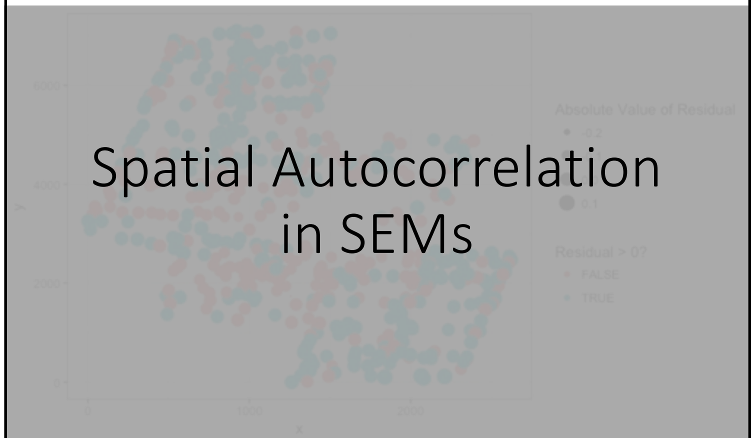


Spatial Autocorrelation in SEMs

Space: The Final SEM?

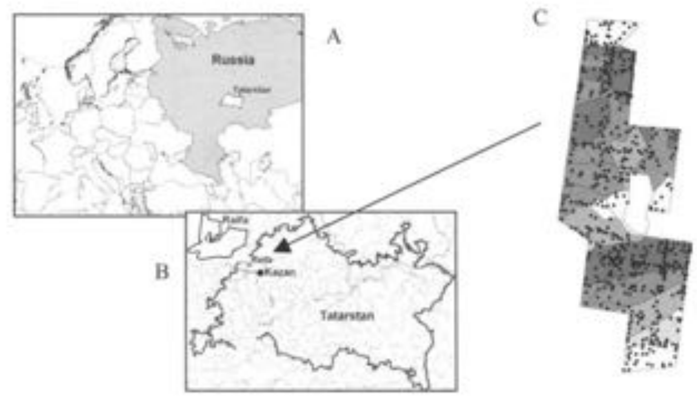
1. Detecting spatial autocorrelation
2. Adjusting for Spatial Autocorrelation with spatial correlation
3. Lagged neighbor Spatial Autoregressive (SAR) models

Spatial modelling of forest community features in the Volzhsko-Kamsky reserve

Spatially explicit forest and environmental data from Russia

Example data from Zuur's books

Example: NDVI in a Boreal System

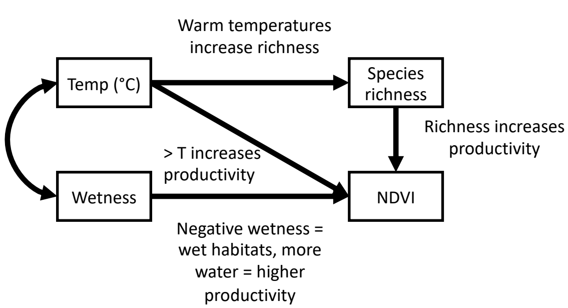


Data Contains Spatial Information

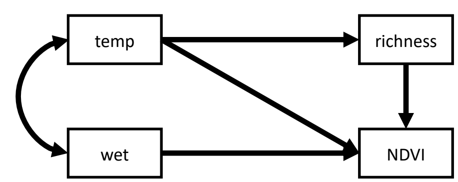
```
> boreal <- read.csv("../Data/boreal.csv")
>
> head(boreal)
  X       x       y richness     NDVI   temp        wet
1 1 2109.70 2093.52       13 0.480180 23.217 -0.0264378
2 2 2190.18 2105.71       21 0.483990 23.217 -0.0234048
3 3 2064.48 2052.77       30 0.489213 23.217 -0.0189264
4 4 2277.34 2103.42       13 0.473226 23.217 -0.0280431
5 5 2347.91 2074.81       13 0.405898 23.635 -0.0292287
6 6 2437.21 2086.95        6 0.424769 23.217 -0.0229209
```

Example: NDVI in a Boreal System

The Model



"Independent"

