# Sensitivity Analysis

EESS 211


# Steps to Modeling

Untested Hypothesis

- Goal Definition
- Conceptualization
- Model Type/Scale Decisions
- Coding
- Calibration ⟵ Data
- Sensitivity Analysis
- "Validation" ⟵ Data
- Application ⟵ Data

Partially Tested Hypothesis

# Sensitivity Analysis:

- Sensitivity Analysis (SA) is the study of how the variation in the output of a model (numerical or otherwise) can be apportioned to different sources of variation in the model input.

# Outline

1) Goals of SA
2) Properties of an ideal SA method
3) Steps for SA
4) Simple Screening
5) Morris' Method
6) Variance based SA measures
7) Examples of other approaches

# Goals of SA

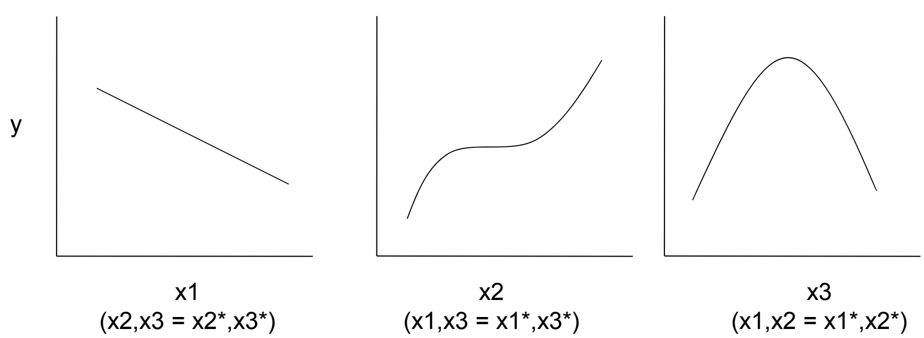
- Find the factors most in need of calibration, and others that can be fixed (factor prioritization)
- Identify factors or equations that are unimportant (factor fixing or “mechanism reduction”)
- Evaluate parameter combinations that can achieve a pre-determined level of variance (variance cutting in risk analysis)
- In general, SA is to help evaluate whether model performance is sufficient and where to invest effort in improving performance

# Choosing an SA technique

Consider: $\hat{\mathrm{Y}} = \hat{\mathrm{F}}(\mathrm{X}, \hat{\beta})$

Where x = (x1,x2,x3)

How would you describe the sensitivity of Y?



## Properties of an ideal SA technique

1. Ability to account for shape and scale of pdf of $x_i$
2. Ability to average effect of $x_i$ for varying $x_{-i}$ (multidimensional average)
3. Model independent (should work regardless of linearity or additivity of model)
4. Ability to treat grouped factors as single factors
5. Quantitative (able to not only order different factors, but say how much more important one is than another)

## Steps for SA

1. Establish goals (e.g., what specific output are you interested in?)
2. Decide which input factors to analyze
3. Choose distribution for each factor
4. Choose SA technique
5. Generate input sample
6. Evaluate model for input sample
7. Analyze model outputs with SA method

## Simple Screening

- Run for hi/low value of each factor (one-at-a-time)
- Total model runs = 2 x K (K = number of factors)
- Usually results in a barplot, tornado plot, etc.
- Problems with this method?

## Simple Screening

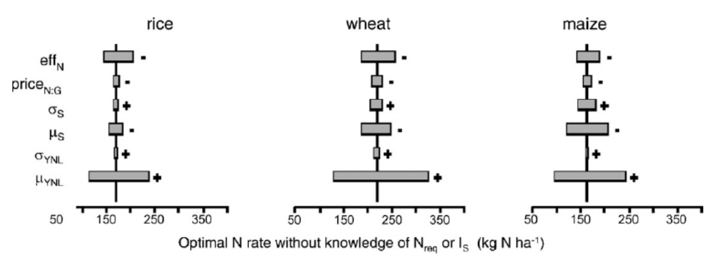



Fig. 4. The sensitivity to model parameters of simulated optimal N rates. The solid vertical line shows the optimal N rate for the baseline parameter values. The bars show the range in computed optimal rates when adjusting each parameter by ±20% of its baseline value. A plus sign indicates that the higher rate was computed when using the higher parameter value (baseline + 20%), while a negative sign indicates that the higher rate was associated with the lower parameter value.

