Structural Equation Modeling in Agronomy and Soil Science

Eric G. Lamb
Dept. of Plant Sciences
University of Saskatchewan

Structural Equation Modeling (SEM)

- Path analysis developed in 1920s by Sewell Wright
- Rejected in favour of an experimentalist approach on philosophical grounds by Pearson and Fisher
- Modern form developed by social scientists in 1970s (LISREL synthesis)
- Since 2000 becoming increasingly used in fields such as ecology

Why Structural Equation Modeling?

- Standard methods (e.g. ANCOVA, multiple regression) work well for:
  - complex manipulative experiments
  - studies where the net effect of a treatment is of interest
- These methods do not perform well for:
  - Observational studies
  - Datasets with intercorrelated response variables
  - Where the goal is to study biological mechanisms leading to an outcome not the net outcome

Correlation Matrix

	Graz	Bio	Rich
Graz	1.0	---	---
Bio	-.5	1.0	---
Rich	0.0	-.6	1.0

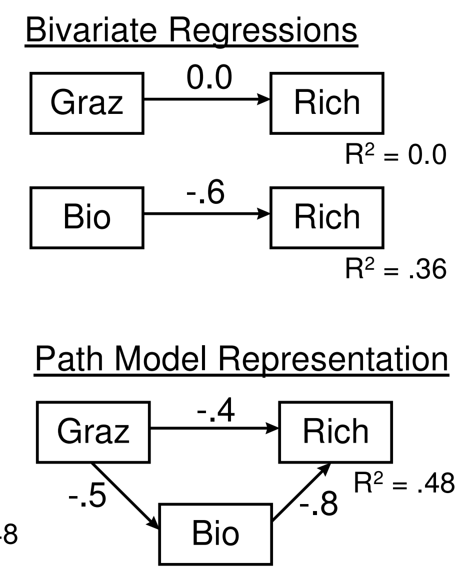
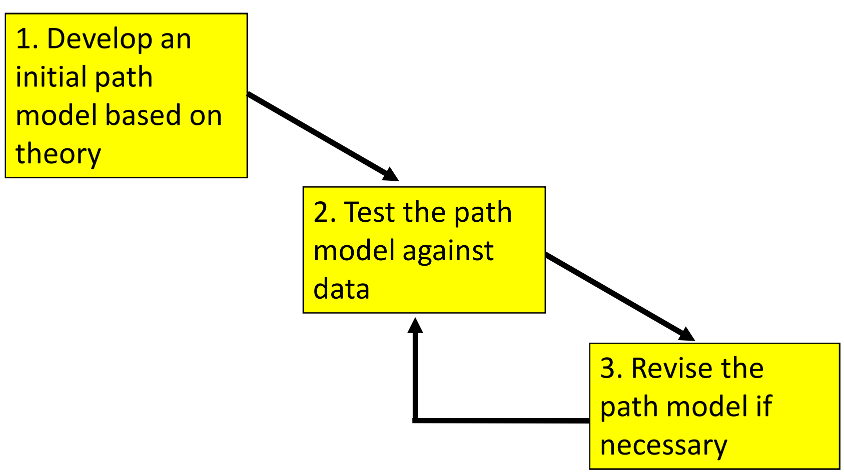


Figure 30.3; McCune and Grace, 2002. Analysis of Ecological Communities. MjM Software Design

How SEM works

- Scientist specifies paths (causal relationships between variables)
- Paths imply a structure to the covariance matrix
- Implied covariance structure can be tested against actual structure
- Agreement between implied and actual structures validates the causal relationships represented by the paths
- Coefficients describing strength and direction of paths can be calculated

