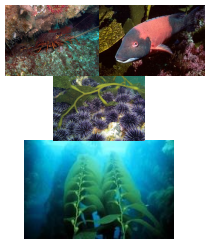
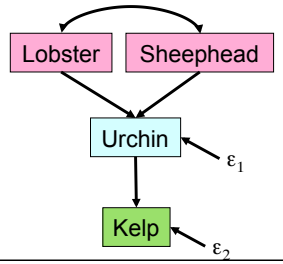
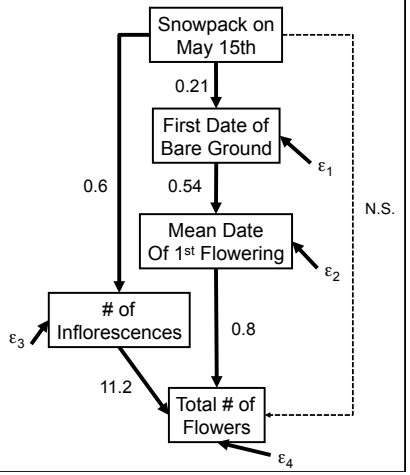
Structural Equation Modeling for Ecology & Evolutionary Biology

Jarrett E. K. Byrnes
University of Massachusetts Boston
http://byrneslab.net/teaching/sem





Climate Change in a Systems Context



Inouye et al 2002 Oecologia

Course Goals

1. Give you a working familiarity with SEM
2. Decide when SEM is right for you
3. Understand the process of model creation, evaluation, and revision
4. Be able to implement SEM in R

Schedule for the Week

M – Introduction to SEM
Model Building
Piecewise Fitting

T – SEM with Likelihood
Model Comparison

W – Multigroup Models
Latent Variables

Th – Composite Variables
Advanced Topics

F – Advanced Topics & Open Lab

S – Project Presentations

Typical Day

9:00 – 10:30 Lecture/Lab I

10:30 – 10:45 Break!

10:45 – 12:00 Lecture/Lab II

12:00 – 1:00 Lunch

1:00 – 2:30 Lecture/Lab III

2:30 – 3:00 Break!

3:00 – 4:30 Lecture

4:30 – 5:30 Work with Your Data

Where You can Learn More about SEM

Grace (2006) Structural Equation Modeling and Natural Systems. Cambridge Univ. Press.

Shipley (2000) Cause and Correlation in Biology. Cambridge Univ. Press.

Kline (2012) Principles and Practice of Structural Equation Modeling. (3rd Edition) Guilford Press.

Bollen (1989) Structural Equations with Latent Variables. John Wiley and Sons.

Hoyle (2012) Handbook of Structural Equation Modeling. Guilford Press.

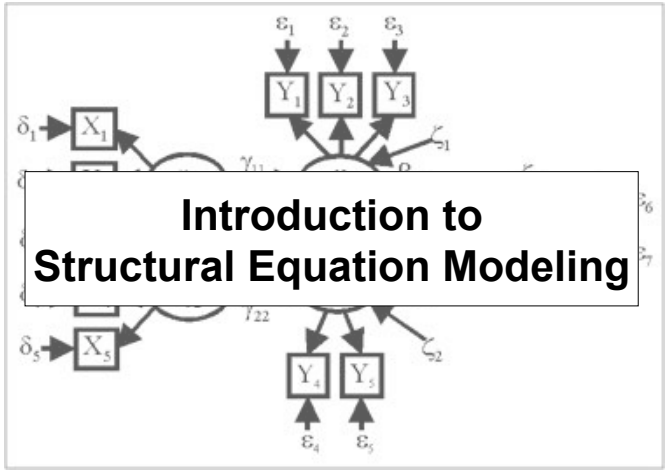
also, see www.structuralequations.org

Software for SEM

1. R with lavaan, piecewiseSEM, sem, or openMX libraries – flexible, can solve models piecewise or using covariance analysis. Many options.
2. AMOS – most user friendly, but, point and click
3. LISREL – original software. Still being updated with many advanced features
4. EQS – competitior to LISREL, has REQS package
5. MPLUS – favorite of advaned users, but, black-boxes many processes
6. WinBUGS, JAGS, or OpenBUGS – VERY flexible. VERY complex. Time to get your Bayes on!

Who am I?

Who are you? Why are you here?

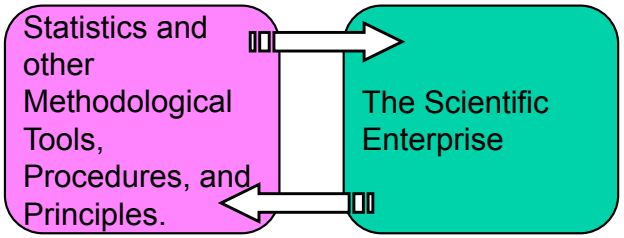


Introduction to Structural Equation Modeling

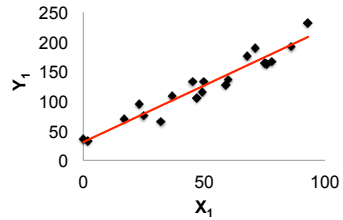
Introduction Outline

1. What is SEM?
2. History!
3. From ANOVA to SEM
4. SEM as Part of a Research Program

The Scientific Enterprise is Influenced by our Statistical Methodology

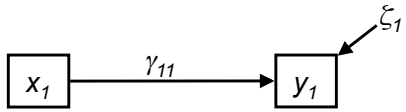


SEM is a form of Graphical Modeling



equation form $y_1 = \gamma_{11}x_1 + \zeta_1$

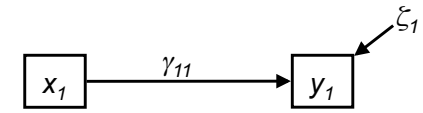
graphical form



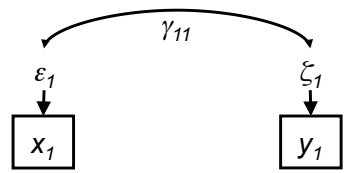
The Structure in SEM implies CAUSALITY

equation form $y_1 = \gamma_{11}x_1 + \zeta_1$

graphical form

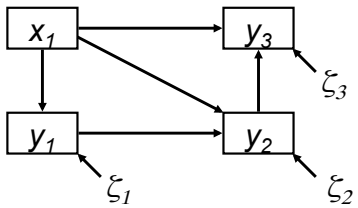


graphical form without causality



SEM: The use of two or more structural equations to evaluate direct and indirect effects in a system

