Overview of the Modeling Process

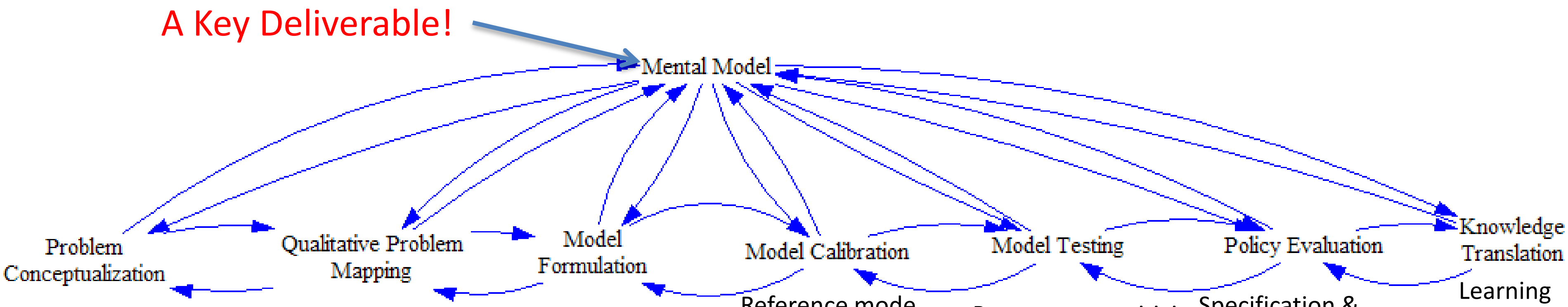
Nathaniel Osgood

Using Modeling to Prepare for Changing Healthcare Needs
Duke-NUS
April 16, 2014

Overview of Modeling Process

- Typically conducted with an interdisciplinary team
- An ongoing process of refinement
- Best: Iteration with modeling, intervention implementation, data collection
- Observation:
  - Traditionally, the focus in ABM has been on insights gained from the model delivered
  - Often it is the *modeling process* itself – rather than the models created – that offers the greatest value

Modeling Process Overview



Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogenous, exogenous, ignored)

Key entities & their relationships

- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

State charts
Stocks & Flows
Process flow operators
Parameter & state variables
Qualitative State charts
Influence
Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Specification of

- Parameters
- Quantitative causal relations
- Flow formulas
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Specification & investigation of intervention scenarios

Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Learning environments (e.g. DISimS)

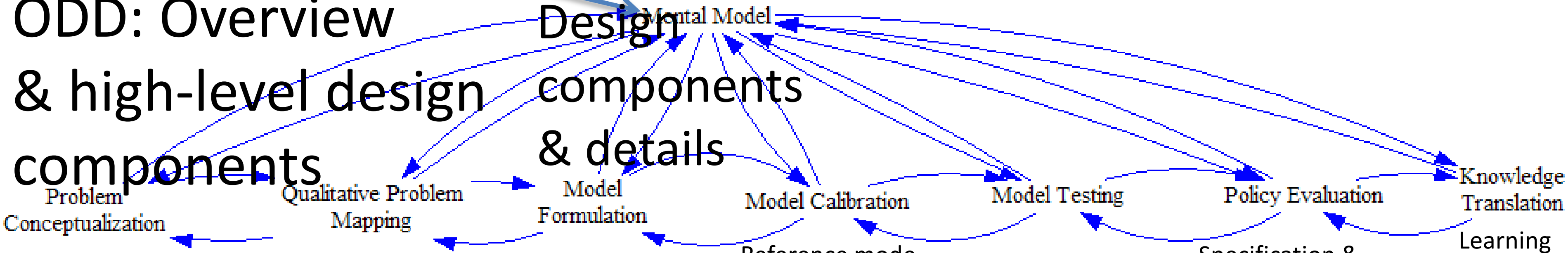
Visualizations

Modeling Process Overview

A Key Deliverable!

