CMPT 394 Lectures Wrap-Up & Resources for Further Study

Nathaniel Osgood

CMPT 394

April 9, 2013

Important Topics Put Aside

- AnyLogic capabilities (see www.tinyurl.com/ConsensusABM/Youtube channel)
  - Data collection & output
  - Debugging
  - Hierarchical metapopulation models
  - Diverse sorts of hybrid models
  - Building user interfaces
  - Performance/computational burden

Conceptually Rich Topics

- Linking dynamic simulation with statistical tools
  - Predictor-corrector models
  - Posterior derivation
- Linking models & data
- Multi-scale modeling
- Real-time predictive simulation
- Model specification & domain specific languages
- Mathematical analysis tools
- Numerical analysis considerations
- Alternative mathematical formalisms (hybrid automata, differential Equation variants, DAEs, etc.)
- Performance
- Eval. study design & stats w/synth. pop experiments
- Planning experiments using modeling

Conceptually Rich Topics

- Linking dynamic simulation with statistical tools
  - Predictor-corrector models
  - Posterior derivation
- Linking models & data
- Multi-scale modeling
- Real-time predictive simulation
- Model specification & domain specific languages
- Mathematical analysis tools
- Numerical analysis considerations
- Alternative mathematical formalisms (hybrid automata, differential Equation variants, DAEs, etc.)
- Performance
- Eval. study design & stats w/synth. pop experiments

Predictor Corrector Models
Benefits of Synergizing Models & Ongoing Measurement via "Closed Loop Models"

Benefits to Data

- Interpreting for implications to other areas of the system not directly measured
- Understanding implications for decision making
- Separating signal from noise: Avoiding overconfidence in measurements
- Generalization/abstraction to broader dynamic patterns of behavior

Benefits to Models

- Preventing model state divergence from actual situation
- Maintaining model "freshness" by repeated re-grounding in measured data
- Better understanding of current situation
- More reliable prospective simulation with the model
- Avoiding overconfidence in model output