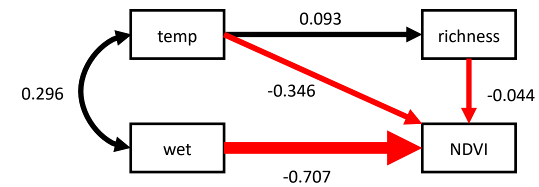



```
rich_lm <- lm(richness ~ temp, data = boreal)
ndvi_lm <- lm(NDVI ~ richness + temp + wet, data=boreal)
boreal.sem <- psem(
  rich_lm,
  ndvi_lm,
  temp %~~% wet,
  data = boreal
)
```

SEM Example. "Independent"

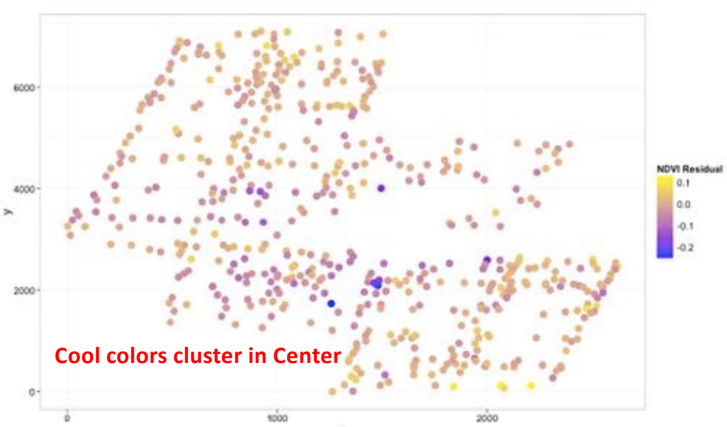
```
---
Tests of directed separation:

        Independ.Claim Estimate Std.Error  DF Crit.Value P.Value
  richness  ~  wet + ...   -35.13   33.7123 530    -1.0421  0.2979

Global goodness-of-fit:

  Fisher's C = 2.422 with P-value = 0.298 and on 2 degrees of freedom

---
Coefficients:

  Response Predictor Estimate Std.Error  DF Crit.Value P.Value Std.Estimate
  richness      temp   1.1707    0.5470 531     2.1402  0.0328       0.0925   *
      NDVI  richness  -0.0004    0.0002 529    -2.0862  0.0374      -0.0440   *
      NDVI      temp  -0.0355    0.0023 529   -15.6564  0.0000      -0.3456 ***
      NDVI       wet  -4.2701    0.1329 529   -32.1406  0.0000      -0.7066 ***
    ~~temp     ~~wet   0.2961        NA 531     7.1436  0.0000       0.2961 ***

  Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
```

What about the residuals?

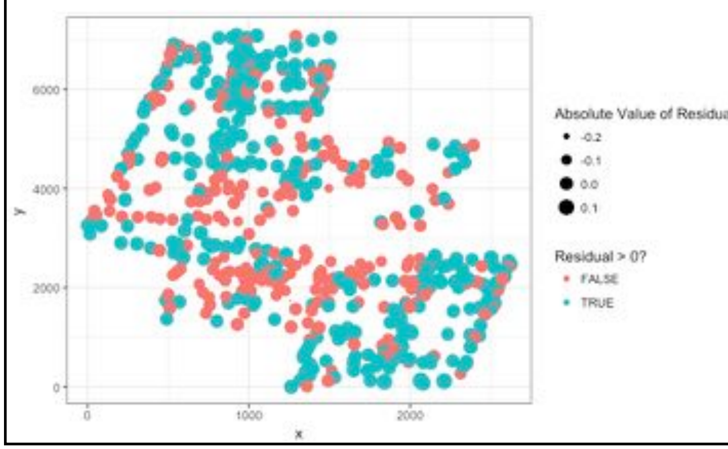



```
source("./residuals.R")

res <- residuals.psem(boreal.sem)
boreal <- cbind(boreal, res)
```

Spatial Correlation in Residuals?



Spatial Correlation in Residuals?



