

Types of Uncertainty in Modeling

Andreas Handel

University of Georgia

2019-07-11

Uncertainty

- Model results have different sources of uncertainty attached to them.
- Not all types of uncertainty are always explicitly acknowledged.
- That's true not only for mathematical/computer models.

Structural Uncertainty

Structural Uncertainty

- Models are simplifications and abstractions of the real world.
- Specific assumptions lead to different models.
- Every model is 'wrong' in some sense, but some might be useful.



Which one is the right model of Georgia?

Structural Uncertainty

- We need to decide which variables and processes to include and which to exclude.
 - Include age or not?
 - Allow for co-infection or not?
 - Include treatment or not?
- Once we have decided on that, we need to decide how to model each process.
 - What type of model? (ODE, IBM, etc.)
 - How to implement each process?

Example variants of specific processes

Interaction: Mass-action or saturating?

Growth: Exponential or Linear or Saturating?

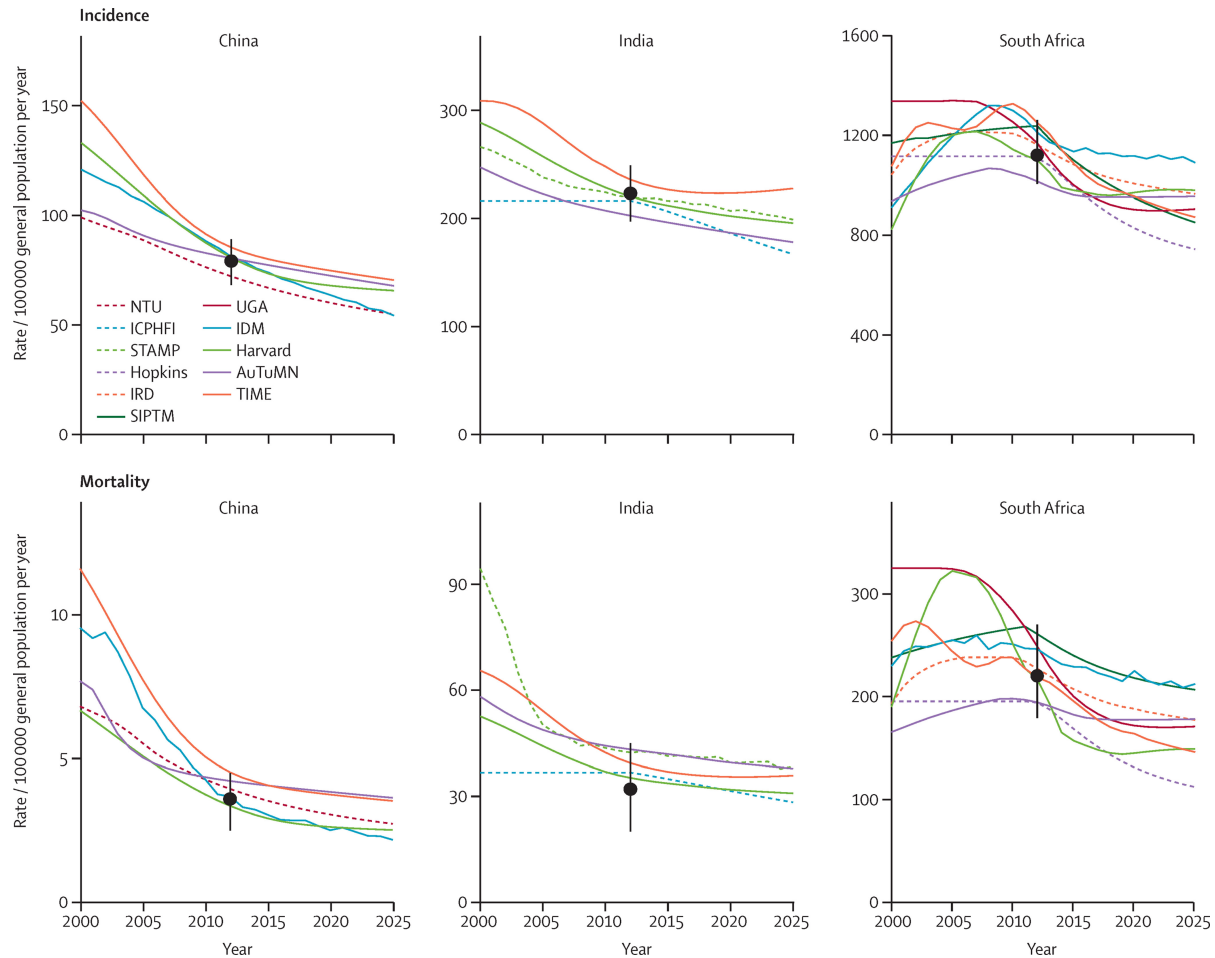
Structural Uncertainty - Example 1

- Multi-group effort to use computer models to predict how different interventions affect TB incidence and prevalence in 2025.