Hypothesis involving indirect effects

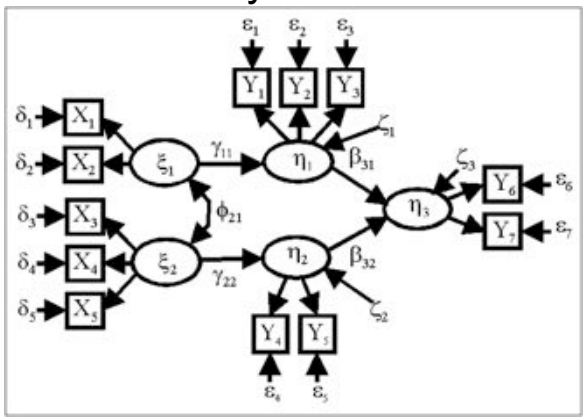


Corresponding Equations

e.g.

$$y_1 = \gamma_{11}x_1 + \zeta_1$$
$$y_2 = \beta_{21}y_1 + \gamma_{21}x_1 + \zeta_2$$
$$y_3 = \beta_{32}y_2 + \gamma_{31}x_1 + \zeta_3$$

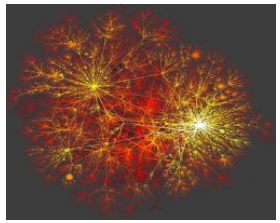
Simple Idea to Attack Complex Systems



SEM is a Framework

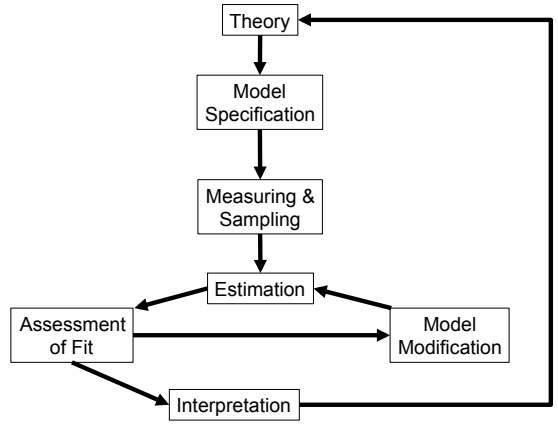
We use statistical and mathematical tools

within the SEM framework



to build scientific understanding about the multiple processes operating in systems

SEM as a Unifying Process

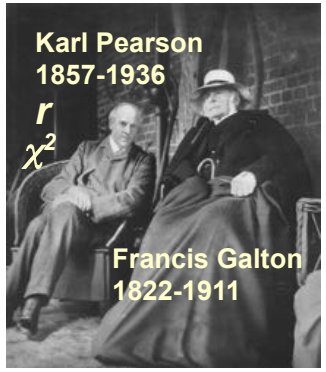


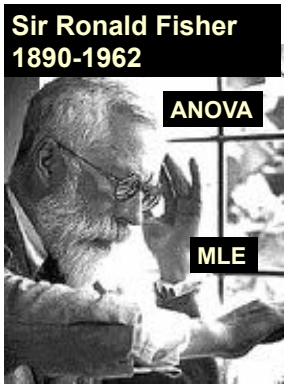
After Grace & Bollen 2005

Introduction Outline

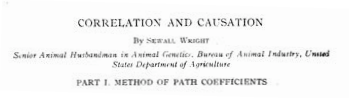
1. What is SEM?
2. History!
3. From ANOVA to SEM
4. SEM as Part of a Research Program

Fit, Correlation, and Testing Models





Sewall Wright 1921 & 1st Gen SEM

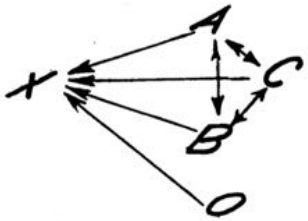


CORRELATION AND CAUSATION

By Sewall Wright

Senior Animal Husbandman in Animal Genetics, Bureau of Animal Industry, United States Department of Agriculture

PART I. METHOD OF PATH COEFFICIENTS



The First Path Diagram?

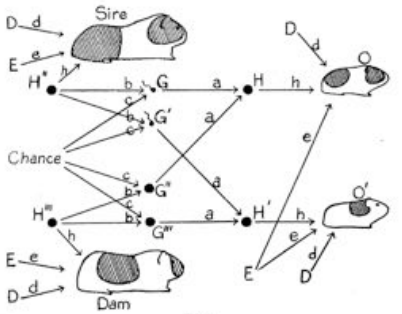


FIG. 3.

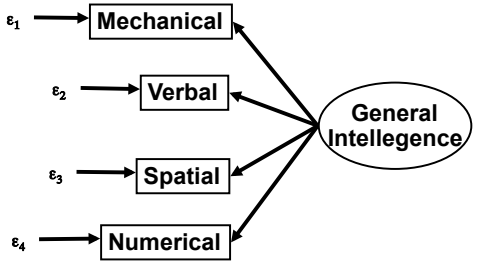
Diagram illustrating the casual relations between litter mates (O, O′) and between each of them and their parents. H, H′, H″, H‴ represent the genetic constitutions of the four individuals, G, G′, G″, and G‴ that of four germ cells. E represents such environmental factors as are common to litter mates. D represents other factors, largely ontogenetic irregularity. The small letters stand for the various path coefficients.