Process of SEM Modeling

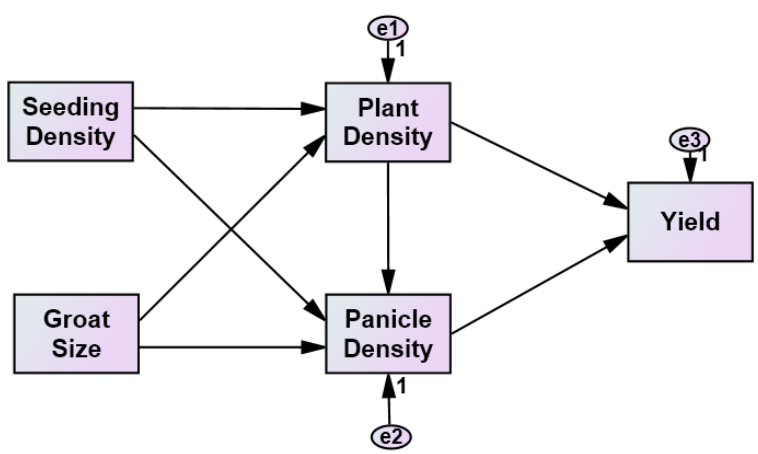


A walk through an SEM

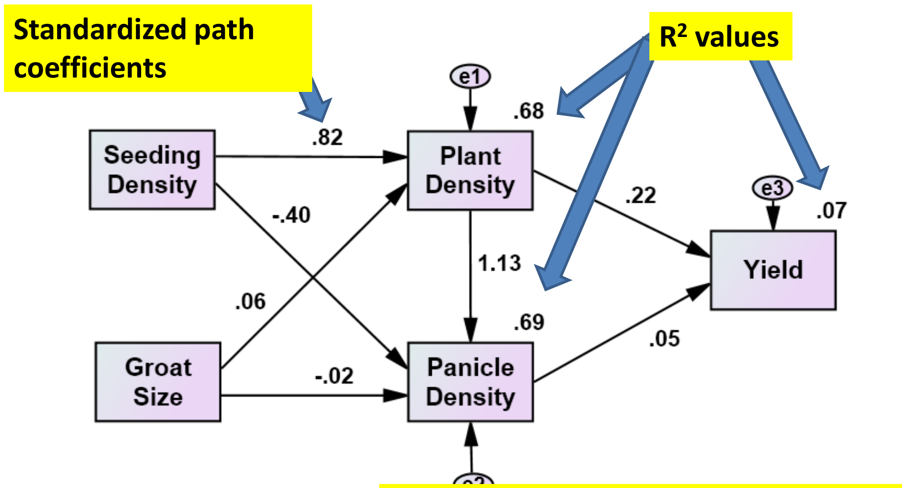
- Field trial examining the effects of seeding density and groat size on oat yield
  - Blocked experimental design
  - 2 sites and 2 years
  - Experimental treatments
    - Seedling density (target densities 50-400 plants /m^{-2})
    - Groat size (18-30)
  - Measurements
    - Plant density
    - Panicle density
    - Yield

Data set kindly provided by Steve Shirtliffe, Dept. of Plant Sciences, University of Saskatchewan

Initial Model



Fitted Model



X^2_3=1.182, p=0.757

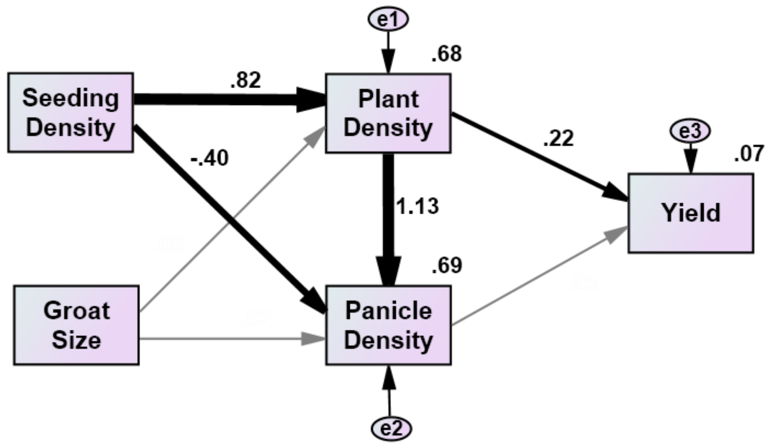
Unstandardized path coefficients are in the units of the original variables and can be used in predictive models

Standardized path coefficients allow easy comparison of the relative importance of paths in a model

Pathway	Unstand. Estimate	S.E.	C.R.	P	Standard. Estimate
Target Density--->Plant Density	0.668	0.035	19.306	<0.001	0.82
Groat Size--->Plant Density	1.313	0.942	1.393	0.164	0.059
Target Density--->Panicle Density	-0.321	0.059	-5.432	<0.001	-0.399
Groat Size--->Panicle Density	-0.415	0.921	-0.45	0.652	-0.019
Plant Density--->Panicle Density	1.11	0.073	15.275	<0.001	1.125
Panicle Density--->Yield	0.403	0.963	0.419	0.675	0.05
Plant Density--->Yield	1.742	0.95	1.834	0.067	0.219