ODD: Overview & high-level design components

ODD: Design components & details



Problem Conceptualization

Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogenous, exogenous, ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

Qualitative Problem Mapping

State charts
Stocks & Flows
Process flow operators
Parameter & state variables
Qualitative State charts
Influence
Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Model Formulation

Specification of
- Parameters
- Quantitative causal relations
- Flow formulas
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Model Calibration

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Model Testing

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered
  - Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Policy Evaluation

Specification & investigation of intervention scenarios

Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Knowledge Translation

Learning environments (e.g. DISimS)

Visualizations

Incremental Model Development

- Great advantages are conferred by building a simulation model in a step-by-step fashion
- With each iteration, the model is modified in some small fashion
- A new version of the model is “docked” against older versions of the model
  - Confirming identical behavior when the changes are disabled
  - Understanding behavior with the new feature enabled
- Frequently these incremental versions
  - Can be demonstrated to system stakeholders
  - Produce insight that inform the next step undertaken

Benefits of Incremental Development

- Greater understanding of where model patterns emerge & interactions
- Faster ability to diagnose bugs
- Flexibility to change direction based on learning
- Capacity to secure feedback from stakeholders (e.g. observations of unexpected emergent model patterns, prioritization of issues)
- Greater clarity in prioritization
- More effective time-boxing
- Enhanced stakeholder confidence
- Improved morale

