

Reflections on the Data Collection and Modeling Experiences

SAMSI Undergraduate Workshop 2008

David Bell
Elaine Spiller

May 21, 2008

Modeling

What is a model (or a mathematical model)?

Modeling

What is a model (or a mathematical model)?

- Mathematics is the language of change
- The use of mathematics to describe a system's behavior.

Why do we need them and what purpose do they serve?

Why do we need them and what purpose do they serve?

- To analyze a system to be controlled or optimized
- To hypothesize about how a system works
- To make predictions at parameter values and/or scales that are difficult to test

Are mathematical models the only models out there?

Are mathematical models the only models out there?

- Economic Models
- Business Models
- Behavioral Models
- These can have a mathematical component, and there are many other possible types of models.

What are the features of a mathematical model?

What are the features of a mathematical model?

- variables
- equations
- parameters
- This is the most basic description.

What are the common pitfalls of mathematical modeling?

What are the common pitfalls of mathematical modeling?

- Oversimplification
- Overcomplexity
- Invalid Assumptions

How do we assess if a mathematical model represents the real world?

How do we assess if a mathematical model represents the real world?

- Is fidelity to the data of most importance (sampling error)?
- What difference is there between the data we collect and the phenomenon of interest?

Can we do without the model and settle for the data only ?

Can we do without the model and settle for the data only ?

- We can create a purely descriptive model with few assumptions, but this has limited usefulness
- We want to model for either inference or prediction

Descriptive

Physical / Biological



What do we do once we choose a model - do we stick to it for eternity?

Data Collection

Why collect data?

Data Collection

Why collect data?

How do these data relate to the subject of interest?

Data Collection

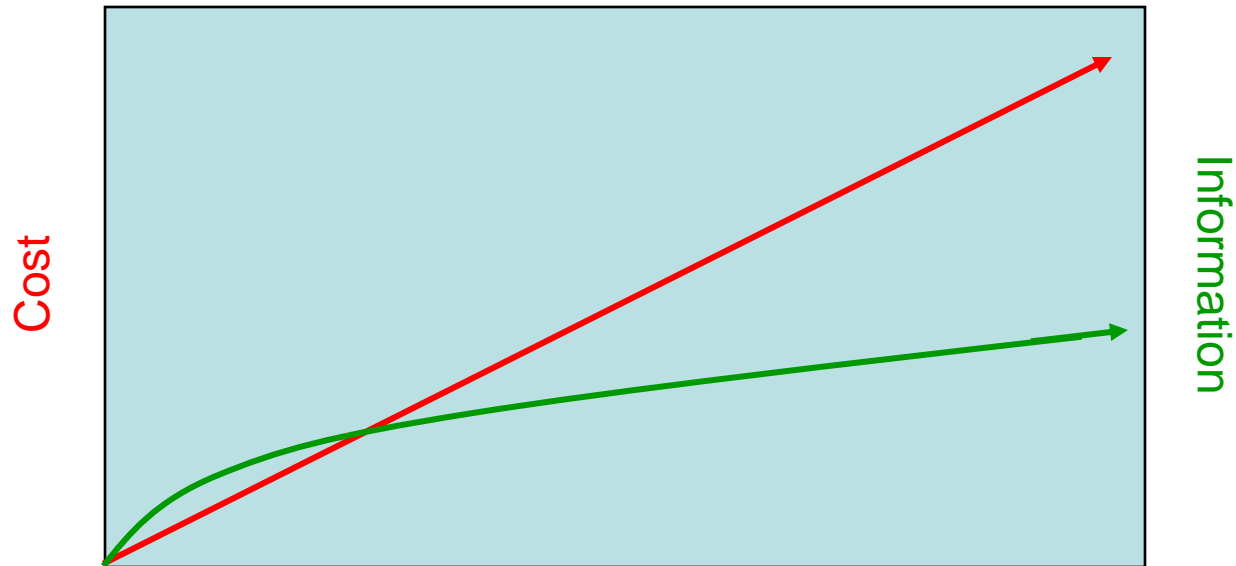
Why collect data?

How do these data relate to the subject of interest?

When should scientists start talking to statisticians?

How much data do we need to answer our question?

How much data do we need to answer our question?



How should we view missing or faulty data?

What comes first: the data or the model?