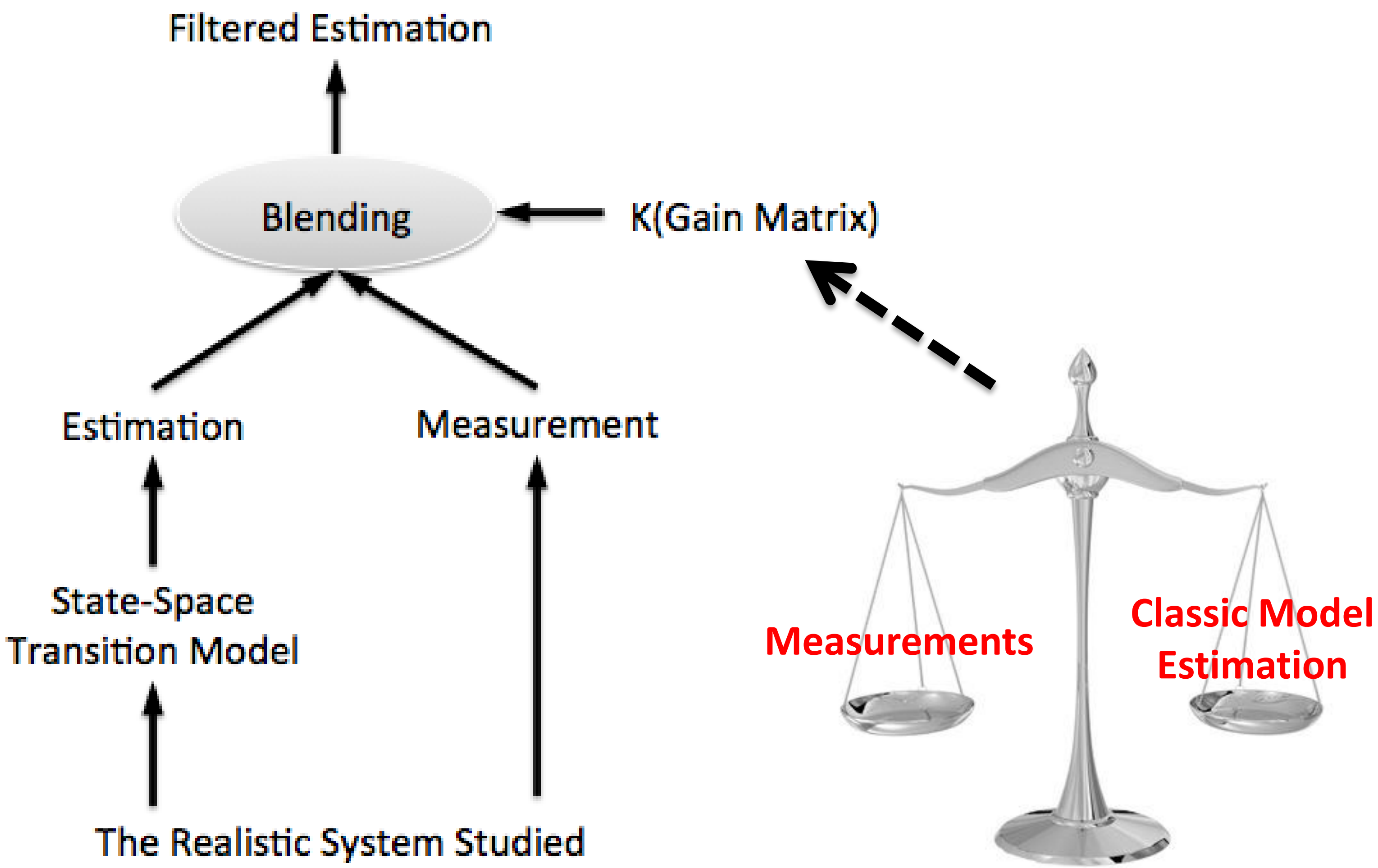
Predictor Corrector Methods

- “Brute force”
  - Parameter point estimates: Recalibrate based on updated data
  - Posterior distribution over parameters: Run full (Batch) Monte Carlo (e.g. MCMC) for each new point
- Recursive (Incremental computation)
  - Point estimates: Kalman Filter
  - Posterior distribution over parameters: Sequential Monte Carlo & Particle Filtering

Why Filter Aggregate Models?

- Typically the state in dynamic models involve both observable & non-observable elements
- We can make limited inference on non-observable components from observable
- Many observations are often required to "triangulate" non-observables
- Inferring non-observables is far easier if there are fewer of them => more aggregate models
  - Almost always required for aggregate measurement
- Ubiquitous sensing does raise the intriguing potential for inferring state at the individual level

Dissecting the Kalman Filter



Slide courtesy of Weicheng Qian

Kalman Filter Equations

$$\dot{X}(t) = f(X(t), t) + w(t)$$

$$\dot{Y}(t = k_i) = h_i(X(t = k_i)) + v(t = k_i)$$

System Theory

$$\hat{x}_k(+) = \hat{x}_k(-) + K_k[z_k - h_k(\hat{x}_k(-))]$$

$$P_k(+) = [I - K_k H_k(\hat{x}_k(-))]P_k(-)$$

Measurement Update Equations

$$H_k(\hat{x}_k(-)) = \left.\frac{\partial h_k(x(t_k))}{\partial x(t_k)}\right|_{x(t_k)=\hat{x}_k(-)}$$

$$K = \frac{PH^T}{HPH^T + R}$$

'Gain' (Weighting) Matrix

$$F(\hat{x}(t), t) = \left.\frac{\partial f(x(t), t)}{\partial x(t)}\right|_{x(t)=\hat{x}(t)}$$

