Webinar: Modeling for a Globally Connected World - What Models are Good for and How they Work

Presented by:
Professor Louis Gross
National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville

With support from the National Science Foundation (DBI-1300426)
Suzanne Lenhart, PhD
Chancellor’s Professor of Mathematics, University of Tennessee

Associate Director for Education and Outreach, National Institute for Mathematical and Biological Synthesis (NIMBioS)
HOW TO INTERACT TODAY

Question appears here

Welcome
Feel free to ask the host and panelists questions

Type your question here...

Type here
NIMBioS.org  A recording of this webinar will be posted within two days
MEET YOUR PRESENTER

Louis J. Gross, PhD

*Chancellor’s Professor* of Ecology and Evolutionary Biology and Mathematics, University of Tennessee

*Director*, The National Institute for Mathematical and Biological Synthesis, (NIMBioS), University of Tennessee

April 14, 2020
Webinar Objectives

• Provide an overview of modeling objectives and the process of developing a model
• Describe different types of models and their applications
• Discuss the limitations of models and how they are evaluated
An SIR-type model of an epidemic that accounts for the “observed” data from testing and the “real” spread, accounting for sensitivity of surveillance to clinical testing errors.

Nina Fefferman, Univ. of Tennessee
A model that uses experimental data on sneezes linked to mathematical analysis of turbulent fluid dynamics

Lydia Bourouiba, MIT
A model that uses a branching process to project the spread of COVID-19 in different US counties.

Lauren Ancel Meyers, Univ. of Texas
Infectious Disease Modeling at NIMBioS

Mathematical modeling of *Leptospira* transmission and intervention strategies

Evaluating the shifts in antimicrobial use practices and resistance resulting from risk mitigation strategy

Climate change and vector-borne diseases

Synthesizing and predicting infectious disease while accounting for endogenous risk

Integrated modeling and analysis of within-host infection and between-host transmission for *Toxoplasma gondii*

Optimal control of neglected tropical diseases
Order of topics

• Science and models
• Methods of investigation and theory
• Constraints on models
• Evaluating models
• Some lessons
What is science?

Science is thought to be a process of pure reductionism, taking the meaning out of mystery, explaining everything away, concentrating all our attention on measuring things and counting them up. I don't like this at all. The scientific method is guesswork, the making up of stories. The difference between this and other imaginative works of the mind is that science is then obliged to find out whether the guesses are correct, the stories true. Curiosity drives the enterprise, and the open acknowledgement of ignorance.

Lewis Thomas - Sierra Club Bulletin, March/April 1982, P. 52
What is science?

Science is thought to be a process of pure reductionism, taking the meaning out of mystery, explaining everything away, concentrating all our attention on measuring things and counting them up. It is not like this at all. The scientific method is guesswork, the making up of stories. The difference between this and other imaginative works of the human mind is that science is then obliged to find out whether the guesses are correct, the stories true. Curiosity drives the enterprise, and the open acknowledgement of ignorance.

Lewis Thomas - Sierra Club Bulletin, March/April 1982, P. 52
What is science?

Science is thought to be a process of pure reductionism, taking the meaning out of mystery, explaining everything away, concentrating all our attention on measuring things and counting them up. It is not like this at all. The scientific method is guesswork, the making up of stories. The difference between this and other imaginative works of the human mind is that science is then obliged to find out whether the guesses are correct, the stories true. Curiosity drives the enterprise, and the open acknowledgement of ignorance.

Lewis Thomas - Sierra Club Bulletin, March/April 1982, P. 52
Expressing Theory

Verbally
Graphically
Mathematically
Through Simulation

**Approaches to Develop Theory**

1. Descriptive: (a) Empirical
   (b) Comparative
2. Mechanistic (a) Compartmental
   (b) Optimization - adaptationist
3. Systems - hierarchy theory
4. Individual- or agent-based
5. Expert systems, machine learning
The “stories” in science are models

A model is a simplification of reality. Think of it as a map - it includes some features that represent what we observe but not others. Modeling is the process of selective ignorance - we select what to include and what to ignore.
The “stories” in science are models

Is this less of a simplification? Is it closer to “reality”? 
You make models all the time:

In the current pandemic, you are taking account of many factors, data from sources you trust, and your personal values/beliefs to decide who to interact with, whether to leave your home, and how often to do so. You may not be making a “formal” calculation of your personal risk of harm, but an underlying “model” is involved.