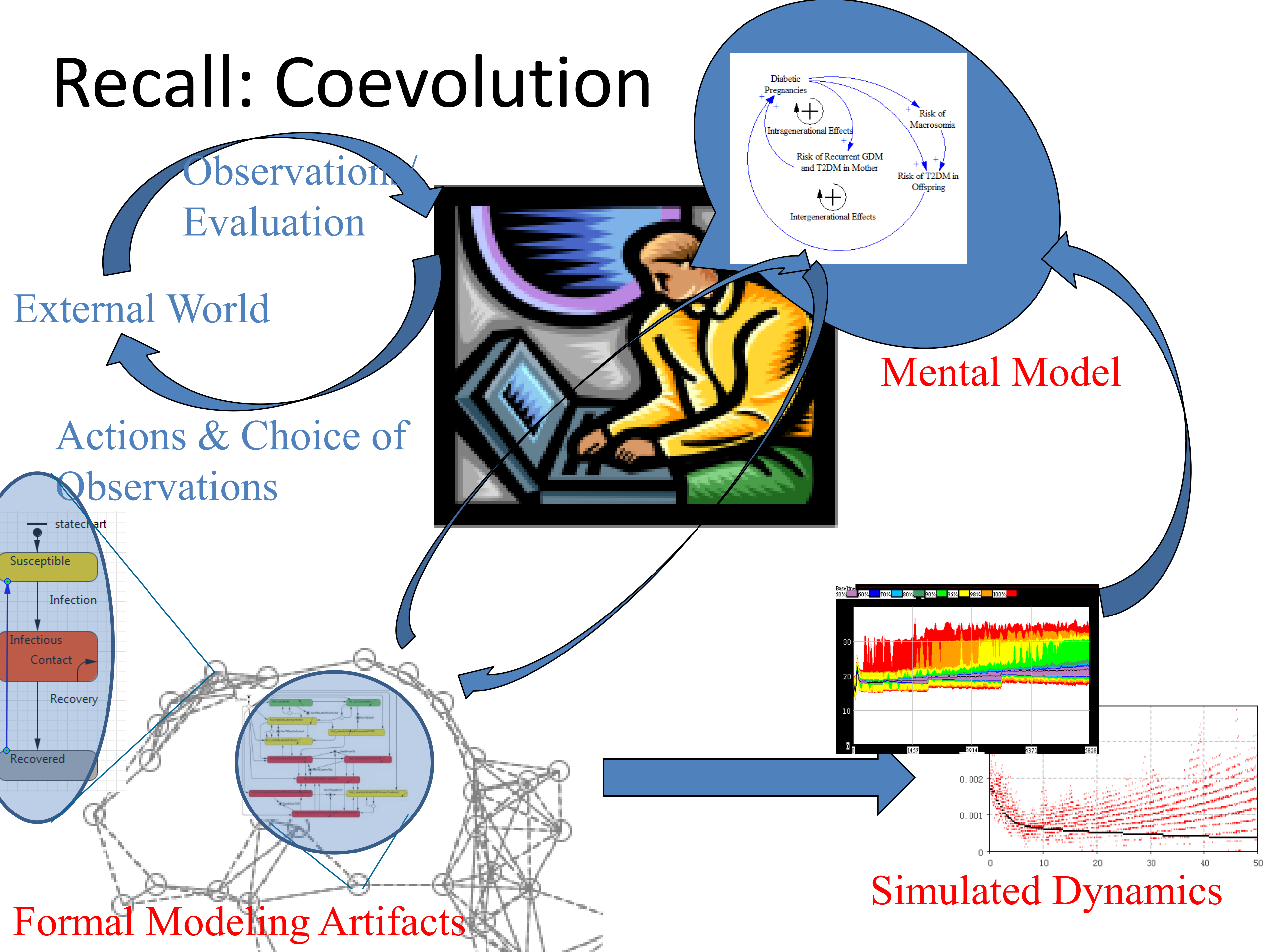
Framework 1: PARTE

- Properties
  - State & characteristics
- Actions
  - Agent interactions with environment (including other agents)
- Rules
  - Characterization of processes affecting agent evolution
- Time
  - Time horizon?
  - Size of timestep or time units
- Environment
  - Spatial & topological (e.g. network) context of agent
  - Dynamics of surrounds

Framework 2: The Overview, Design concepts, and Details (ODD) Protocol for ABM Design

- Consensus protocol derived from panel for ABM modelers
- Primary focus: *Specification* protocol
  - To help understand, communicate & reproduce ABMs
- Secondary benefit: Process for ABM design
- 3 broad components
  - Overview: model goals & high level scope & design
  - Design concepts: Different aspects of design being considered
  - Remaining elements
- Reference: Railsbeck & Grimm

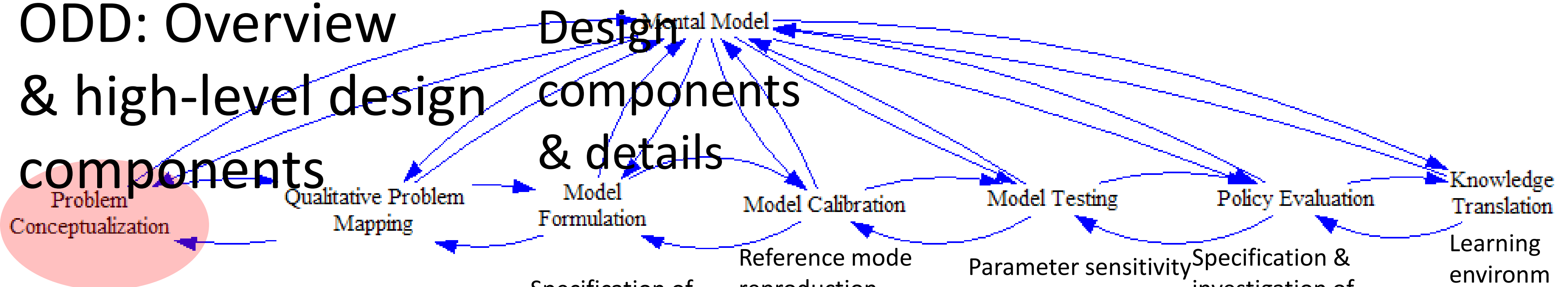
Recall: Coevolution



ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Specification & investigation of intervention scenarios
Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Learning environments (e.g. DISimS)

Visualizations

Identification of Questions/ “The Problem”

- All models are simplifications and “wrong”
- Some models are useful
- Attempts at perfect representation of “real-world” system generally offer little value
- Establishing a clear model purpose is critical for defining what is included in a model
  - Explaining reference modes
  - Understanding broad trends/insight?
  - Understanding policy impacts?
  - Ruling out certain hypotheses?
- Think explicitly about model boundaries
- Adding factors often does not yield greater insight
  - Often simplest models give greatest insight
  - Opportunity costs: More complex model takes more time to build=>less time for insight

