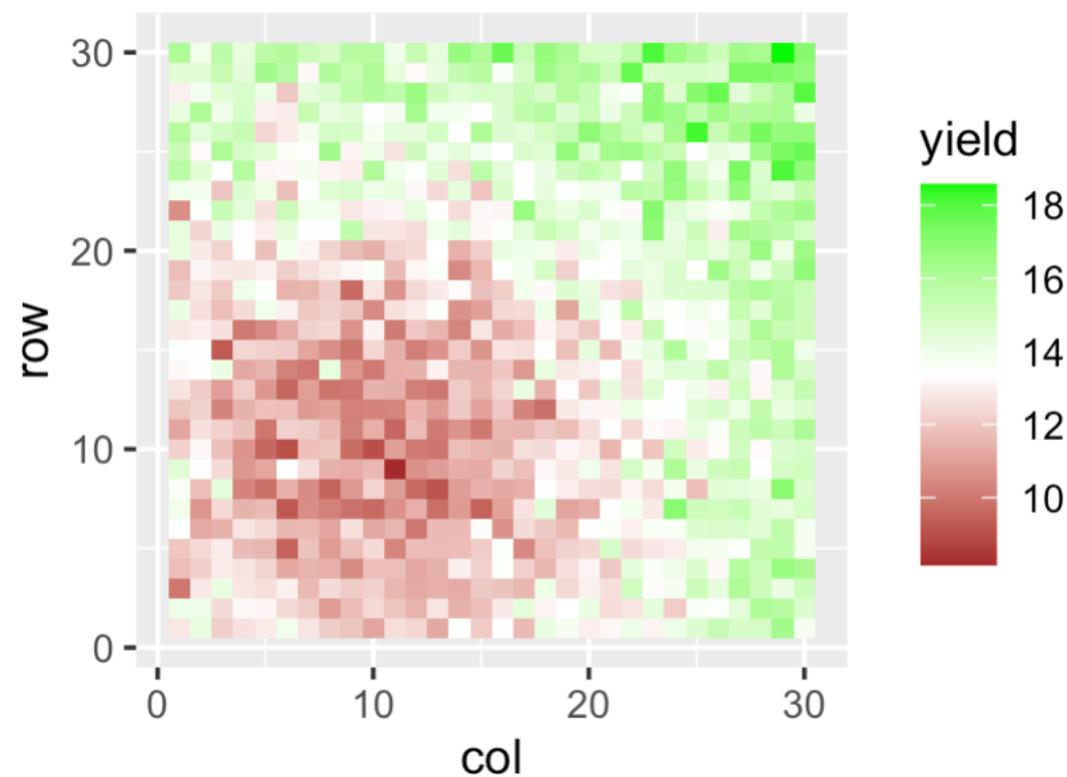
GAM can help account for heterogeneity in the field!

Experiment: Do genetic knockout lines of rice perform better than wildtype lines of rice?