Moran's I: Calculate a Distance Matrix

1. Distance matrices tell us how close points are in space
   - ape library calculates matrix and Moran's I

```
library(ape)
distMat <- as.matrix(dist(
              cbind(boreal$x, boreal$y)))
```

2. We take the inverse, as closer points have greater similarity
   - The diagonal is 0, as there is no correlation within a point

```
distsInv <- 1/distMat
diag(distsInv) <- 0
```

Moran's I with Residuals

```
> mi.ndvi <- Moran.I(boreal$ndvi_residuals, distsInv)
> mi.ndvi
$observed
[1] 0.08014145

$expected
[1] -0.001879699

$sd
[1] 0.003986118

$p.value
[1] 0
```

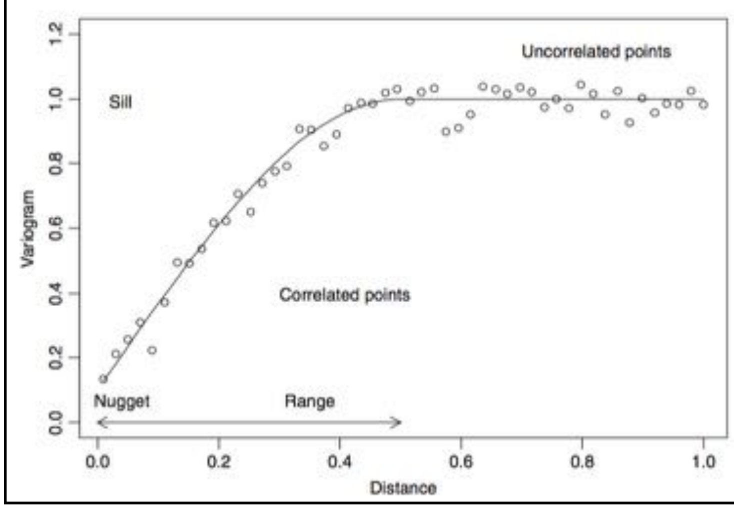
Data is more spatially correlated than expected – need a correction

Moran's I with Residuals

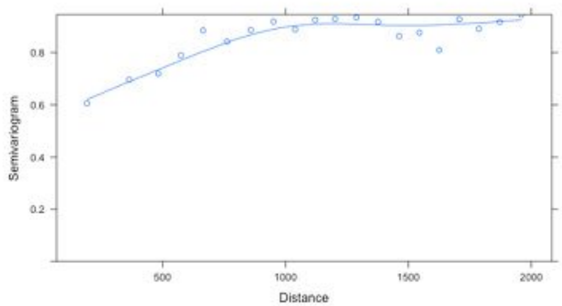
```
> Moran.I(boreal$richness_residuals, distsInv)
$observed
[1] 0.03853411

$expected
[1] -0.001879699

$sd
[1] 0.003998414

$p.value
[1] 0
```

Data is more spatially correlated than expected – need a correction

Variograms to Examine Correlation



Variograms to Examine Correlation

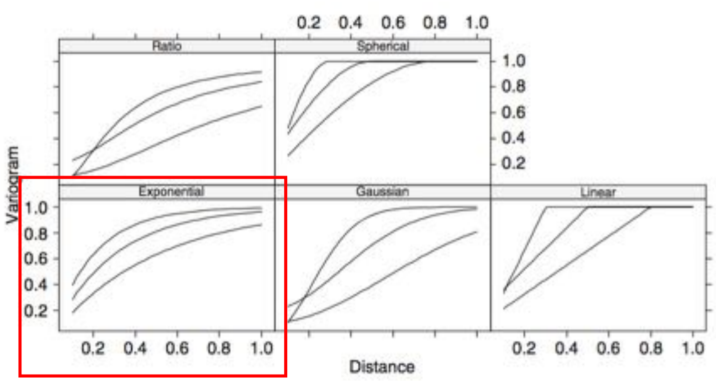



```
library(nlme)
ndvi_gls<- gls(NDVI ~ richness + temp + wet, data=boreal)

plot(Variogram(ndvi_gls, form=~x+y,
               robust=T, maxDist=2000,
               resType="normalized"))
```

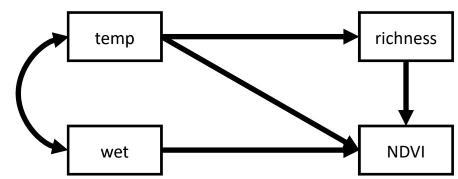
Space: The Final SEM?