$$\dot{\hat{x}}(t) = f(\hat{x}(t), t)$$

$$\dot{P}(t) = F(\hat{x}(t), t)P(t) + P(t)F^T(\hat{x}(t), t) + Q(t)$$

System evolution between measurements

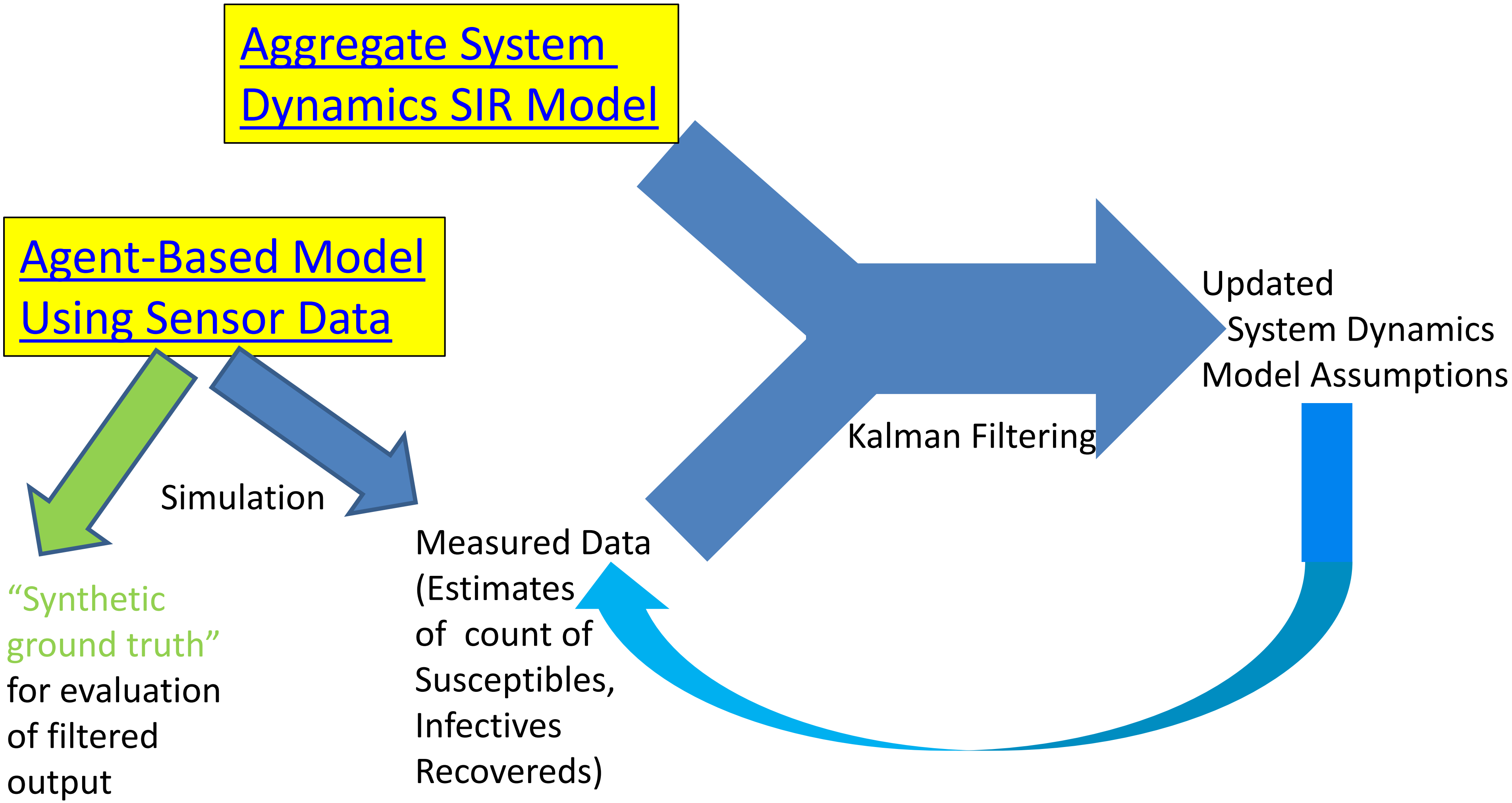
Evaluating Using a Synthetic Population

- Analytic approaches (and study designs) are often challenging and costly to test in the real world
  - Expensive to establish study
  - Time consuming
  - Ethical barriers
  - Lack of definitive knowledge of how conclusions compare to some “ground truth”
- We can often evaluate such approaches using “synthetic populations” drawn from simulation models
  - Here, the simulation model helps to identify potential weaknesses of study designs & analysis approaches