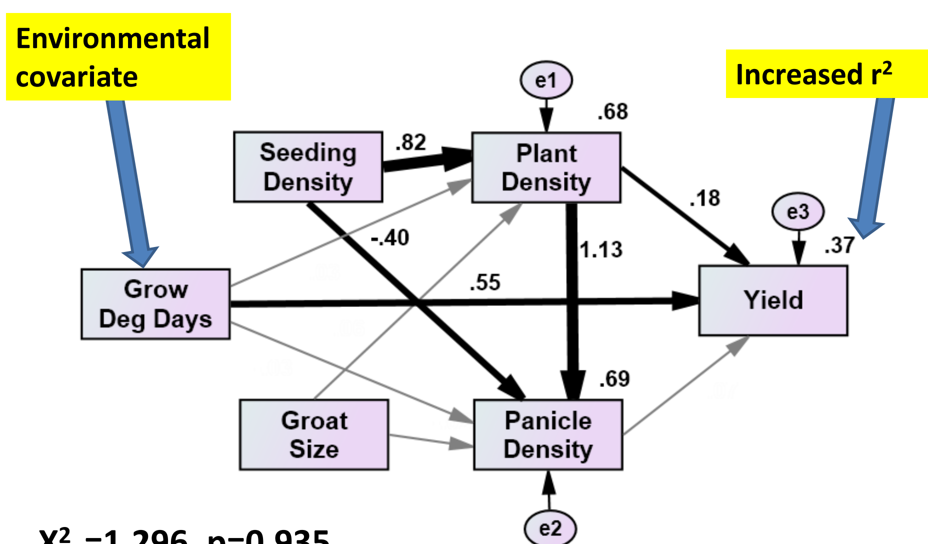
Test of path significance (whether the unstandardized estimate is different from zero)

Fitted Model



X^2_3=1.159, p=0.763

More complex model accounting for site-year effects

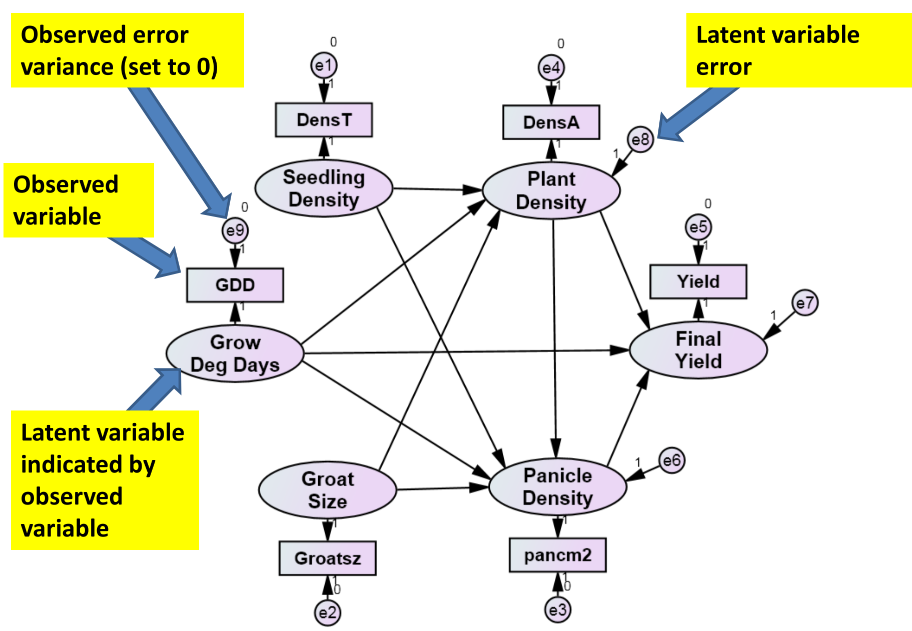


X^2_5=1.296, p=0.935

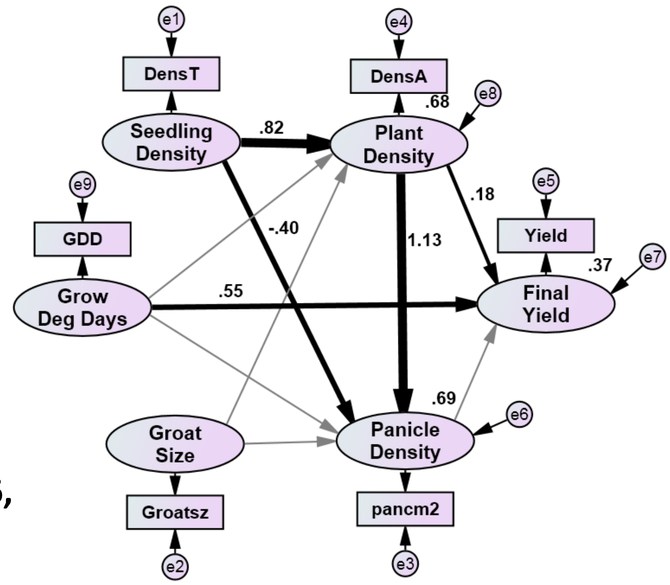
Observed vs. Latent Variable Models

- Observed variable models
  - Measured variable closely matches conceptual variable
  - Assumes little/no measurement error
- Latent variable models
  - variable of conceptual interest not measured but correlated variables measured
  - Measurement error known to be present