1. Detecting spatial autocorrelation
2. Adjusting for Spatial Autocorrelation with spatial correlation
3. Lagged neighbor Spatial Autoregressive (SAR) models

What is the Shape of Our Correlation

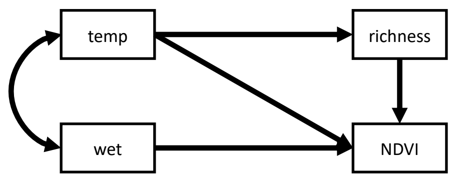


Spatial SEM



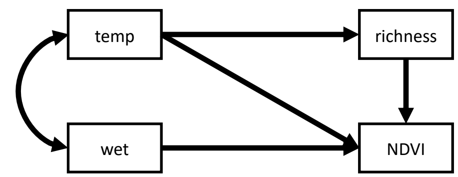

```
#Fit using spatial autocorrelation
spaceCor <- corExp(form =~ x+y, nugget=T)

ndvi_gls_space<- gls(NDVI ~ richness + temp + wet,
                     correlation = spaceCor,
                     data=boreal)

rich_gls_space <- gls(richness ~ temp,
                     correlation = spaceCor,
                     data = boreal)
```

Spatial SEM



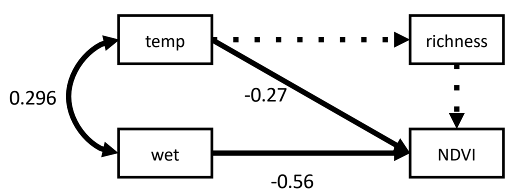

```
boreal_space.sem <- psem(
  ndvi_gls_space,
  rich_gls_space,
  temp %~~% wet,
  data = boreal
)
```

Spatial SEM




```
> dSep(boreal_space.sem)
          Independ.Claim  Estimate Std.Error  DF Crit.Value   P.Value
1 richness  ~  wet + ... -35.20974  42.07507 533 -0.8368315 0.4030645
```

Spatial SEM



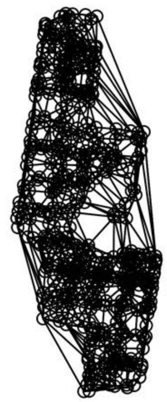

```
> coefs(boreal_space.sem)
  Response Predictor Estimate Std.Error  DF Crit.Value P.Value Std.Estimate
1     NDVI  richness  -0.0002    0.0002 533    -1.6250  0.1047      -0.0301
2     NDVI      temp  -0.0282    0.0033 533    -8.4356  0.0000      -0.2746 ***
3     NDVI       wet  -3.4060    0.1590 533   -21.4266  0.0000      -0.5636 ***
4 richness      temp  -0.0357    0.7765 533    -0.0459  0.9634      -0.0028
5   ~~temp     ~~wet   0.2961        NA 531     7.1436  0.0000       0.2961 ***
```

Space: The Final SEM?

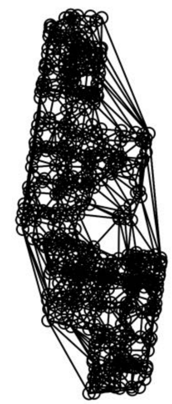
1. Detecting spatial autocorrelation
2. Adjusting for Spatial Autocorrelation with spatial correlation
3. Lagged neighbor Spatial Autoregressive (SAR) models

Nearest Neighbor Spatial weights

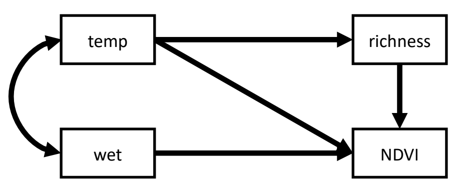
```
library(sp)
library(spdep)

# Determine nearest neighbors
boreal_sp <- boreal

coordinates(boreal_sp) <- ~x+y

#get neighbors
#many functions if using
#regions instead of points
nb <- tri2nb(boreal_sp)

plot(nb, coordinates(boreal_sp))
```