## Properties of an ideal SA technique

1. Ability to account for shape and scale of pdf of $x_i$
2. Ability to average effect of $x_i$ for varying $x_{-i}$ (multidimensional average)
3. Model independent (should work regardless of linearity or additivity of model)
4. Ability to treat grouped factors as single factors
5. Quantitative (able to not only order different factors, but say how much more important one is than another)

## Morris' Method

- Tries to deal more effectively with nonlinearity and interactions
- Can think of as individually randomized O-A-T experiments

# Morris' Method

Define an Elementary Effect

$d_i\,(x) = [y(x_1,x_2,\ldots,x_{i-1,}x_i+\Delta,x_{i+1},\ldots x_k) - y(x)]\,/\Delta$

What is $\Delta$?

Define p = # of levels to consider for each factor (e.g. 4)

$\Delta$ is some multiple of 1/(p-1), e.g. 1/3 or 2/3

# Morris' Method

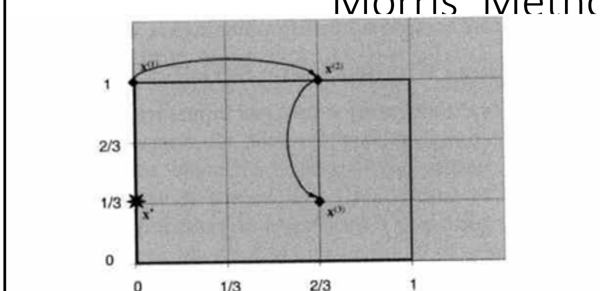



Figure 4.3 An example of trajectory in the two-dimensional space.

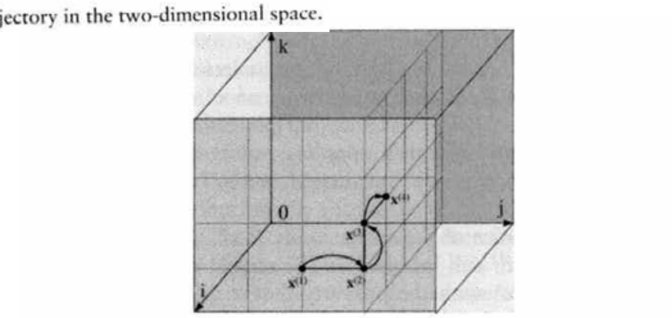



Figure 4.2 An example of trajectory in the input factor space when $k = 3$

## Morris' Method

1. Start by generating an initial vector $x^*$
2. Then randomly increase 1 or more $x_i$ by $\Delta \rightarrow x^{(1)}$
3. Then randomly increase 1 or more $x_i$ by $\Delta \rightarrow x^{(2)}$ so that we have $x^{(1)}$ and $x^{(2)}$ that differ in 1 factor
4. Repeat step 3 k times to compute k elementary effects (can think of this as a trajectory in sample space)
5. Repeat steps 1-4 r times to average for different trajectories

Total # runs = (k+1)r

## Morris' Method

- We can use average or average absolute value, of $d_i$ 's to measure factor importance, and standard deviation to estimate nonlinearity of effect
- Useful for identifying few important factors with limited computations.

## Properties of an ideal SA technique

1. Ability to account for shape and scale of pdf of $x_i$
2. Ability to average effect of $x_i$ for varying $x_{-i}$ (multidimensional average)
3. Model independent (should work regardless of linearity or additivity of model)
4. Ability to treat grouped factors as single factors
5. Quantitative (able to not only order different factors, but say how much more important one is than another)

## Variance Based Measures

$y=F(x_1,x_2,\ldots x_k)$

Total uncertainty in y is measured as Var(y)

How much would Var(y) go down if I could tell you a certain factor, $x_i$ ? i.e. what is $\mathrm{Var}(Y|X_i)$

If it is sensitive to xi, it should go down by
(A) a lot
(B) a little
(C) either A or B