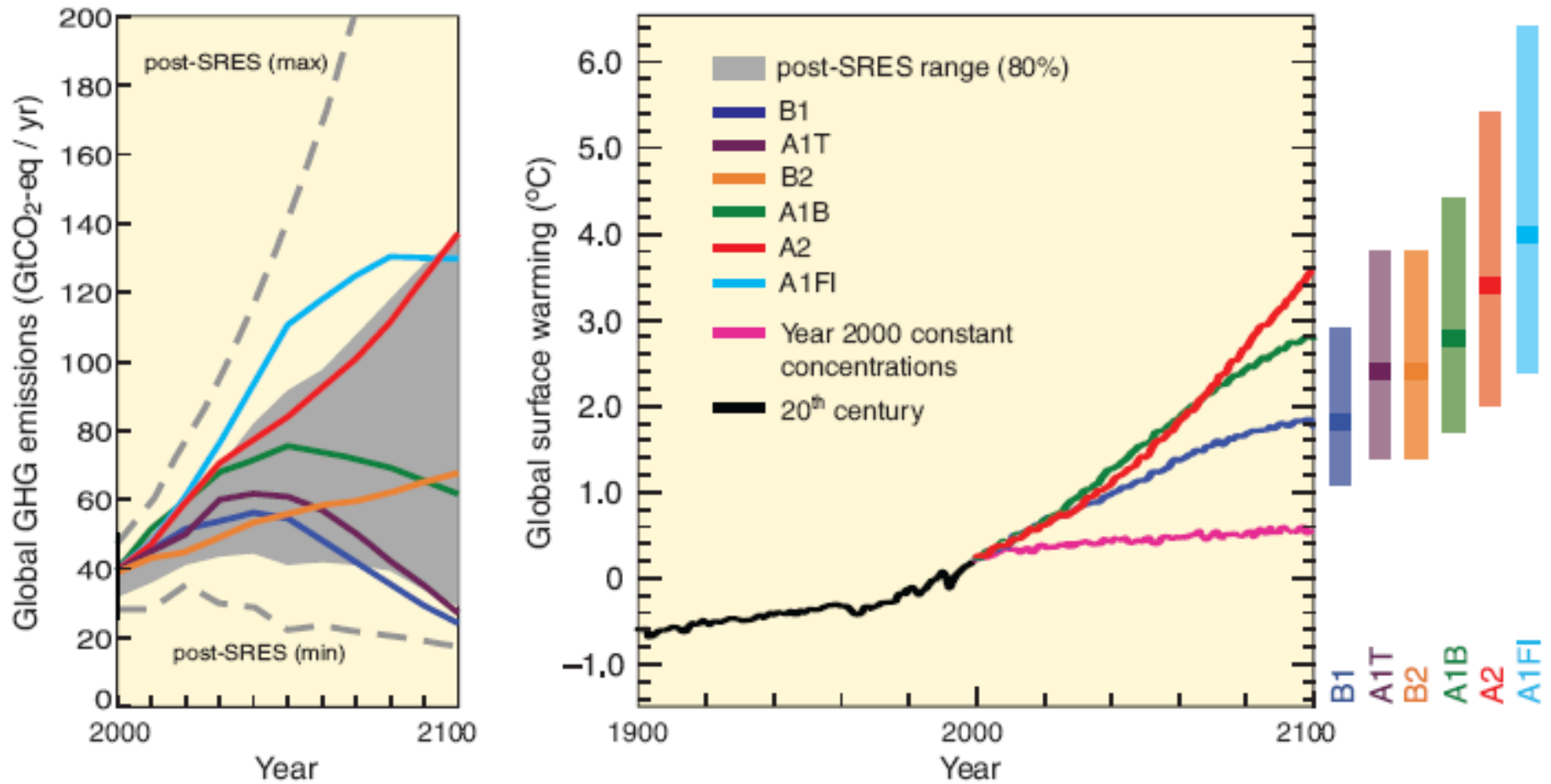
What comes first: the data or the model?

– Sort of like the “Chicken or the Egg”

Real World Example: Global Warming

- There is a scientific consensus that human activity contributes to global warming.
- However, different climate models (with large mathematical components) tend to give different answers regarding climate change.
- Different models may predict different rates of climate change, or may measure climate change differently.

**Scenarios for GHG emissions from 2000 to 2100 (in the absence of additional climate policies)
and projections of surface temperatures**



from IPCC. 2007. Summary for Policymakers

What could be the reason that they are giving different projections ?

How might these models differ?

What would you do with these different predictions?

How could one reconcile the differences ?

One Last Thing...

“All models are false, but some models are useful.”

-George E. P. Box