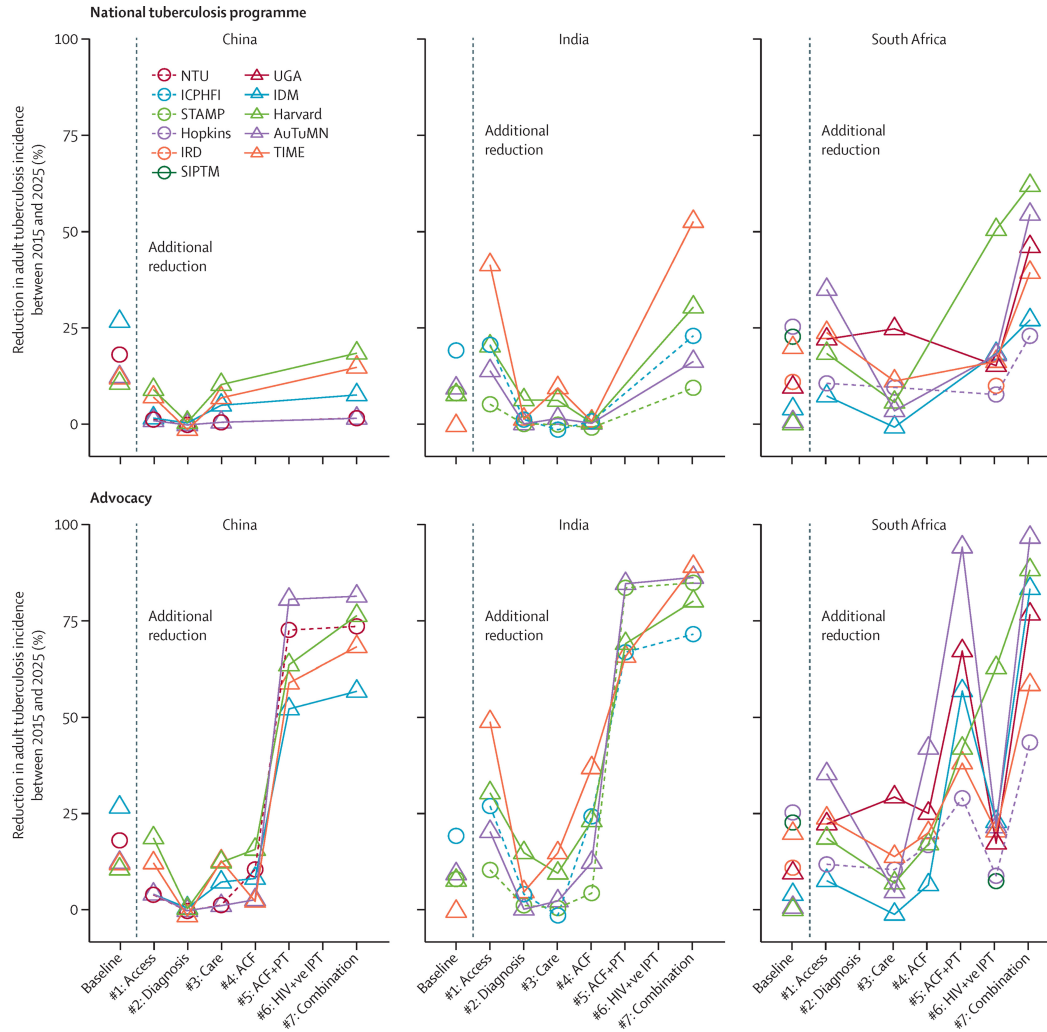


- More details:
 - Houben et al "*Feasibility of achieving the 2025 WHO global tuberculosis targets in South Africa, China, and India: a combined analysis of 11 mathematical models.*" Lancet Global Health, 2016.
 - Menzies et al "*Cost-effectiveness and resource implications of aggressive action on tuberculosis in China, India, and South Africa: a combined analysis of nine models.*" Lancet Global Health, 2016.