Wright 1920 PNAS

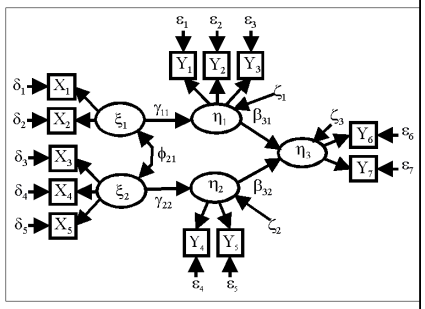
A Parallel Tradition: Spearman & Factor Analysis in 1904

"GENERAL INTELLIGENCE," OBJECTIVELY DETERMINED AND MEASURED.

By C. Spearman.



Jöreskog & 2nd Generation SEM

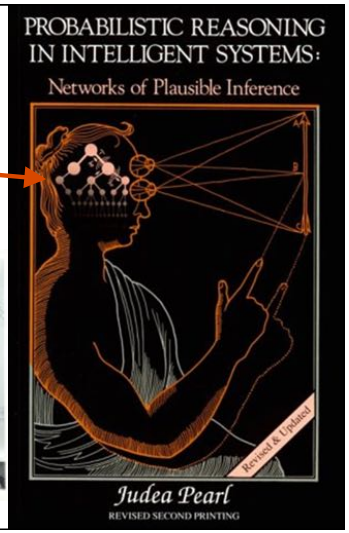


1. Model fit using covariance matrix of the data
2. Estimation of parameters via Maximum Likelihood
3. Can assess and compare fit of a multivariate model

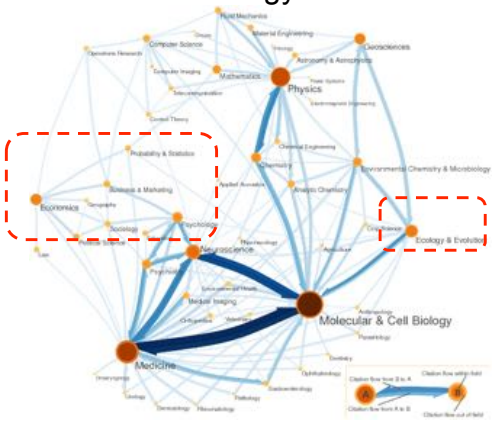
Judea Pearl and 3rd Geration SEM

The networks of the mind represent our causal thinking about systems.

1. SEM with a graph theoretic framework
2. Causality is central
3. Methodological flexibility via piecewise approaches

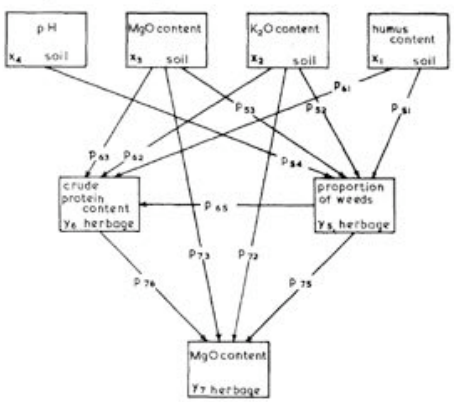


Why hasn't SEM received more attention in ecology?

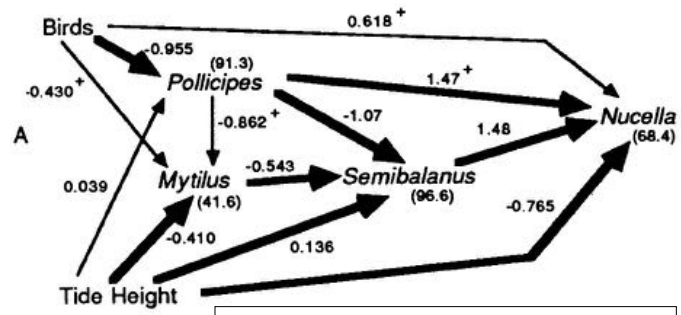


eigenfactor.org

Path Analysis in Ecology: Ferrari 1963



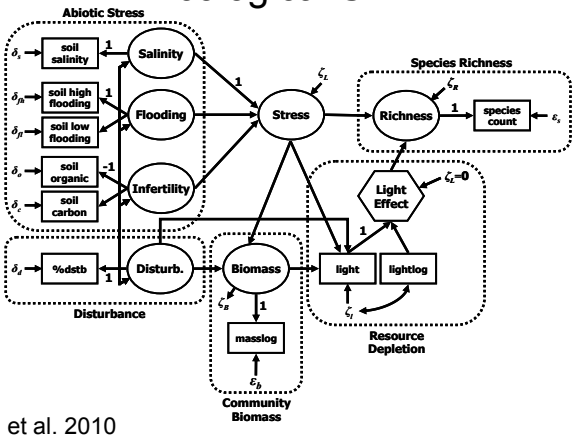
Path Analysis and Food Webs



1) Coupling of observational data and causal manipulations
2) Evaluation of multiple hypotheses

Wootton, 1994 Ecology

Grace, Pugesek and Modern Ecological SEM



Grace et al. 2010

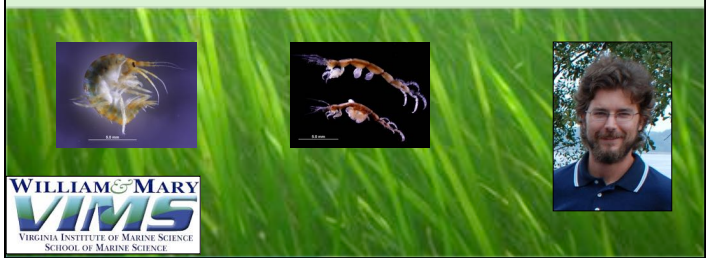
Introduction Outline

1. What is SEM?
2. History!
3. From ANOVA to SEM
4. SEM as Part of a Research Program

Field-based Evidence for the Importance of Small Herbivores in a Seagrass Ecosystem:

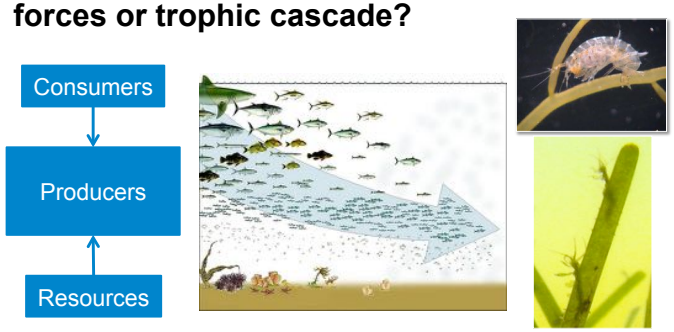
An Examination Using Structural Equation Modeling

Matthew A. Whalen[1], J. Emmett Duffy[1], James B. Grace[2]
1 - Virgina Institute of Marine Science
2 - USGS



Introduction/Questions

Are seagrasses controlled by bottom-up forces or trophic cascade?



Subtext: Is nutrient runoff or overfishing causing seagrass declines?

Field Experiment

1. Investigate proposed food-web interactions
2. Test the relative impact of top-down and bottom-up forces

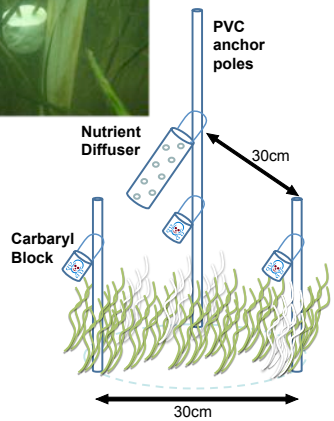
Manipulation: Nutrients X Grazers

Location: Cuba Island
- nearly monospecific eelgrass bed
- constant depth
- large enough for experiment

Duration: Summer 2009 for 6 weeks

Experiment being replicated around the world! ZEN!

Experimental Reduction of Small Herbivores



Experimental Design:

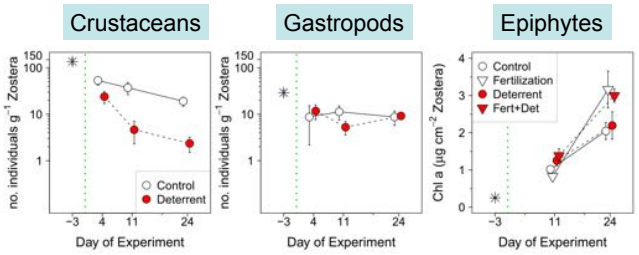
Treatments:
- **pesticide to reduce crustacean grazers**
- **nutrient addition**
- **combination**
- **controls**

8 reps @ 5 trts = 40 plots

Basic Results:

Crustaceans: reduced 58-96%
Algal biomass: increased 130-748%
Nutrients: inconsistent effects

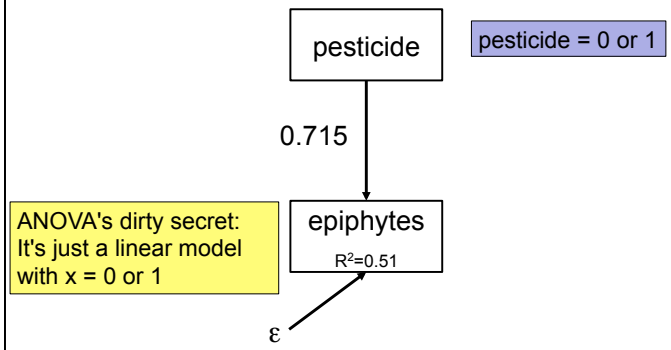
Repeated Measures ANOVA Effects Plot



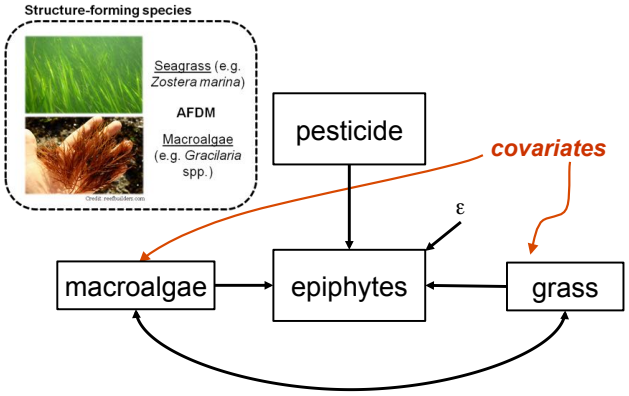
Death By F-Table

	Response														
	Crustacean Mesograzers			Gastropod Mesograzers			Micro-algal Epiphytes			Eelgrass Growth			Eelgrass Shoot Density		
Source	df	F	P	df	F	P	df	F	P	df	F	P	df	F	P
Deterrence (Det)	1	114.28	**<0.001**	1	1.67	0.200	1	50.30	**<0.001**	1	0.02	0.899	1	6.68	**0.015**
Fertilization (Fert)	1	0.02	0.902	1	0.07	0.793	1	0.87	0.354	1	0.00	0.995	1	1.24	0.275
Sampling Date	2	4.66	**0.012**	2	20.20	**<.0001**	2	1.76	0.178	3	32.98	**<0.001**			
Det × Fert	1	0.13	0.720	1	2.13	0.148	1	0.17	0.677	1	0.63	0.432	1	0.10	0.751
Det × Date	2	7.29	**0.001**	2	2.45	0.093	2	10.52	**<0.001**	3	0.70	0.555			
Fert × Date	2	1.49	0.230	2	1.07	0.349	2	1.08	0.344	3	0.27	0.845			
Det × Fert × Date	2	0.76	0.470	2	0.29	0.747	2	1.36	0.263	3	2.32	0.083			
Residual	84			84			82			65			28		