So you are deciding what is “best” for you
You make models all the time:

What decision do you make in more “normal” times when faced with:

So you are deciding what is “best” for you quite regularly
Models in Biology

- Physiology
- Neurobiology
- Development
- Model Systems

- Disease
- Microbiology
- Genetics

NIMBioS
National Institute for Mathematical and Biological Synthesis

The University of Tennessee, Knoxville
Taxonomy of Models

This could be based on the model objectives, on the general approaches used, or on the methodology. One possibility is:

- **conceptual**
- **verbal**
- **quantitative**
- **physical** (e.g. real, such as a physical model for an animal to evaluate heat-loading)
- **biological** (including animal models used for experiments, cell lines, and tissue cultures)

*Note that modeling text books typically classify models based on mathematical approach.*
Possible Model Objectives

1. Suggest observations and experiments
2. Provide a framework to assemble bodies of facts - provide a means to standardize data collection
3. "Allows us to imagine and explore a wider range of worlds than ours, giving new perceptions and questions about how our world came to be as it is" F. Jacob - The Possible and the Actual, 1982
4. Clarifies hypotheses and chains of argument
5. Identifies key components in systems
6. Allows simultaneous consideration of spatial and temporal change
7. Extrapolate to broad spatial or long temporal scales for which data can not easily be obtained
8. Prompts tentative and testable hypotheses
9. Serves as a crude guide to decision making in circumstances where action cannot wait for detailed studies
10. Provides an antidote to the helpless feeling that the world is too complex to understand in any generality - provides a means to get at general patterns and trends
11. To predict how a system will behave under different management, and control the system to meet some objective
Models and tradeoffs

No one model can do everything
Environmental Modeling

Data sources:
- GIS map layers (Vegetation, hydrology, elevation), Weather, Roads, Species densities

Models:
- Statistical
- Differential equations
- Matrix
- Agent-based

Simulation:
- Matlab, C++, Distributed, Parallel

Evaluation/Analysis:
- Visualization, corroboration, sensitivity, uncertainty

Management input:
- Harvest regulation
- Water control
- Reserve design

Monitoring:
- Species densities
- Animal telemetry
- Physical conditions
Describing Models

There is no single protocol to describe models. Some are described graphically.

Describing Models

Some are described mathematically with definition of the variables and parameters

\[
\frac{dN_1(t)}{dt} = a_1(a_2 - N_1) - f(N_1)N_2, \quad (1)
\]

\[
\frac{dN_2(t)}{dt} = b_1f(N_1)N_2 - b_2N_2 - u(t)N_2,
\]

\[
f(N_1) = \frac{a_1a_3N_1}{1 + a_3a_4N_1}, \quad (2)
\]

Describing Models

Described through the code that simulates the system and associated documentation and metadata