Field design: randomized



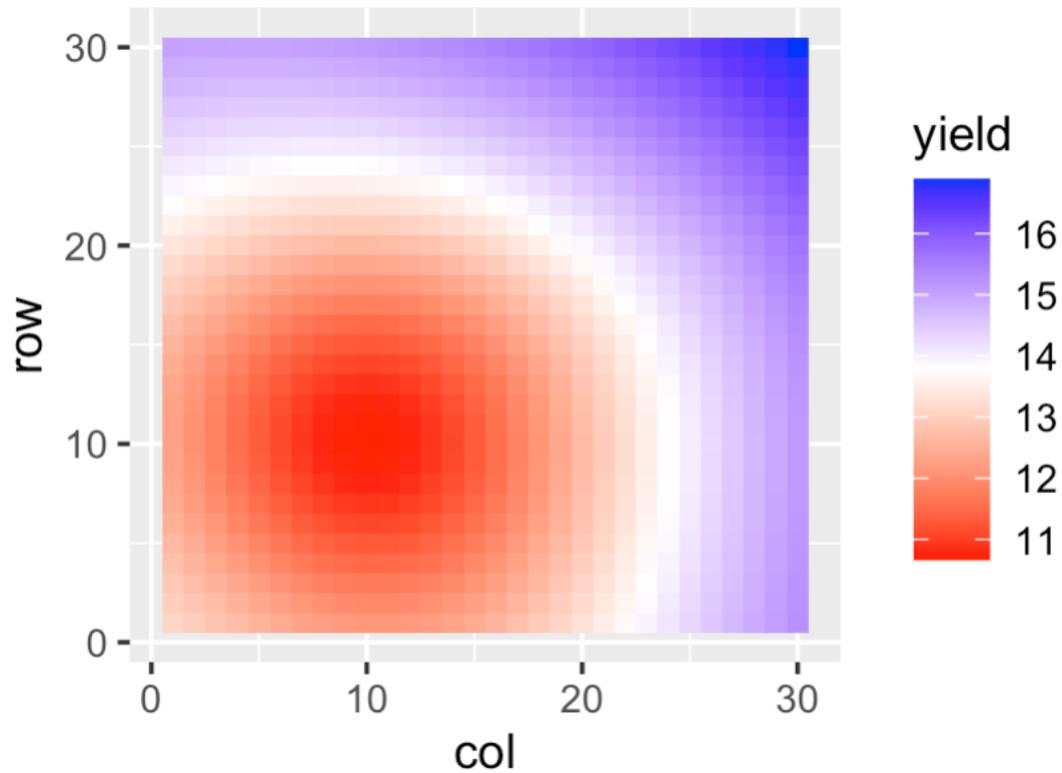
Results of yield on the field



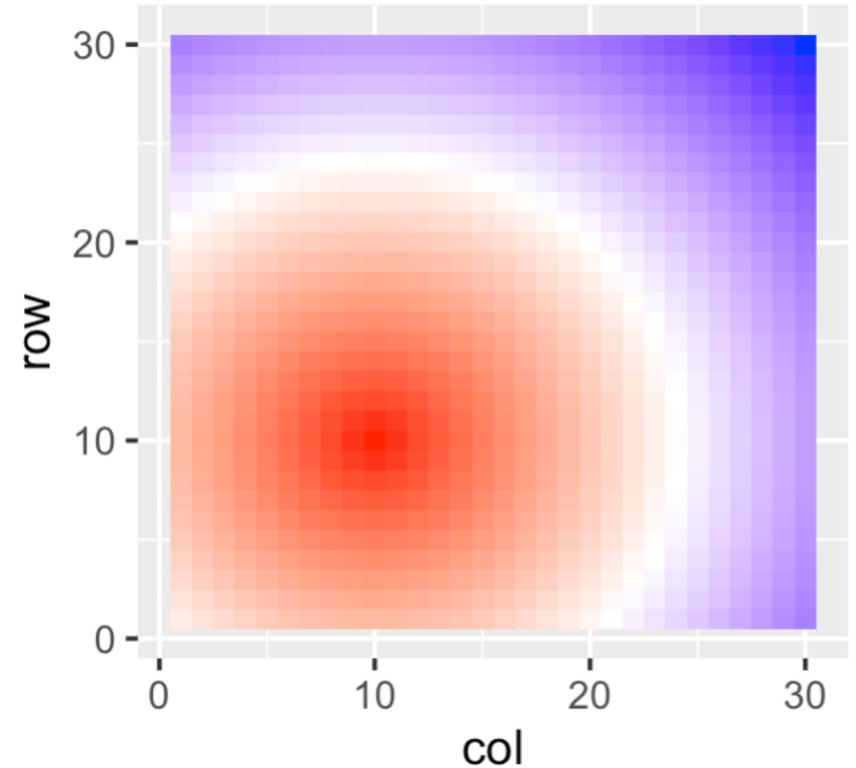
GAM can help account for heterogeneity in the field!

```
model<-gam(yield~accession+s(row,col, k=20), data=ko)
```