Structural Uncertainty - Example 1

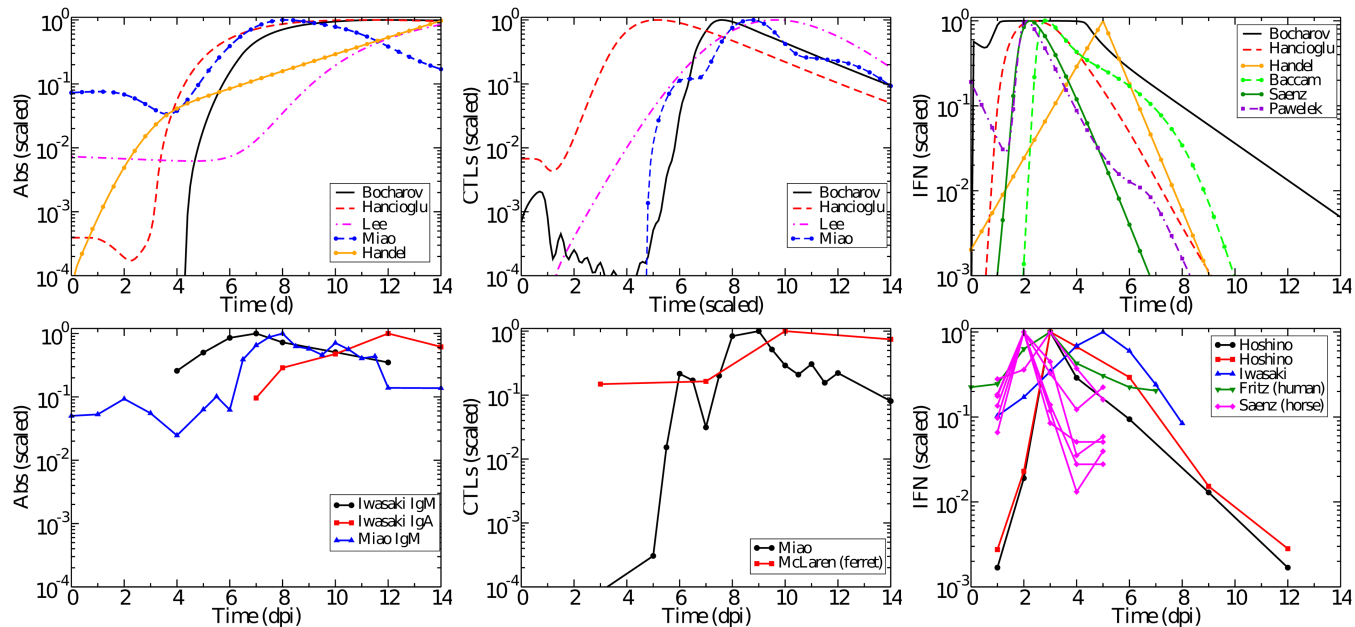


Structural Uncertainty - Example 1



Structural Uncertainty - Example 2

Dobrovolny et al compared different influenza models and assessed how they matched experimental data.



Top: models, bottom: data

More details: Dobrovolny et al (2013) PLoS One. *Assessing Mathematical Models of Influenza Infections Using Features of the Immune Response.*

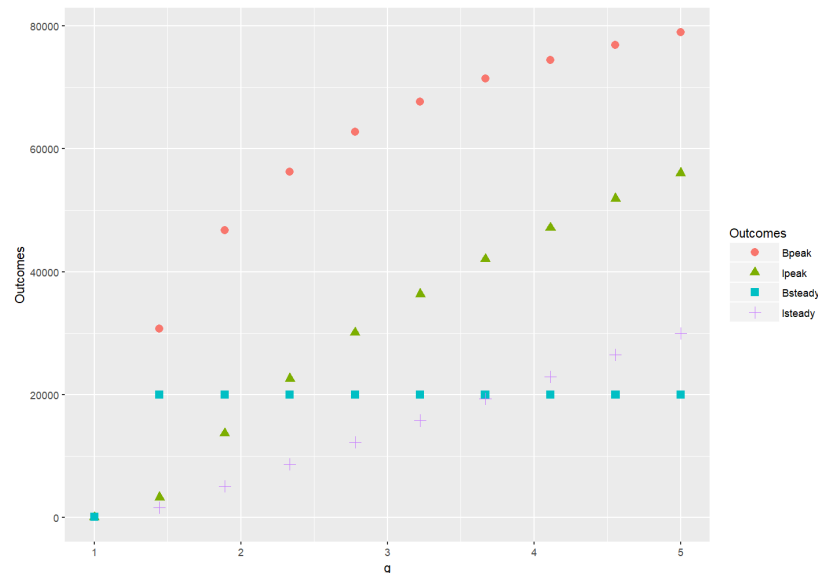
Structural Uncertainty - Practice

- The *Model variant exploration* app in DSAIRM explores the impact of different model formulations.