# Morris' Method Example

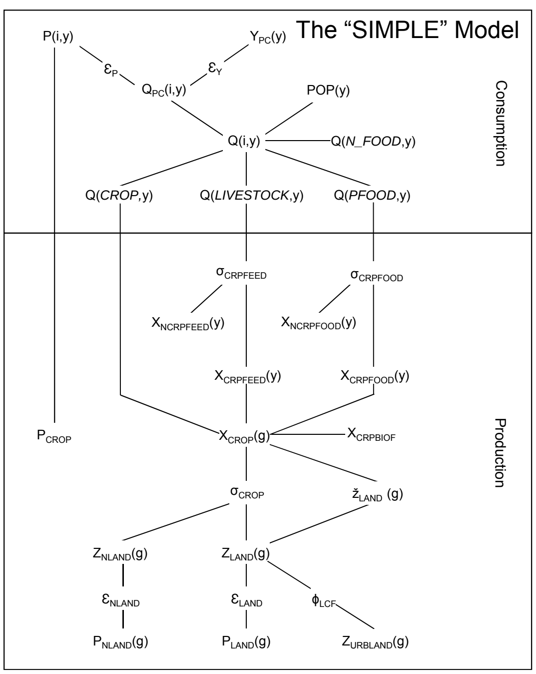
Question: how much carbon is saved (and at what cost) by investing in agricultural productivity?

Quick facts:

- •converting 1 ha of land to agriculture releases about 300 metric tons of $CO_2$
- •current global cropland area about 1.5 billion ha (roughly 1 ha per 5 people)
- •demand for crop production will roughly double by 2050 (depending on price)

•To evaluate this question, need a model that treats a few key links:

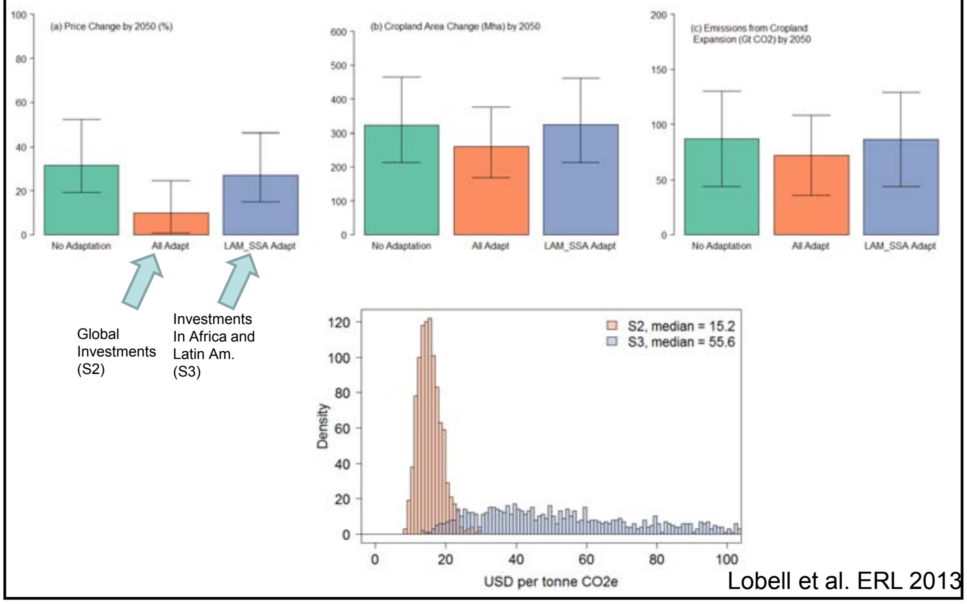
- •Investments → technology / yields
- •Yields → land conversion rates (incl. feedbacks from demand)
- •Land conversion → land use emissions