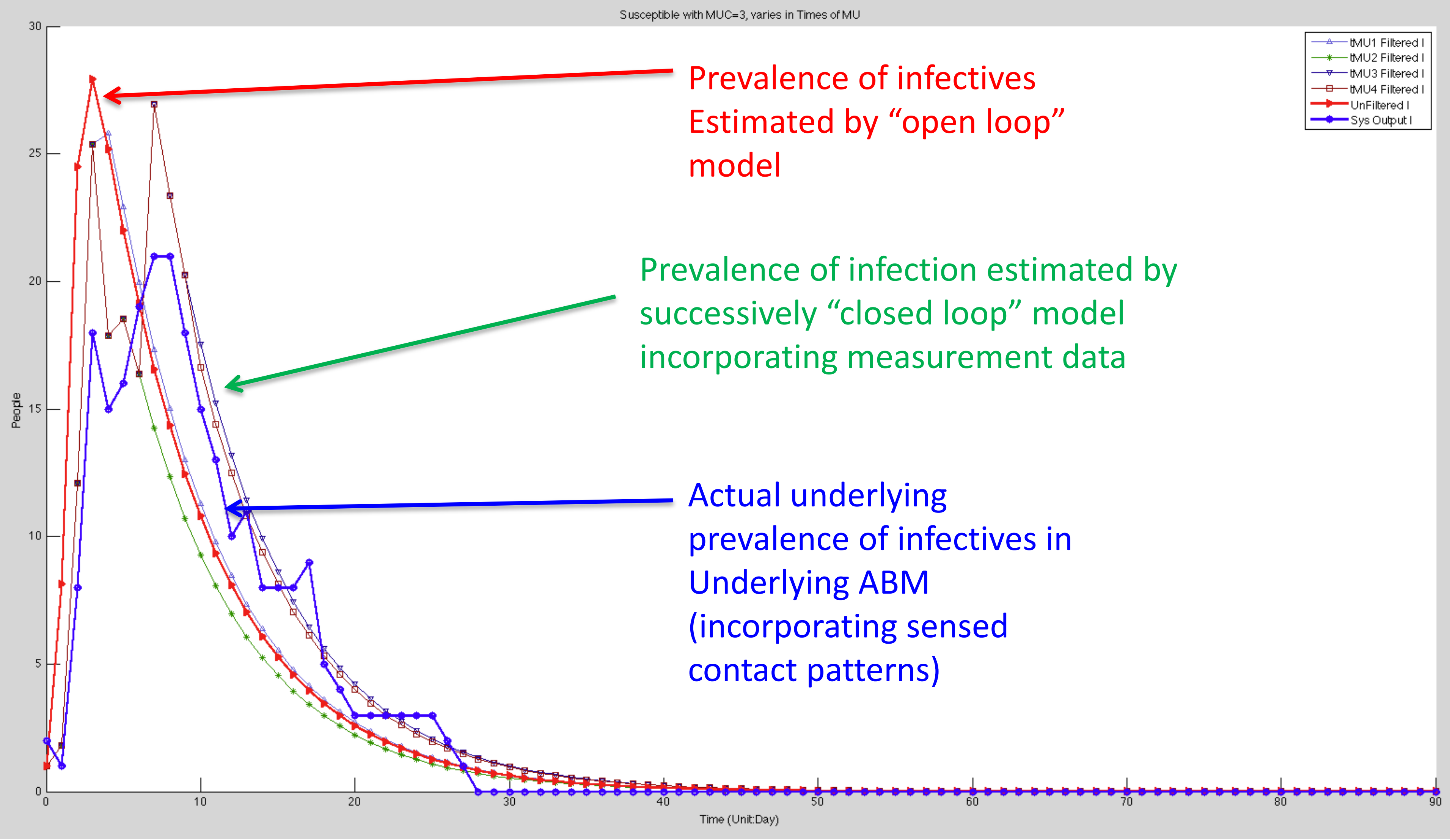
Synthetic Population Studies

- Establish a “synthetic population” for a “virtual study”
- Perform simulation, simulating study design of interest
  - Actual underlying situation is blinded from researcher
  - Collect data from the synthetic population similar to what would collect in the external world
  - Optionally, may actually simulate roll out and dynamic decision protocols
- Analysis procedures being evaluated are applied to the data from the synthetic population
- We compare the findings from those analysis procedures to the underlying “ground truth” in the simulation model

Performing the Filtering



Projecting Forward from Updates



Conceptually Rich Topics

- Linking dynamic simulation with statistical tools
  - Predictor-corrector models
  - Posterior derivation
- Linking models & data
- Multi-scale modeling
- Real-time predictive simulation
- Model specification & domain specific languages
- Mathematical analysis tools
- Numerical analysis considerations
- Alternative mathematical formalisms (hybrid automata, differential Equation variants, DAEs, etc.)
- Performance
- Eval. study design & stats w/synth. pop experiments

Traditional Parameter Estimation via Calibration in (Dynamic) Simulation Models

- Point estimation of parameter values
- Optional: Estimation of confidence intervals surrounding estimate
- Limitations
  - Single dynamic model assumed
  - Single point estimates (doesn't make apparent how this yields estimate in certain output measures – e.g. comparison between interventions)
  - Even with estimate of intervals, assuming that unimodal
    - No understanding of global shape of likelihood
  - Error metric for estimation (e.g. square of estimate) imposes distributional assumptions on errors