Parameter Uncertainty

Exploring the impact of model parameters

- Often, we do not know the values for the model parameters very well.
- Sometimes we can obtain estimates for parameter values from fitting to data, but the right data is often not available.
- Instead of running our model for just one set of parameter values, we could run it using different values that are within reasonable ranges.
- If we have a small model, we could systematically explore model behavior for each parameter.



Exploring the impact of model parameters

- Once models get big, it would take too long to thoroughly scan over all parameters.
- Often we can get reasonable ranges from the literature, but not exact values.
- Sometimes, we might be *mainly* interested in how results change as we vary one parameter, but we also want to know how uncertainty in other parameters affect our outcome.
- Other times, we might want to use our model to make predictions. We need to figure out how uncertainty in parameters affects our predictions.

Uncertainty & Sensitivity Analysis

Varying multiple inputs/parameters over a usually broad range is called a (global) uncertainty & sensitivity analysis.

- **Uncertainty Analysis:** Given uncertainty in the inputs (parameters), how much uncertainty is there in the outputs/results?
- **Sensitivity Analysis:** How much do individual inputs contribute to the uncertainty in outputs/results?

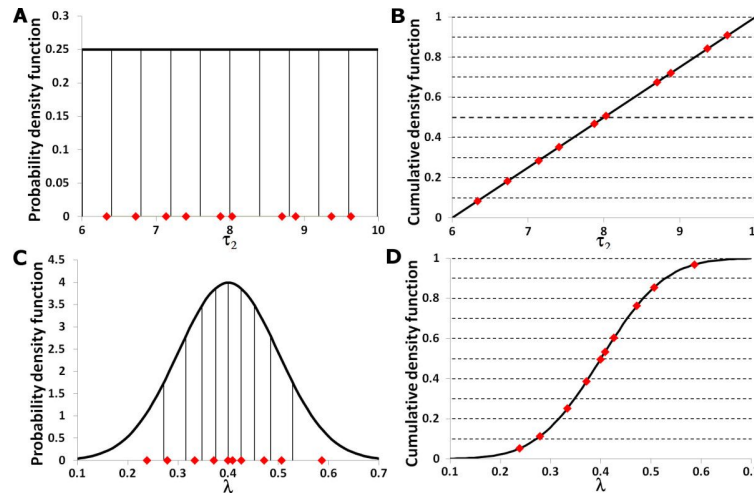
Doing Uncertainty & Sensitivity Analysis

- Determine ranges of uncertainty for each input (parameter and initial conditions).
- Repeatedly draw samples of parameter values from the specified distributions.
- Run model for each parameter sample, record outcomes.

Parameter ranges

For each parameter, specify its distribution:

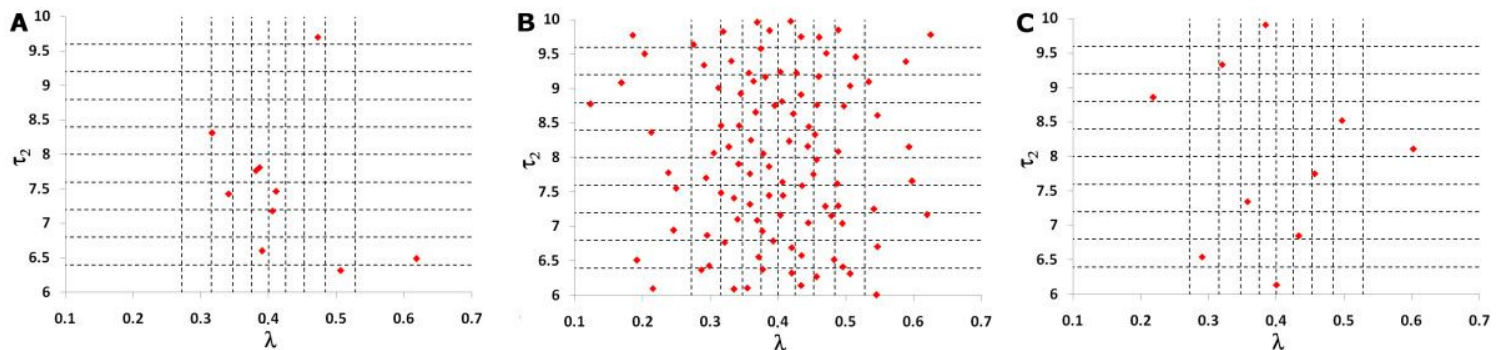
- Single value if we are very certain.
- Uniform distribution between some min and max values if we know very little.
- A 'peaked' distribution (e.g. gamma, log-normal) if we are fairly certain about some parameters.