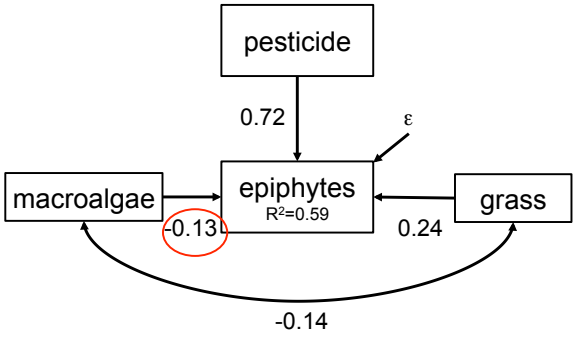
Graphical Illustration of ANOVA for Epiphyte Response



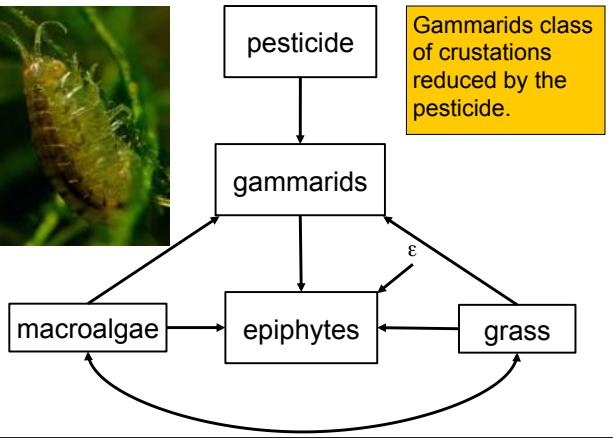
ANCOVA with Macroalgae and Seagrass as Covariates



ANCOVA Results

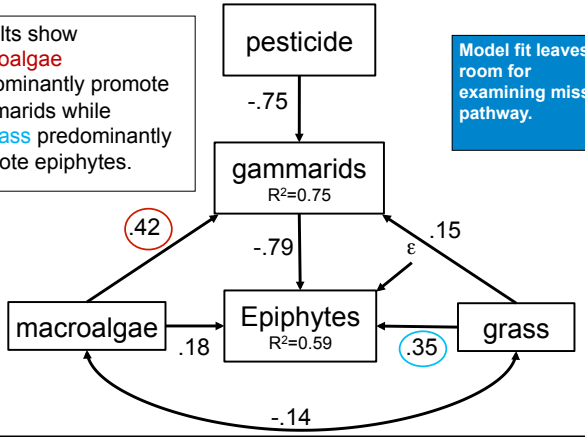


Mediation in SEM



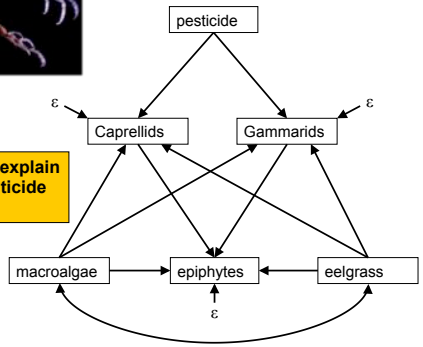
Mediation in SEM

Results show macroalgae predominantly promote gammarids while eelgrass predominantly promote epiphytes.



Model fit leaves room for examining missing pathway.

Two Mediator Model



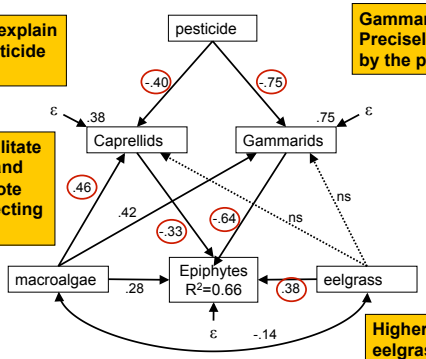
Do Caprellids explain remaining pesticide effect?

Final Model Results

Caprellids do explain remaining pesticide effect.

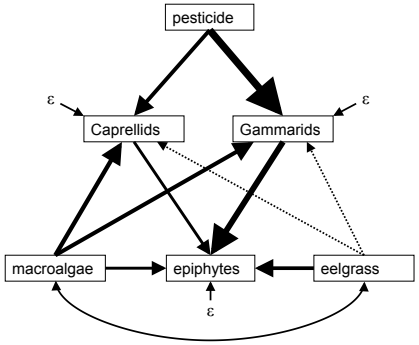
Gammarids more Precisely controlled by the pesticide.

Macroalgae facilitate all amphipods and indirectly promote herbivory, protecting eelgrass.



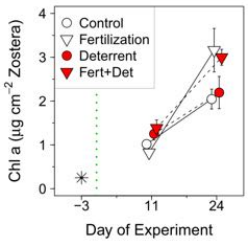
Higher density of eelgrass associated with greater density of epiphytes.