Prior & Posterior Parameter Distributions

- Before calibration, we often have some sense as to where the parameters fall
  - We can encode this with a "prior" distribution
    - This is called a "prior" because we can formulate it prior to observing the data or settling on a model
- Calibration can give us an updated distribution, which we call the "posterior" distribution
  - This takes into account not only our best guesses for the parameter values, but also the likelihood that the model(s) used could explain the observed data given certain parameter values

Posterior Distributions for Parameters (θ): Moving Beyond Point Estimates

- A posterior distribution over θ helps us understand the relative probabil. of parameter values in light of
  - The observed data (y)
  - One or more models under consideration
  - Any pre-existing expectations of the relative likelihood of different parameter values (a prior distribution over θ)
- Given a posterior over θ, we can e.g. derive
  - Point estimates (MLE, mean, Minimum var, &c) of parameters & model output
  - Likelihood intervention A is "better" than intervention B
  - "Credibility intervals" for diverse model output

Markov Chain Monte Carlo Helps us Reliably Derive Parameter Posterior Distributions

- MCMC is a principled way of generating samples from θ according to posterior distribution over θ
- These samples from the posterior distribution of θ are generated over and over again $\Rightarrow$ Monte Carlo
  - For each such sample, we can e.g.
    - Accumulate a running sample mean of e.g.
      - θ
      - Model outputs
      - Intervention
    - Record the likelihood at that point
    - Record the shape of the distribution
- A "Markov Chain" is used to generate the θ samples

High-Level Operation of Using MCMC for Parameter Estimation of Simulation Model

- Elements to specify before algorithm to use in algorithm
  - Assign prior distributions (over parameters, models, etc.) may or may not be parametric (using ‘hyperparameters’)
  - Formulate likelihood formulae that indicate relative likelihood of seeing particular values of empirical data (y) in light of values of parameters (θ) & simulation model output
    - These could be either parametric or non-parametric
  - Conceptually (only), apply Bayes’ rule to get rule for likelihood of parameter θ in terms of data (y) and priors
- A Markov Chain is used to generate samples of parameters (θ), with (asymptotic stationary) distribution proportional to posterior distribution
  - For each generated parameter sample, we sample output measures of interest & accumulate statistics, etc.