The Region

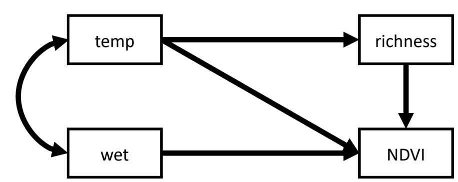


SAR Approach




```
# Run regression models (with spatial weights)
spatial_weights <- nb2listw(nb)

rich_lag <- lagsarlm(richness ~ temp,
                     data = boreal_sp,
                     listw = spatial_weights,
                     tol.solve = 1e-11)

ndvi_lag <- lagsarlm(NDVI ~ richness + temp + wet,
                     data = boreal_sp,
                     listw = spatial_weights,
                     tol.solve = 1e-11)
```

SAR Approach

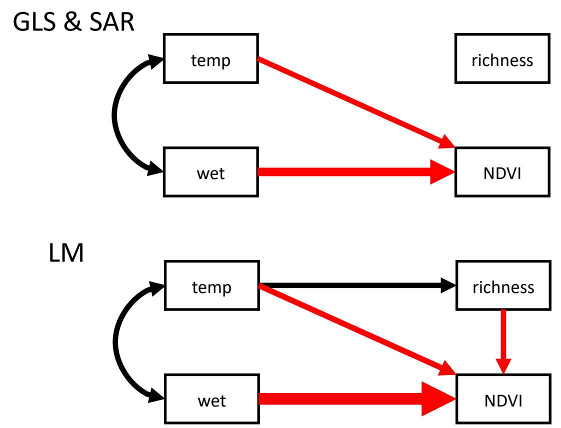



```
boreal_space_lag.sem <- psem(

  rich_lag,

  ndvi_lag,

  temp %~~% wet,

  data = boreal_sp

)
```

SAR Approach

```
---
Tests of directed separation:

         Independ.Claim Estimate Std.Error DF Crit.Value P.Value
  richness  ~  wet + ... -25.6601   31.2746 NA    -0.8205  0.4119

Global goodness-of-fit:

  Fisher's C = 1.774 with P-value = 0.412 and on 2 degrees of freedom

---
Coefficients:

  Response Predictor Estimate Std.Error  DF Crit.Value P.Value Std.Estimate
  richness      temp   0.6808    0.5094  NA     1.3364  0.1814       0.0538
      NDVI  richness  -0.0003    0.0001  NA    -1.8619  0.0626      -0.0338
      NDVI      temp  -0.0207    0.0023  NA    -9.1514  0.0000      -0.2016 ***
      NDVI       wet  -3.1033    0.1519  NA   -20.4340  0.0000      -0.5135 ***
    ~~temp     ~~wet   0.2961        NA 531     7.1436  0.0000       0.2961 ***

  Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
```

SAR Approach



Why not to think about autocorrelation

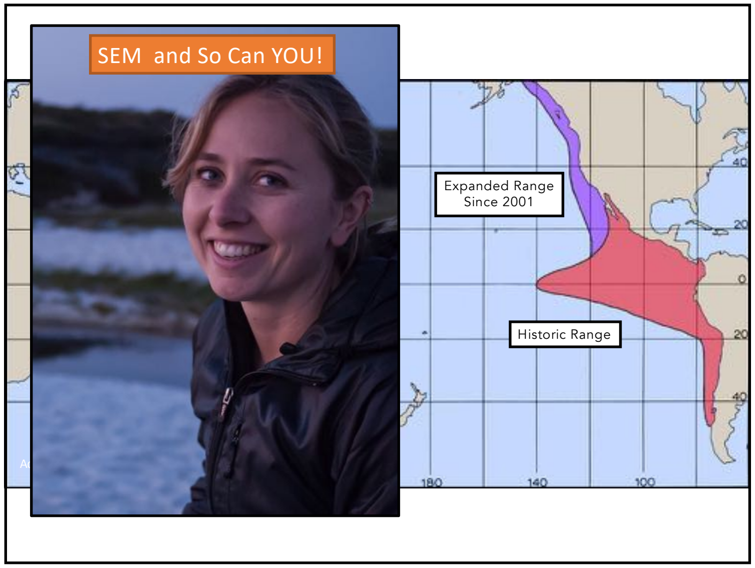
There are two key issues regarding space:

(1) Are their things to learn about the other factors that could explain variations in the data that vary spatially?