We can represent coefficients with line thickness…

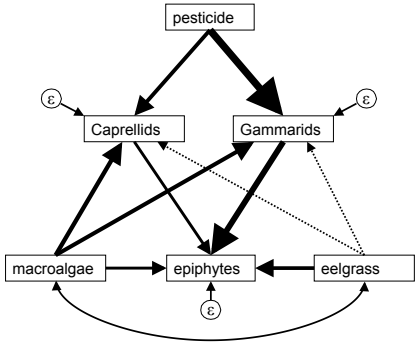


From ANOVAs to SEM

Our model results imply that behind this summary of mean responses…

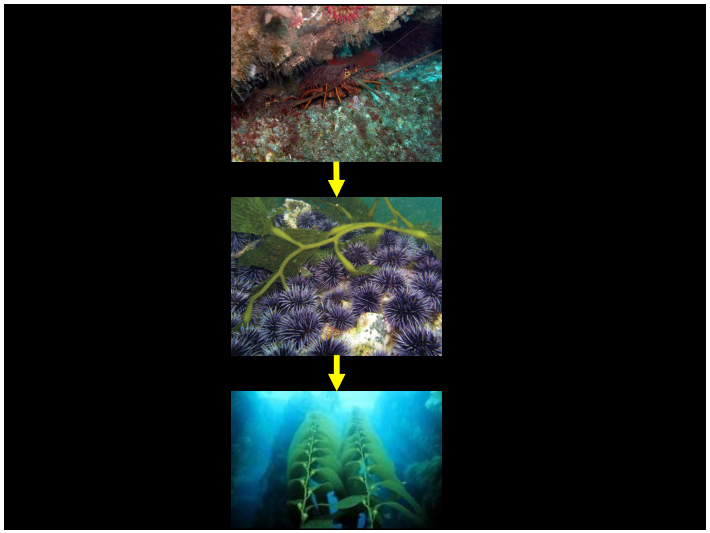


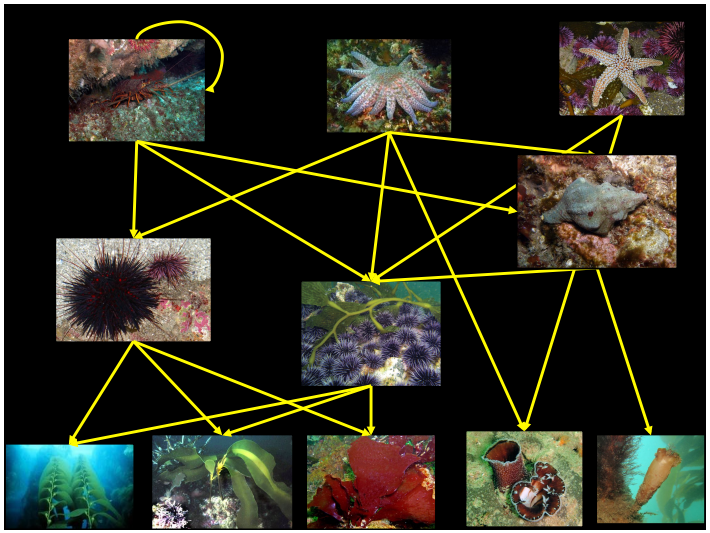
…is a network of effects like this.

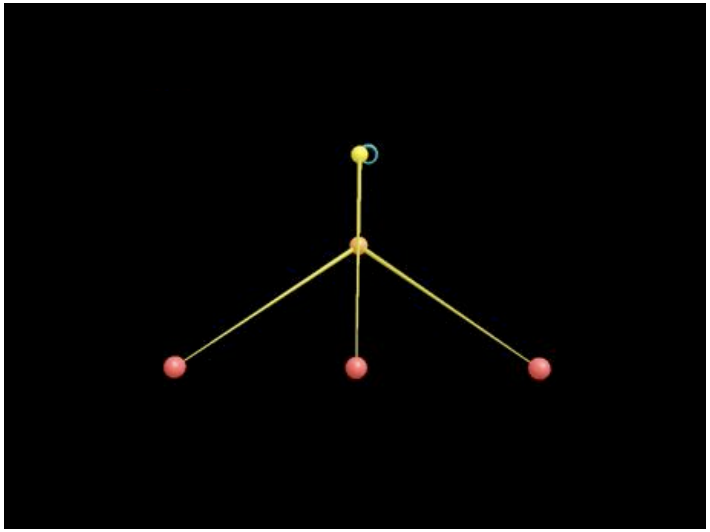


Introduction Outline

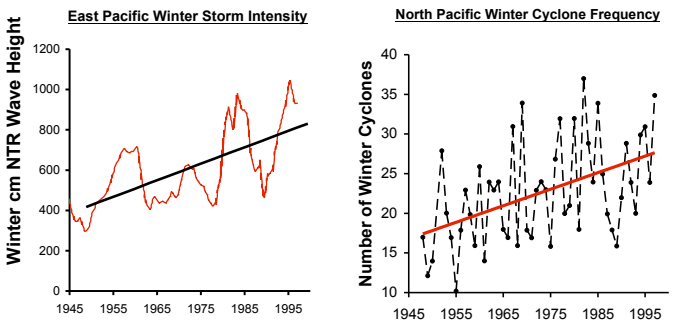
1. What is SEM?
2. History!
3. From ANOVA to SEM
4. SEM as Part of a Research Program







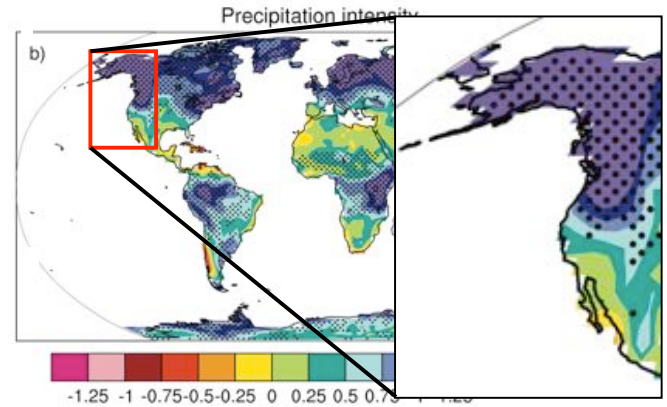
Storm Intensity and Frequency Increased over the Last 50 Years