## 15 key parameters

| Parameters | Description | Mean | Max | Min | Sources of data used in computing model parameters |
|---|---|---|---|---|---|
| ELANDW | Elasticity of land supply wrt land lents | 0.11 | 0.14 | 0.08 | Ahmed et al. (37) and Gurgel et al (38) |
| ENLANDW | Elasticity of non-land supply wrt non-land rents | 1.34 | 1.74 | 0.94 | Keeney & Hertel(39) |
| LCFURBW | Conversion factor of urban land to cropland | 0.23 | 0.30 | 0.16 | USDA (40), Liu &Tian (41), Weber & Puissant (42) |
| NCRP2CRPW | Emission factors (MgCO2eq/ha) non-crop to crop conversion | 286.97 | 373.06 | 200.88 | Plevin et al (43) |
| CRP2NCRPW | Emission factors (MgCO2eq/ha) crop to non-crop conversion | -142.19 | -99.53 | -184.84 | Plevin et al (43) |
| ECROPW | Elasticity of substitution in production of crops | 0.55 | 0.72 | 0.39 | Keeney & Hertel (39) |
| EINVW | Elasticity of TFP wrt agricultural investments | 0.30 | 0.38 | 0.21 | Nelson et al. (24) |
| EIY('Crops','Intercept') | Elasticity of crop demand wrt income (intercept of line of elasticity vs. income) | 0.88 | 0.93 | 0.83 | Muhammad et al. (20) |
| EIY('Livestock','Intercept') | Elasticity of livestock demand wrt income | 1.05 | 1.09 | 1.01 | |
| EIY('Proc_Food','Intercept') | Elasticity of processed food demand wrt income | 1.20 | 1.24 | 1.15 | |
| EIY('Non_Food','Intercept') | Elasticity of non food demand wrt income | 1.56 | 1.65 | 1.46 | |
| EIP('Crops','Intercept') | Elasticity of crop demand wrt price | -0.74 | -0.71 | -0.77 | |
| EIP('Livestock','Intercept') | Elasticity of livestock demand wrt price | -0.83 | -0.80 | -0.85 | |
| EIP('Proc_Food','Intercept') | Elasticity of processed food demand wrt price | -1.17 | -1.13 | -1.21 | |
| EIP('Non_Food','Intercept') | Elasticity of non food demand wrt price | -1.14 | -1.07 | -1.21 | |

# Morris' Method Example

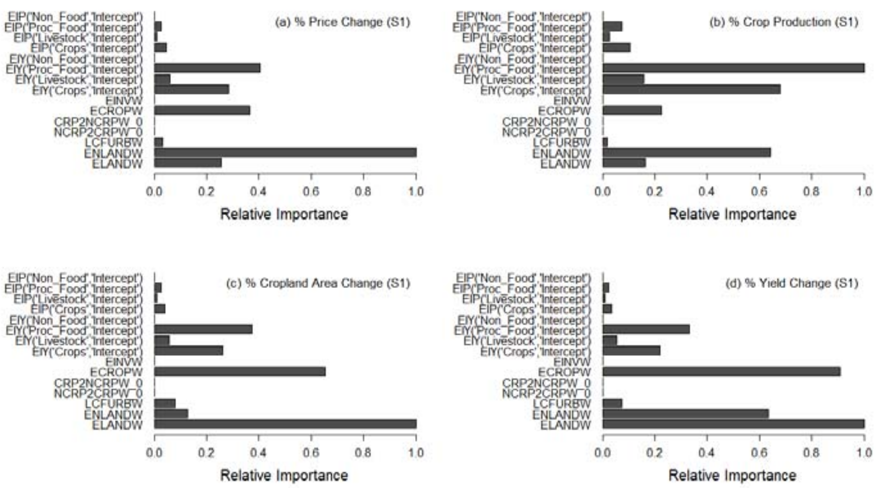



Lobell et al. ERL 2013

# Morris' Method Example

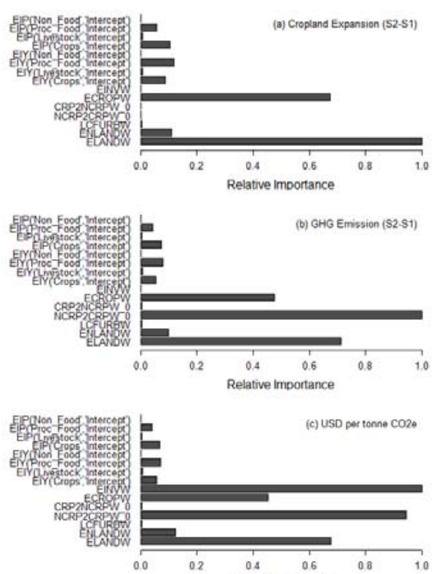
- Want to know how we might go about reducing the uncertainties (so use SA)
- Model is fairly simple but still takes some time to run
- Model has several interactions and nonlinearities
- So need a method that can capture nonlinearities and interactions but with few # runs (so use Morris)

# Morris' Method Example




Lobell et al. ERL 2013

# Morris' Method Example




Lobell et al. ERL 2013

# Examples of some other approaches

Running a model with and without some inputs fixed
e.g., predicting soil water in White and Nemani 2004

Their goal is to determine where effort should be spent to improve soil moisture predictions

Control: Run with variable weather and variable leaf area
Exp CLVM: Only variable weather (temp, and precip)
Exp CMVL: Only variable vegetation leaf area