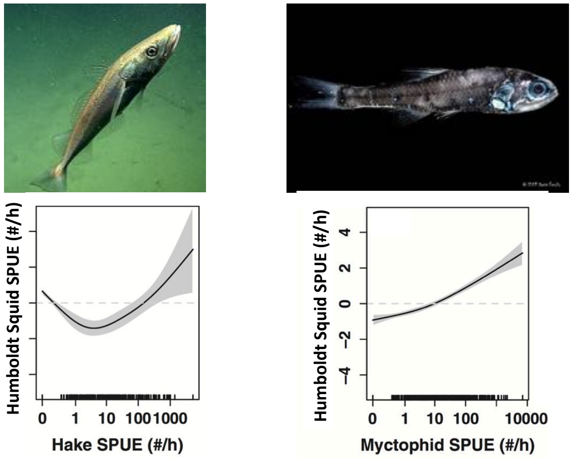
(2) Do we have nonindependence in our residuals?

<u>Recent reference on the subject</u>:
Hawkins, BA (2011) Eight (and a half) deadly sins of spatial analysis. Journal of Biogeography. doi:10.1111/j.1365-2699.2011.02637.x

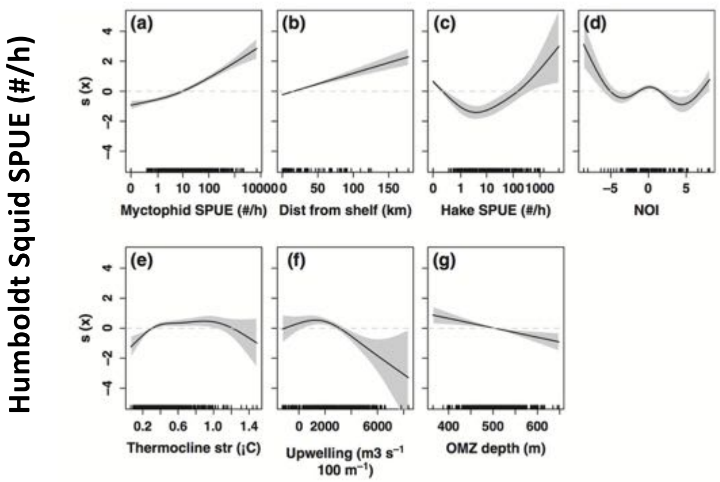
Stewart, J.S., Hazen, E.L., Bograd, S.J., Byrnes, J.E.K., Foley, D.G., Gilly, W.F., Robison, B.H., Field, J.C., 2014. Combined climate- and prey-mediated range expansion of Humboldt squid (*Dosidicus gigas*), a large marine predator in the California Current System. Global Change Biology



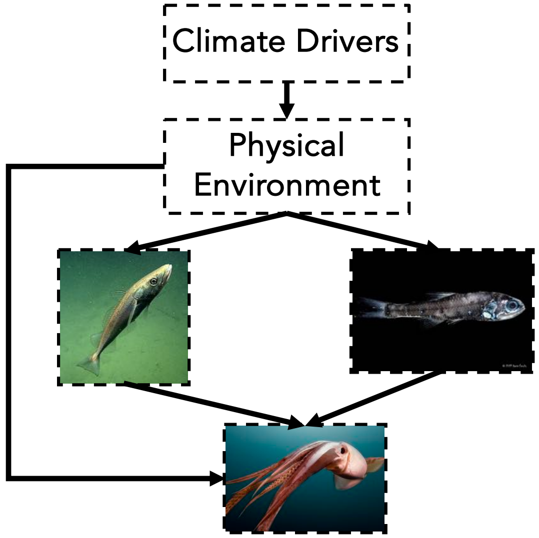
Driven by Prey?