Latent variable model equivalent to observed variable model

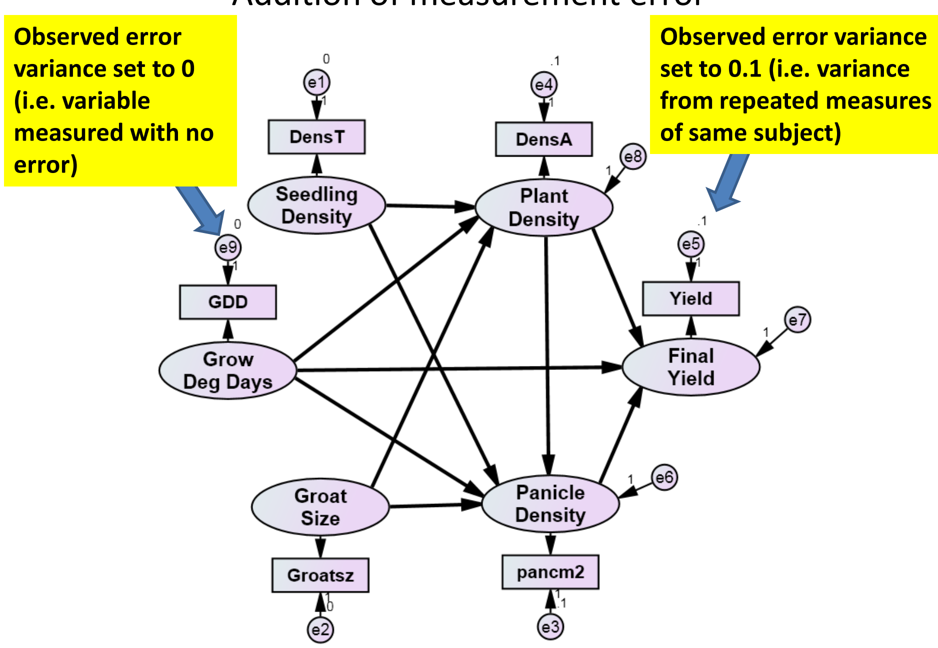


Latent variable model equivalent to observed variable model

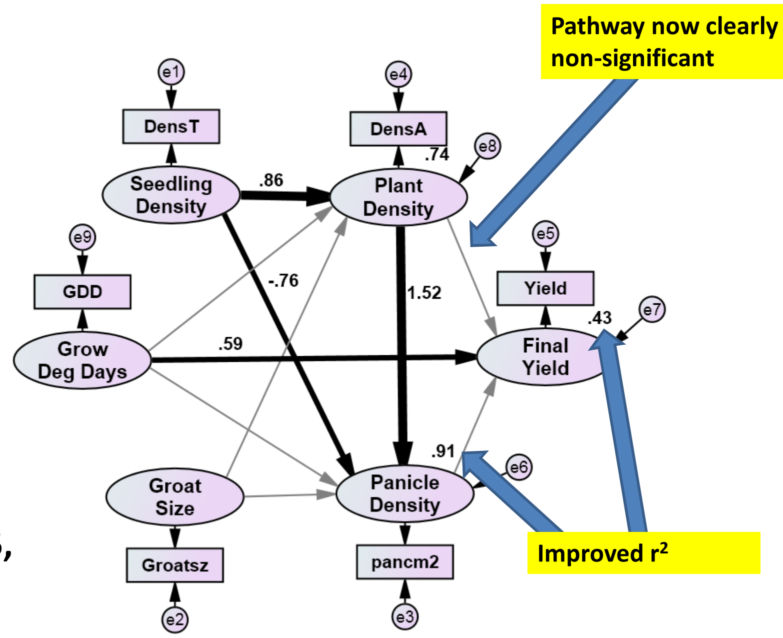


X^2_5=1.296,
p=0.935

Addition of measurement error

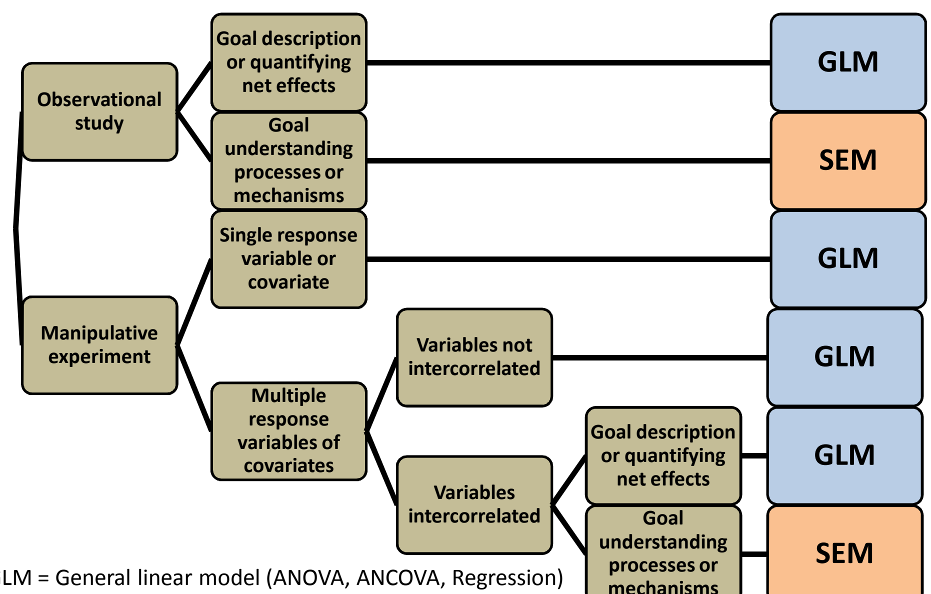


Fitted model with measurement error



X^2_5=1.605, p=0.901

When should SEM be applied?



GLM = General linear model (ANOVA, ANCOVA, Regression)
SEM = Structural Equation Modeling

Structural equation modeling: a search for generality*

- Scientific advancement and theory development enhanced by:
  - Consideration of measurement error and correspondence between theory and measurement
  - Models flexible and easily expanded
  - Flexible stats to meet the requirements of research not the reverse
  - Capacity for multivariate predictive models

*Based on Grace (2006), chapter 12

Resources: SEM Software

- SPSS: AMOS*
  - Best software for users new to SEM
  - Very capable software with superb graphical interface
  - free student version: http://www.amosdevelopment.com/
- R package: SEM library
  - Limited capability but full version freely available
  - http://www.r-project.org/
- M-Plus
  - powerful dedicated SEM software for complex modeling
  - Owners get extremely good technical support
- SAS: proc CALIS

*Amos input files and associated data files for the models in this presentation are available at: http://homepage.usask.ca/~egl388/Lamb_Biostats_Links.html

Resources: websites

- Example AMOS input files and data from this presentation
  - http://homepage.usask.ca/~egl388/Lamb_Biostats_Links.html
- Further examples and tutorials
  - http://www.structuralequations.com/index.html
  - http://www.jamesbgrace.com/