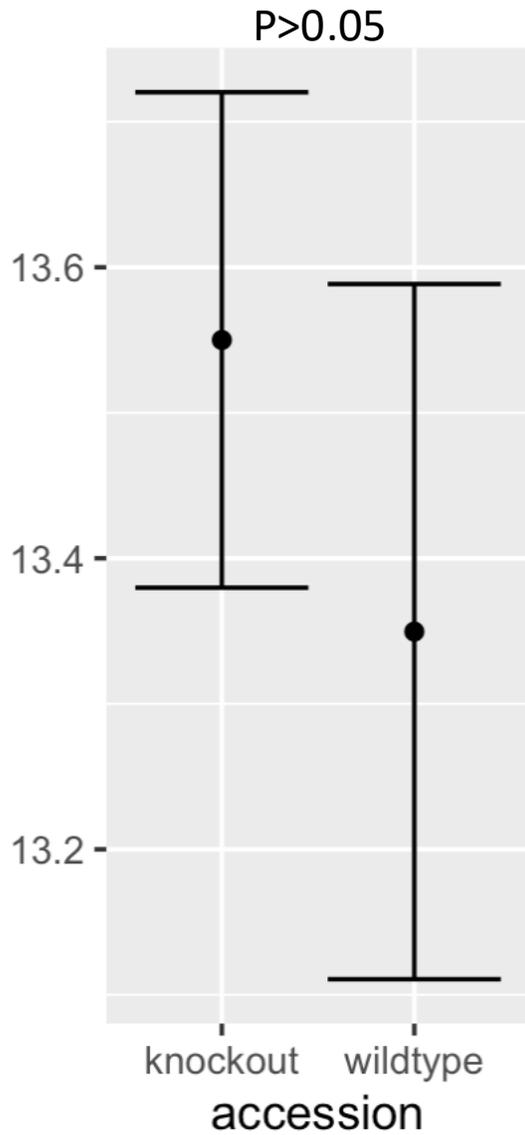
Model prediction:



“Real” effects:

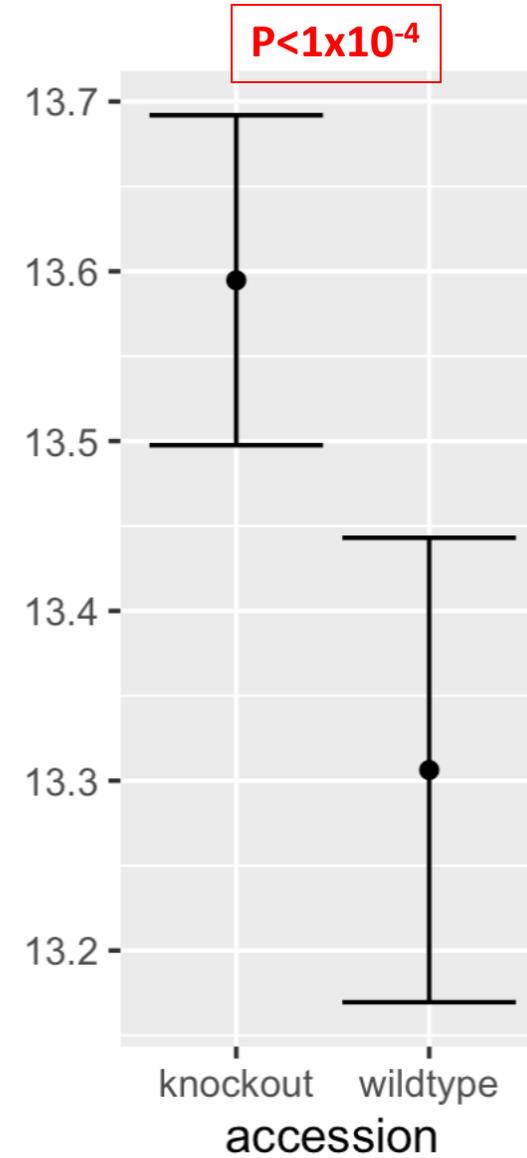


```
model<-gam(yield~accession, data=ko)
```



FYI: Same as lm()

```
model<-gam(yield~accession+s(row,col, k=20), data=ko)
```



Non-Linearity part 2

- **poly(x, 3) vs “x+ I(x²) + I(x³)” ...**
- **Other types of splines and smoothing methods**
- **Other basis functions**
- **Effect of variable scaling in GAM**
- **Model selection with GAM (select=T)**