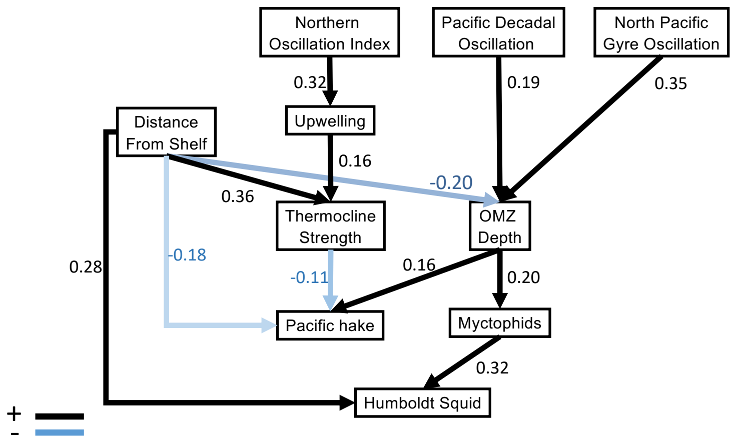
Stewart et al. 2014 GCB



Stewart et al. 2014 GCB



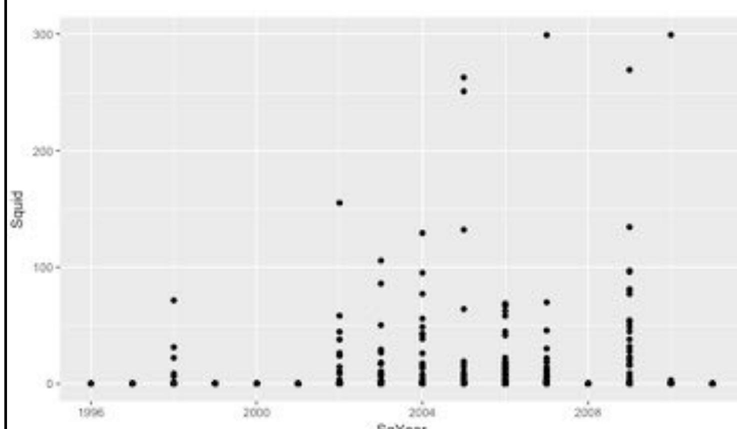
Both Indirect Effects and Spurious Correlations with Prey



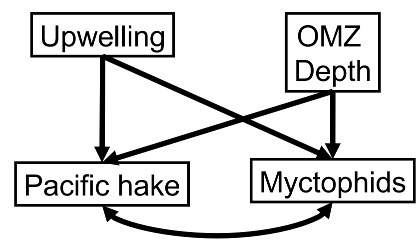
The Data

```
> squid <- read.csv("../Data/squid_stewart.csv")

> head(squid)
  UniDiveNum   Squid DepthOMZ Temp25m UIwin  NOI   DayL Strat   LatN  Hake Loligo    Myct
1       19.3  0.0000  445.085 10.2400  3077 1.11 13.369    21 36.712 0.000      0   5.368
2       20.3 21.7539  439.390  9.8035  3417 1.11 13.395    24 36.701 0.000      0  28.475
3       21.3  8.7137  450.880 10.0570  3640 1.11 13.427    33 36.642 0.000      0   2.739
4       22.3  0.0000  423.740  9.6890  3768 1.11 13.470    20 36.791 0.181      0   2.709
5       23.3  0.0000  421.960 10.1820  3897 1.11 13.503    37 36.779 0.000      0   0.000
6       24.3  0.0000  492.380 10.1480  3925 1.11 13.532    56 36.701 0.000      0 543.158
  SqMonth SqYear Distkm ROV_ID   chla tcline   tcstren     tcval    tcoxy ocline    ocval
1       5   2009  5.229      3 1.5889     22 0.3874163 10.378564 5.247444   34.5 4.584194
2       5   2009  2.581      3 1.7582     22 0.5015197 10.230897 4.763088   38.5 3.948747
3       5   2009  7.450      3 1.5723     31 0.2943885  9.695900 4.646011   27.5 4.443057
4       5   2009  0.136      3 1.0145     33 0.2439991  9.322725 3.369191   35.0 3.445972
5       5   2009  1.550      3 1.1726     34 0.2833411  9.598154 3.923525   35.0 3.832471
6       5   2009  2.395      3 1.0132     52 0.1746201  9.456308 3.952009   56.0 3.679309
```

The Data



Exercise: Use Space, Time, or Both



squid <- read.csv("../Data/goodsquid.csv")

1. Fit this model! If only using space, might want to filter.
2. Is there temporal or spatial autocorrelation (`SqYear, LatN`?)
3. Attempt to handle any correlation using either random effects, spatial, or temporal autocorrelation.