- Listserv archives for SEM technical discussion
  - http://bama.ua.edu/archives/semnet.html

Resources: Further reading

•**Grace, J.B. (2006)** ***Structural equation modeling and natural systems. Cambridge University Press, U.K.***

•Very readable practical introduction to SEM

•**McCune, B. & Grace, J.B. (2002)** ***Analysis of ecological communities. MjM Software Design, Gleneden Beach, Oregon.***

•Chapter introducing the basics of SEM

•**Shipley, B. (2000)** ***Cause and correlation in biology. Cambridge University Press, U.K.***

•Introduction to SEM inc. detailed discussion of causality and model development

•**Grace, J.B. et al. (2010) On the specification of structural equation models for ecological systems.** ***Ecological Monographs, 80, 67-87.***

•SEM specification for testing theoretical constructs

•**Grace, J.B. & Bollen, K.A. (2005) Interpreting the results from multiple regression and structural equation models.** ***Bulletin of the Ecological Society of America, 283-295.***

•Brief introduction to SEM including model interpretation

•**Guillen-Portal, F.R. et al. (2006) Compensatory mechanisms associated with the effect of spring wheat seed size on wild oat competition.** ***Crop Sci, 46, 935-945.***

•Example of a current application of SEM in crop science

•**Kline, R.B. (2005)** ***Principles and practice of structural equation modeling, 2nd edition. Guilford Press, New York, New York.***

•Basic SEM text

•**Grace, J. & Bollen, K. (2008) Representing general theoretical concepts in structural equation models: the role of composite variables.** ***Environmental and Ecological Statistics, 15, 191-213.***

•Introduction to more complex SEM methods particularly useful in the natural sciences