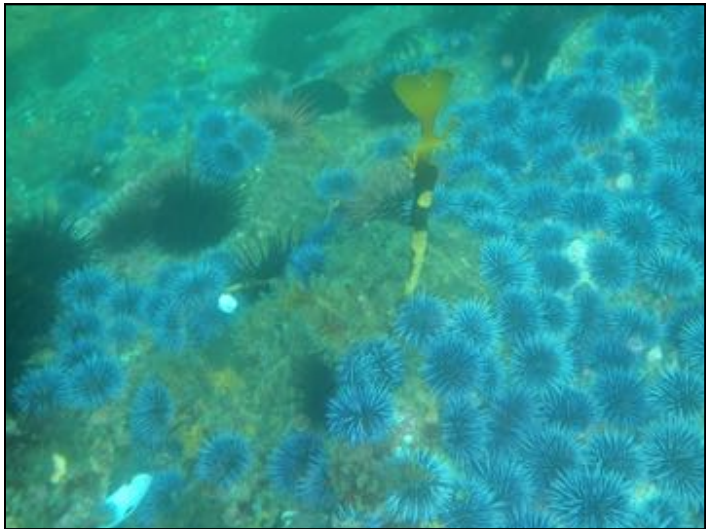
Bromirski *et al*. 2003

Graham and Diaz 2001

Climate Change Prediction: The Largest Storm of the Year Will Get Stronger



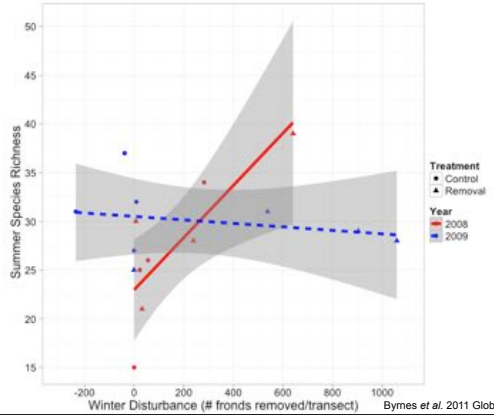
IPCC AR 4



Testing Causality: Repeated Kelp Disturbance Experiment

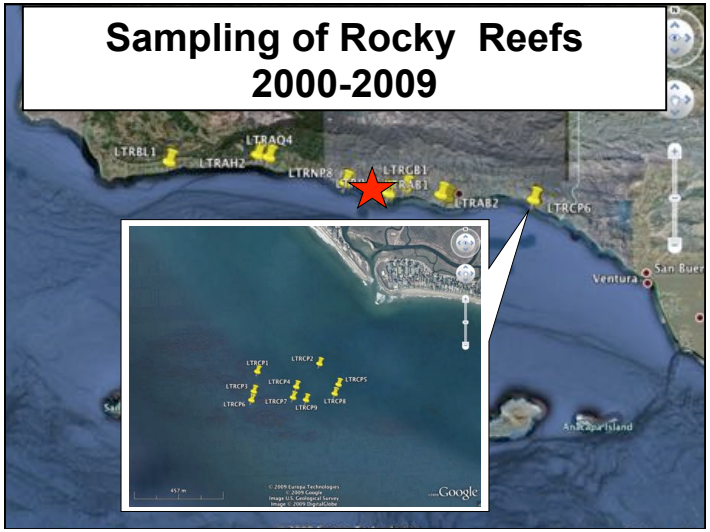
- 4 Reefs selected in 2008 with paired 40x40m areas
- Giant Kelp removed in experimental plots every January to simulate disturbance

Effect of Kelp Removal on Richness Vary With Repeated Removals

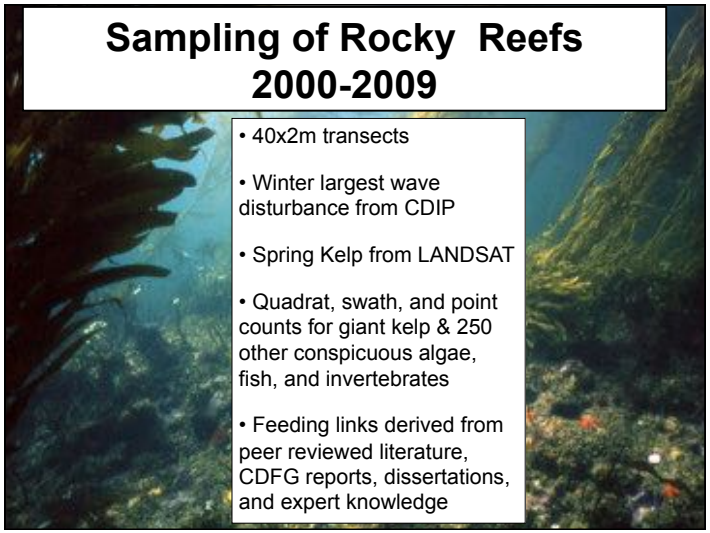


Byrnes *et al.* 2011 Global Change Biology

Sampling of Rocky Reefs 2000-2009

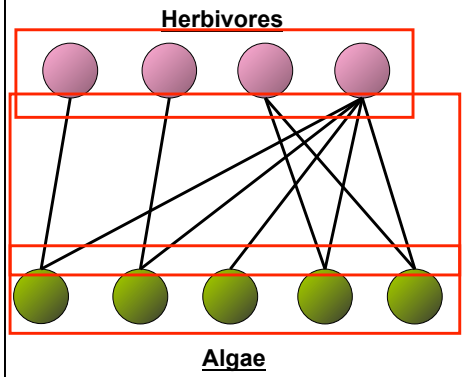


Sampling of Rocky Reefs 2000-2009



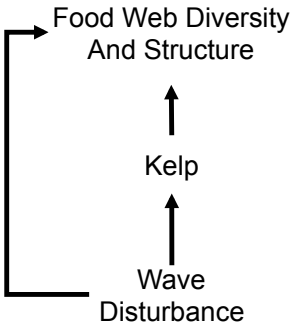
- 40x2m transects
- Winter largest wave disturbance from CDIP
- Spring Kelp from LANDSAT
- Quadrat, swath, and point counts for giant kelp & 250 other conspicuous algae, fish, and invertebrates
- Feeding links derived from peer reviewed literature, CDFG reports, dissertations, and expert knowledge

Quantifying Food Web Structure

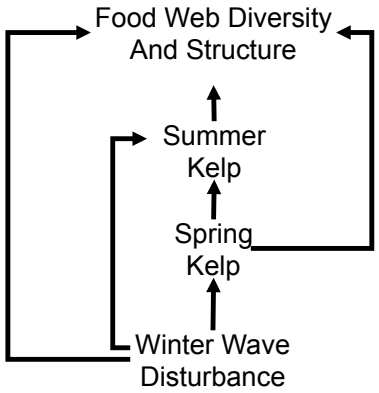


- Species Richness - by functional group
- Linkage Density
- etc…

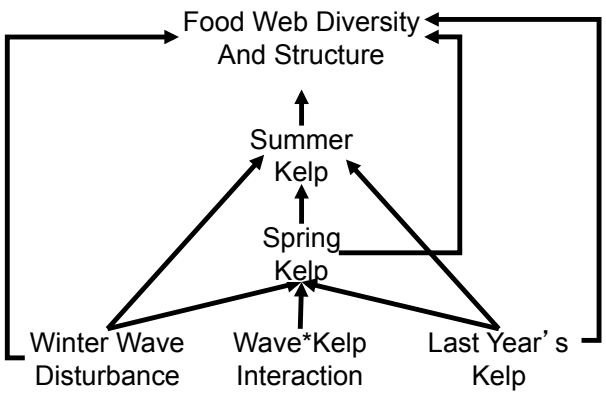
Direct and Indirect Effects of Waves on Food Webs: Simple, Right?