Describing Models

Some are described by the statistical methodology used, data and metadata

<table>
<thead>
<tr>
<th>Event Number</th>
<th>Event</th>
<th>Rate</th>
<th>Change(s) to state variable(s) (AX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Infection of uninfected host by pathogen 1</td>
<td>$F_1 \Delta t + o$ (Δt)</td>
<td>$I_0 \rightarrow I_2 - 1$ $I_1 \rightarrow I_1 + 1$</td>
</tr>
<tr>
<td>2</td>
<td>Infection of uninfected host by pathogen 2</td>
<td>$F_2 \Delta t + o$ (Δt)</td>
<td>$I_0 \rightarrow I_0 - 1$ $I_2 \rightarrow I_2 + 1$</td>
</tr>
<tr>
<td>3</td>
<td>Infection by pathogen 1 of host singly infected by pathogen 2</td>
<td>$F_1 \Delta t + o$ (Δt)</td>
<td>$I_2 \rightarrow I_2 - 1$ $I_1,2 \rightarrow I_1,2 + 1$</td>
</tr>
<tr>
<td>4</td>
<td>Infection by pathogen 2 of host singly infected by pathogen 1</td>
<td>$F_2 \Delta t + o$ (Δt)</td>
<td>$I_2 \rightarrow I_2 - 1$ $I_3,2 \rightarrow I_3,2 + 1$</td>
</tr>
<tr>
<td>5</td>
<td>Death of host singly infected by pathogen 1 and replacement with an uninfected host</td>
<td>$\mu_1 \Delta t + o$ (Δt)</td>
<td>$I_1 \rightarrow I_1 - 1$ $I_3 \rightarrow I_3 + 1$</td>
</tr>
<tr>
<td>6</td>
<td>Death of host singly infected by pathogen 2 and replacement with an uninfected host</td>
<td>$\mu_2 \Delta t + o$ (Δt)</td>
<td>$I_2 \rightarrow I_2 - 1$ $I_0 \rightarrow I_0 + 1$</td>
</tr>
<tr>
<td>7</td>
<td>Death of coinfections by pathogen 1 and replacement with an uninfected host</td>
<td>$\mu_1 \Delta t + o$ (Δt)</td>
<td>$I_3,1 \rightarrow I_3,1 - 1$ $I_0 \rightarrow I_0 + 1$</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pbio.3000551.t001

<table>
<thead>
<tr>
<th>Pathogens with n distinct types, strains, or clones</th>
<th>Observed counts, $O_k$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human papillomavirus</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0  1  2  3  4  5  6  7  8  9</td>
<td>N</td>
</tr>
<tr>
<td>25</td>
<td>2,933 140 64 26 102 39 12 2 2 -</td>
<td>5,412</td>
</tr>
<tr>
<td>Anther smut (M. violaceum)</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0  1  2  3  4  5  6  7  8  9</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>285 74 60 32 14 3 3 2 1 1</td>
<td>475</td>
</tr>
<tr>
<td>B. afzelii on bank voles</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0  1  2  3  4  5  6  7  8  9</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>807 33 26 13 10 11 6 - - -</td>
<td>906</td>
</tr>
<tr>
<td>Malaria (P. vivax)</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0  1  2  3  4  5  6  7  8  9</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>1,023 404 291 208 118 50 16 5 1 1</td>
<td>2,117</td>
</tr>
</tbody>
</table>

Abbreviation: NISP, Noninteracting Similar Pathogens

https://doi.org/10.1371/journal.pbio.3000551.t002

An example of a general protocol developed specifically for agent-based models but useful across other model types is the **ODD Protocol**. This is an update of a protocol first published in 2006 for “Overview, Design Concepts and Details”
### Elements of the updated ODD protocol

1. Purpose
2. Entities, state variables, and scales
3. Process overview and scheduling
4. Design concepts
   - Basic principles
   - Emergence
   - Adaptation
   - Objectives
   - Learning
   - Prediction
   - Sensing
   - Interaction
   - Stochasticity
   - Collectives
   - Observation
5. Initialization
6. Input data
7. Submodels
Constraints on models

*Data constraints:* Available data may not be sufficient to specify appropriate functional forms, interrelationships, or parameters. May force aggregation of components. May not be sufficient to elaborate criteria for evaluation of model performance.

*Effort constraints:* Resource constraints may limit the amount of detail it is feasible to include. Limits on time modelers and collaborators may invest as well as pressure to produce results.

*Computational constraints:* Despite great enhancements in computational resources, there are many problems still not feasible to carry out computationally.