# Examples of some other approaches

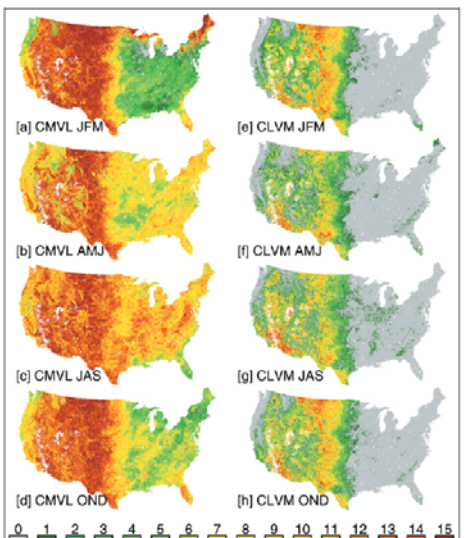



**Figure 3.** Number of years (1–15) with statistically significant (5% level) differences in seasonal soil water between experimental and control simulations. Experimental simulations are either climatological meteorology variable LAI (CMVL) or climatological LAI variable meteorology (CLVM). Seasons are January, February, and March (JFM); April, May, and June (AMJ); July, August, and September (JAS); and October, November, and December (OND).

What do you think they conclude?

# Examples of some other approaches

Characterizing the sensitivity of multiple outputs (e.g. a time series)

314 M. Lamboni et al. / Field Crops Research 113 (2009) 312–320

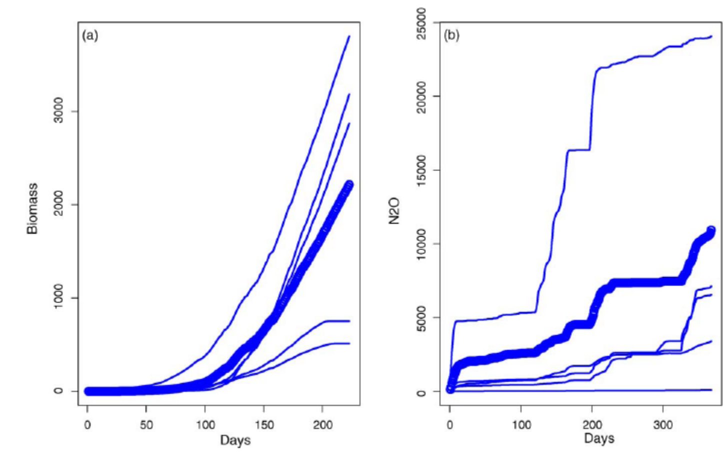



**Fig. 1.** Daily simulated values of the winter wheat dry matter model in g/m$^2$ (a) and of the CERES-EGC $N_2O$ emissions in g N ha$^{-1}$ day$^{-1}$ (b), for the nominal values of the parameters (thick lines) and for a sample of other possible values drawn within the uncertainty ranges.

# Examples of some other approaches

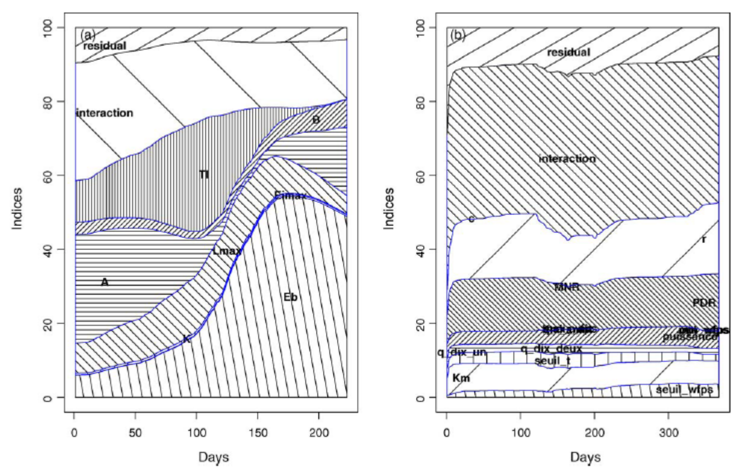



**Fig. 2.** Time-dependent pie charts of sensitivity indices for the WWDM model (a) and for the CERES-EGC model (b). Residual indicates the interactions between three or more parameters.