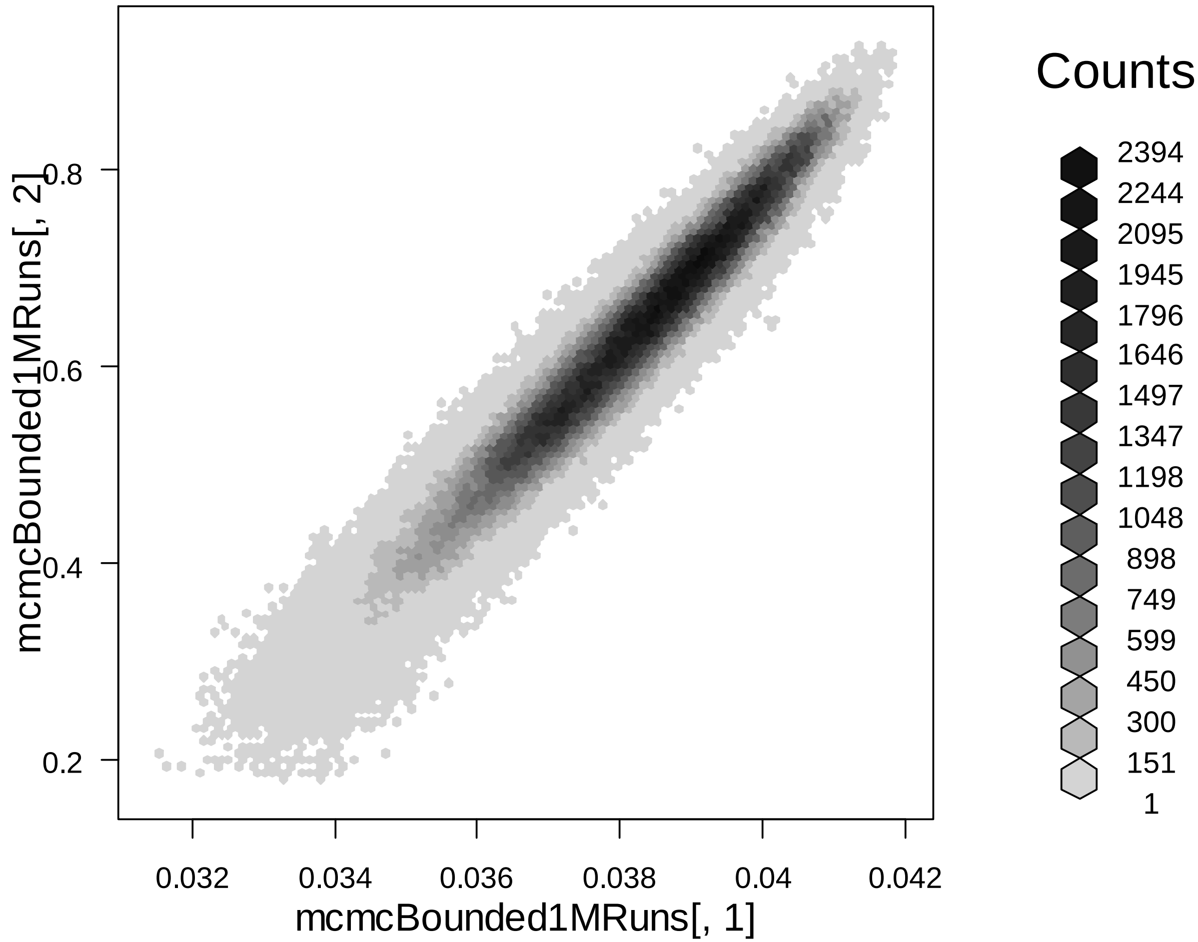
Rough Markov Chain Operation

- The Markov Chain generation process involves
  - Drawing a candidate value from θ (e.g. just generated from random perturbation around last generated sample
  - Assessing (relative) posterior of this candidate value of θ
    - Run the simulation model
      - Required b/c formulated likelihood statements for $P(y|\theta)$ generally depend on simulation model output (for that value of parameters θ)

Takes significant work!

    - Compute relative posterior of candidate θ using $P(y|\theta)P(\theta)$
      - $P(\theta)$ is just the value of the prior at this point
      - $P(y|\theta)$ is in generate calculated using not just θ and y, but also outputs and intermediate values of simulation model produced for value of θ
    - Accept or reject this candidate θ based on ratio between this relative posterior at θ vs. last generated sample
      - If accepted, release this as our new sample!
      - If rejected, emit existing value and go on to the next candidate!
- Run through "burn in" time before using results

2D Projection of Density from Higher Dimensional Sampling



Conceptually Rich Topics

- Linking dynamic simulation with statistical tools
  - Predictor-corrector models
  - Posterior derivation
- Linking models & data
- Multi-scale modeling
- Real-time predictive simulation
- Model specification & domain specific languages
- Mathematical analysis tools
- Numerical analysis considerations
- Alternative mathematical formalisms (hybrid automata, differential Equation variants, DAEs, etc.)
- Performance
- Eval. study design & stats w/synth. pop experiments

Performance

- Model scaling
- Concurrency
  - Problem decomposition
  - Embarrasing parallelism
- Language impact on parallelizability

Other ABM Modeling Frameworks

- ABM
  - Repast Simphony
  - Netlogo
- Others
  - SD: Vensim/Powersim/Ithink/Berkeley Madonna
  - Insightmaker
  - http://www.systemswiki.org/index.php?title=Health_Care_System_Dynamics_Insights

Books

- ABM
  - Railsback, S.F. and Grimm, V. *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton: Princeton University Press. ISBN 978-0-691-13674-5. 2012
  - Epstein, J.. *Computational Social Science.*
- Networks
  - *Newman, M. Structure & Dynamics of Networks*
  - *Valente. Social Networks in Health.*
  - *Popular:*
    - *Strogatz,S. Sync*
    - *Barabasi. Linked.*
    - *Watts. Small Worlds.*
- System Dynamics
  - J. Morecroft.
  - Sterman, *Business Dynamics*.
  - Ford. *Modeling the Environment.*

Conferences

- Winter Simulation Conference
  - With 2012 hosting the first meeting of the AnyLogic Users' Group
- Microsimulation International
- International Conference of the System Dynamics Society
- Social Computing, Behavioral-Cultural Modeling and Prediction (SBP)
- International Health Informatics Symposium (IHI)

Community Contributed Content

- AnyLogic Users Groups
- Health regions of SystemsWiki