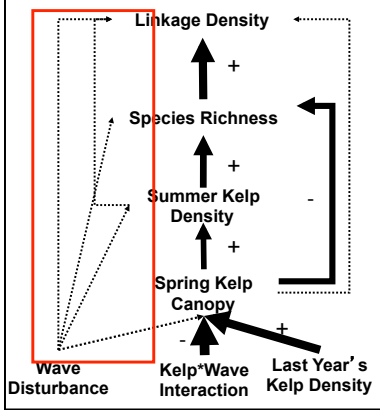
Things Get More Complex…



The Full Model



Wave Disturbance Indirectly Related to Food Web Structure



4) More species = more feeding links per species

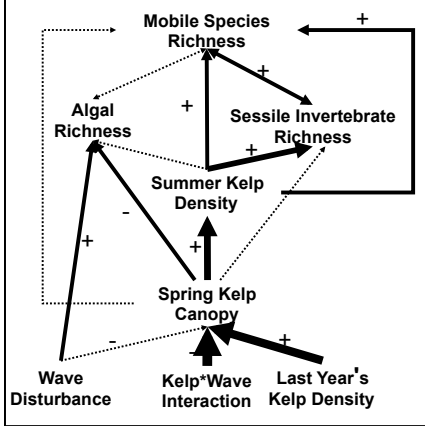
3) Kelp density increases richness, but spring canopy decreases richness

2) More kelp in the spring = more kelp in the summer

1) Big waves remove kelp (where there is kelp)

Byrnes *et al.* 2011 Global Change Biology

Effect of Waves Differs for Algae v. Animals



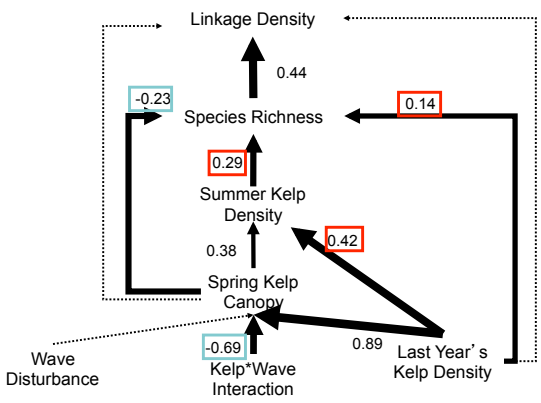
1) Waves DECREASE sessile invert and mobile species richness via kelp removal

2) Waves INCREASE algal richness by altering light availability

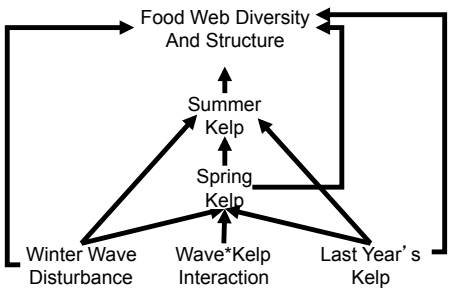
3) Increase in algal richness due to waves outweighs decrease in animal richness

Byrnes *et al.* 2011 Global Change Biology

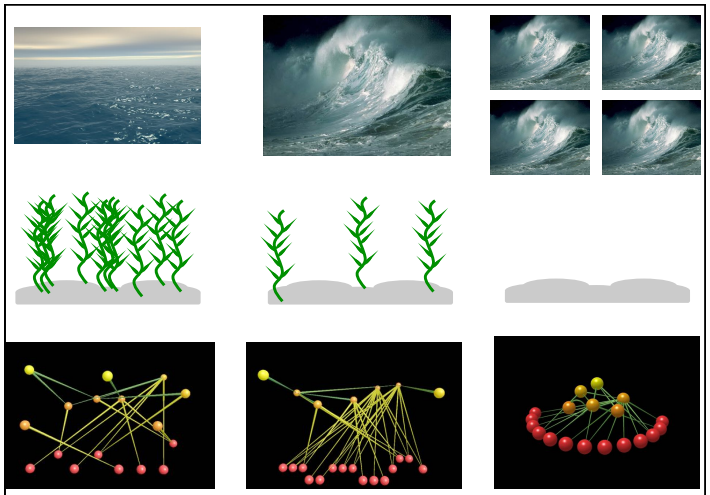
Last Year's Kelp Influences This Year's Food Web Structure



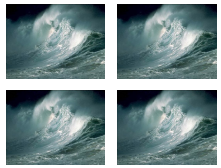
SEM and Simulations



Spring Kelp = Waves + Last Year's Kelp + Waves*Last Year's Kelp
Summer Kelp = Waves + Last Year's Kelp + Spring Kelp
Richness = Waves + Last Year's Kelp + Spring Kelp + Summer Kelp
Etc…..



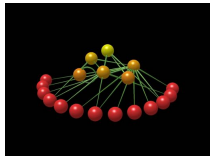
Baseline Forest **One Storm** **Many Storms**



Climate Change May Simplify Kelp Forest Food Webs

1. Free resources from wave disturbance may initially promote diversity and complexity

2. Loss of foundation species leads to simplified food webs



Many Storms

QUESTIONS & COFFEE