Hoare et al 2008 TBMM

Parameter sampling

- Exhaustively trying all parameter value combinations takes too long.
- Randomly sampling is not efficient, it might leave areas of parameter space unexplored.
- A method called Latin Hypercube Sampling creates samples that efficiently span the parameter space.



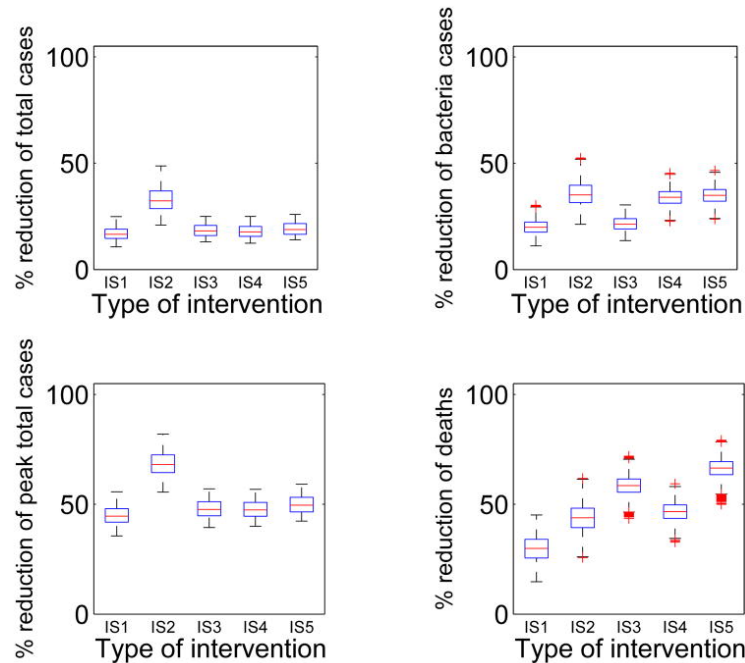
Hoare et al 2008 TBMM

Uncertainty Analysis - Example

- Question: What is the impact of different antiviral and antibacterial treatment strategies on deaths and cases during an influenza pandemic in the presence of bacterial co-infection?
- ODE equation model with 13 variables/equations and 47 parameters.
- Investigated the impact of 5 different treatment strategies on 4 outcomes: reduction in total cases, bacteria cases, peak cases and deaths.
- More details: Handel et al (2009) *Epidemics "Intervention strategies for an influenza pandemic taking into account secondary bacterial infections"*.

Uncertainty Analysis - Example

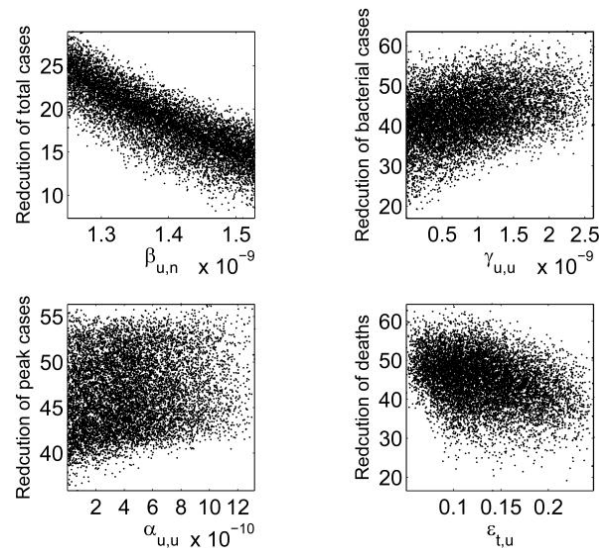
- For each parameter sample, we run the simulation and record the result.
- We can then see how uncertainty in inputs affects the results.
- A convenient way to represent the results is by using boxplots.