*Other constraints:* ethical or other societal considerations.
Given the above, the entire modeling process involves evaluation of alternative approaches to assess the most appropriate procedures for the questions of concern. This is part of the process of selective ignorance involved in constructing models.

Just as public policy decisions involve a balancing act between various alternatives which satisfy to varying degrees the desires of different stakeholders, realism in modeling involves balancing different approaches to meet a goal.

Realistic modeling is the science of the actual rather than the science of the ideal.
Model evaluation – some terminology

**Verification** - model behaves as intended, i.e. equations correctly represent assumptions; equations are self-consistent and dimensionally correct. Analysis is correct. Coding is correct - there are no bugs.

**Calibration** - use of data to determine parameters so the model "agrees" with data. This is specific to a given criteria for accuracy. Some call this Tuning or Curve-fitting.

**Corroboration** - model is in agreement with a set of data independent from that used to construct and calibrate it.

**Validation** - model is in agreement with real system it represents with respect to the specific purposes for which it was constructed. Thus there is an implied notion of accuracy and domain of applicability.

**Evaluation (testing)** - appropriateness to objectives; utility; plausibility; elegance; simplicity; flexibility.
Evaluation and model objectives

Given the many objectives for models, we should expect there to be multiple criteria for evaluating whether a model is useful.

Before developing a model in any detail, criteria should be established for evaluating its use.

Evaluation procedures should account for constraints of Data, Effort and Resources, Computation.

Evaluation criteria should be taken into consideration in assessing methods, level of detail, scale, and what to ignore in deciding on a model.
Evaluating different types of models

Models for theory development –

General, some realism, little precision.

Make qualitative comparisons to patterns, not quantitative ones, over some parameter space. No calibration or corroboration performed, except theoretical corroboration (meaning that model agrees with the general body of theory in the field).
Evaluating different types of models

Descriptive models-
  Precise, little realism, not general
Statistical hypothesis testing; time series analysis methods applied.

Models for specific systems -
  Realism, some precision, not general
Quantitative comparisons, constrained by available data. Compare component-by-component if data are available.
CMIP5 promotes a standard set of model simulations from 20 different GCMs in order to:

• evaluate how realistic the models are in simulating the recent past,
• provide projections of future climate change on near term (out to about 2035) and long term (out to 2100 and beyond), and
• understand factors responsible for differences in model projections, including quantifying some key feedbacks such as those involving clouds and the carbon cycle.

Example of complex model evaluation – Global Climate Models

https://pcmdi.llnl.gov/mips/cmip5
Global Mean Surface Temperature to 2050

Climate models assume static representative concentration pathways (RCPs) for anthropogenic GHG emissions.

Despite the importance of evaluation, many published models lack explicit criteria or consideration of this. Why?

1. It’s difficult and requires potentially different skill sets from those constructing and using models.
2. Science is very much a human enterprise and it is natural that once one has devoted considerable effort to developing a particular model, it is difficult to critique yourself.
3. Modern settings with a great amount of team effort to develop models or experimental protocols can constrain individuals who do not wish to be an outcast in a lab.
Criteria I use in Reviewing Modeling Papers

1. Are the models appropriate to the biological questions being addressed?
2. Are the underlying biological questions of potential interest to a significant fraction of the journal’s audience?
3. Does the mathematics/model teach us anything new that is biologically significant?
4. Is the mathematics/modeling correct?
5. If the paper is strictly theoretical, does it point out broadly useful new insights?
6. Are the model parameters and variables estimable from observations?
7. Is there some effort devoted to model evaluation?
Take home lessons

• Model evaluation for all types of biological models is relatively rare.
• Set criteria for model evaluation prior to expending a lot of effort on a model.
• Tie evaluation criteria to model objectives.
• Encourage consideration of evaluation in all your educational initiatives.
• Multiple models are good – encourage this.
• Consider whether an evaluation has been done or discussed whenever you review a paper.
Thank you for your participation!

Questions/comments? Please use the Questions Button on Zoom to post these.