Handel et al (2009), Epidemics

Sensitivity Analysis

- Uncertainty analysis tells us how uncertainty in inputs affects uncertainty in outputs.
- If we want to know in more detail how a specific input affects a specific output, we can move on to sensitivity analysis.
- Using the same simulation results, we now plot scatterplots for income/outcome pairs of interest instead of boxplots.



Handel et al (2009), Epidemics

Sensitivity Analysis

- If the scatterplot shows a monotone relation, we can summarize it with a single number, a correlation coefficient
- Correlation Coefficients (CC) indicate how correlated a given output is with a given input.
- CC are between -1 and 1. Large CC means strong (negative) correlation, CC ≈ 0 means no correlation.
- Input-output relations are often nonlinear, therefore computing the linear correlation is often not the best measure.
- Using a Rank CC is usually more suitable.
- Partial Rank Correlation Coefficients (PRCC) are even better if multiple inputs/parameters are changed at the same time.

U/S Analysis - Summary

- Uncertainty in parameters can be fairly easily quantified.
- By sampling over parameters, one can obtain confidence intervals or similar measures of uncertainty.
- **This does not account for structural or inherent uncertainty!**

Parameter uncertainty - Practice

- The *U/S analysis* app explores the concept in more detail.

Inherent/Stochastic Uncertainty

Introduction

- Deterministic models (both continuous and discrete-time) give you the same result for a set of parameters and starting conditions no matter how often you run them.
- Real systems are stochastic/random, which means they have some inherent variability.
- The words Stochasticity, Randomness, Noise, Variability are often used interchangeably.

Sources of Stochasticity

- Random (external) noise (e.g. measurement error).
- Fluctuations in parameters (e.g. due to temperature or circadian rhythms).
- Unpredictability of event occurrence (e.g. births, deaths).

When is stochasticity important

- For small numbers.
- When we want to answer questions such as "What is the probability that X will happen?"

Stochastic compartmental models

We can fairly easily formulate any compartmental model (e.g. and ODE model) as a stochastic model.

- All variables take on discrete values (0,1,2,...).
- Increases or decreases are dictated by inflows and outflows.
- At each time step, one of the possible inflows/outflows occurs, with probability based on the size of the term.

Some terminology

- The events that happen are often called reactions or transitions.
- The inflow and outflow terms are called propensities, multiplied by the time step they are probabilities.

Example 1

Event type	Transitions	Propensity
Infection	$S \Rightarrow S-1, I \Rightarrow I+1$	bSI
Recovery	$I \Rightarrow I-1, R \Rightarrow R+1$	gI
Births	$S \Rightarrow S+1$	m
Death of susceptible	$S \Rightarrow S-1$	nS
Death of infected	$I \Rightarrow I-1$	nI
Death of recovered	$R \Rightarrow R-1$	nR

Example 2

Event type	Transitions	Propensity
Production of U	$U \Rightarrow U+1$	n
death/removal of U	$U \Rightarrow U-1$	U
infection	$U \Rightarrow U-1, V \Rightarrow V-1, I \Rightarrow I+1$	bUV
death if I	$I \Rightarrow I-1$	I
production of V	$V \Rightarrow V+1$	pI
removal of V	$V \Rightarrow V-1$	V

Comments

- Stochastic models are not much harder to implement, but they require more computational time since we need to run ensembles.
- One can fit stochastic models to data, but that's tricky and takes a long time.
- The `adaptivetau` package in R makes implementing stochastic models easy.
- The `pomp` package in R is a good tool for fitting stochastic models.

Stochastic uncertainty - Practice

- The apps in the *Stochastic models* section of DSAIDE provide more details.
- The *Stochastic Model* and *Influenza antivirals and resistance* apps in DSAIRM illustrate stochastic models.

Summary

- Uncertainty in modeling results can come from different sources.
- Structural uncertainty is arguably the most important but rarely explored. (Similar problems exist outside of modeling, e.g. how good a model are mice or elite undergrad students for the general population?)
- Parameter uncertainty is increasingly included in modeling projects and should be explored if a model is used for prediction.
- Stochastic models are also becoming more common but are still somewhat harder to work with.