Model Purpose & System Dynamics Models

- The smaller vocabulary of System Dynamics models provides practical limits to the level of detail we can pack into our model
- The danger here is often how broad we make this model

Model Purpose & Agent-Based Models

- The flexibility & generality & computational universality of ABM supports the creation of arbitrarily rich models
- It is very easy to add deep levels of detail (e.g., on processes underlying agent behaviour)
  - Adaptive behavior based on context
  - Sophisticated decision rules
  - Many properties/attributes

General Principles

- Typically: $\exists$ high opportunity cost to investing in a given model area: Given limited time, it takes away from richness elsewhere – and often from learning
- Given this flexibility & cost, it is especially critical to wield the "logical knife" of model purpose
- YAGNI (**Y**ou **A**in't **G**onna **N**eed **I**t): Start simple & add as one develops confidence in & understanding of model

Common Division

- Endogenous
  - Things whose dynamics are calculated as part of the model
- Exogenous
  - Things that are included in model consideration, but are specified externally
    - Time series
    - Constants
- Ignored/Excluded
  - Things outside the boundary of the model

Motivations for Including Endogenous Factors

- Maintaining factors as endogenous (rather than pre-specified as exogenous) lends
  - Extra flexibility for more accurately capturing effects of
    - Interventions
    - Alternative exogenous scenarios
  - Greater robustness in the context of changes
  - Support for translations to other contexts
- Keeping greater detail requires more data & implementation work, but allows our models to be translated to other contexts & times

Example: Smoking

- Hard-coded uptake rates
- Uptake rates dependent on prevalence of smoking in the group
- Uptake rate dependent on preferences (including peer pressure, with positive feedback)
- Uptake rates dependent on social networks and preferences

Example: Aggregate Mixing Between Individuals

- hard-coded rate or distribution of contacts between individuals
- term which allows for risk perception

Example

$$\begin{bmatrix} x_{11} & x_{12} & 1-x_{11}-x_{12} \\ x_{21} & x_{22} & 1-x_{21}-x_{22} \\ x_{31} & x_{32} & 1-x_{31}-x_{32} \end{bmatrix}$$

- Mixing Matrix (specifies fraction of population A's contact that occur with populations B & C
- Preference matrix
  - Scales to capture fluctuating population captures relative preference
  - can't specify where to test
- Mobility-based methods with mobility patterns hard-coded
  - this is challenged for interventions which change e.g. mixing opportunities and mobility
- Mobility-based methods with preference-based mobility model

Fecal-Oral Transmitted Zoonoses

- approximate contact as occurring animal-to-animal
  - can't capture effects of interventions that change likelihood of transmission due to environmental factors
  - requires much less data detail
- approximate contact as occurring via spatial disaggregation of environment and hard-coded mobility model
  - need grounded mobility model
  - can capture effects of interventions that
    - clean up fecal material
    - spread fecal material in such a way to degrade it
  - because mobility patterns are hard-coded, can't capture effects of changing shape, character of environment
- approximate contact as occurring via spatial disaggregation of environment and preference-based mobility model
  - now can simulate effects of both
    - interventions that clean up environment
    - new space design

Pattern Oriented Modeling

- ABMs occupy an arbitrarily rich model space
- To gain insight, it important to leverage the broad set of information we know about a system
  - We need to accompany general model purpose by a broad set of *patterns to be explained*
- Patterns (often called "stylized facts") may include e.g.
  - Similar to classic System Dynamics reference modes
    - Quantitative time series patterns
    - Qualitative (e.g. presence of oscillations, rising, asymmetries, etc.)
  - Patterns of heterogeneity (disparities, stratification, deg. dist)
  - Spatial/topological patterns (waves, clustering, phenomenology)
  - Multi-scale phenomena
- ***We seek a model that will explain (or at least exhibit consistency with, "stay true to") such patterns***

The Value of Patterns in Building a Model

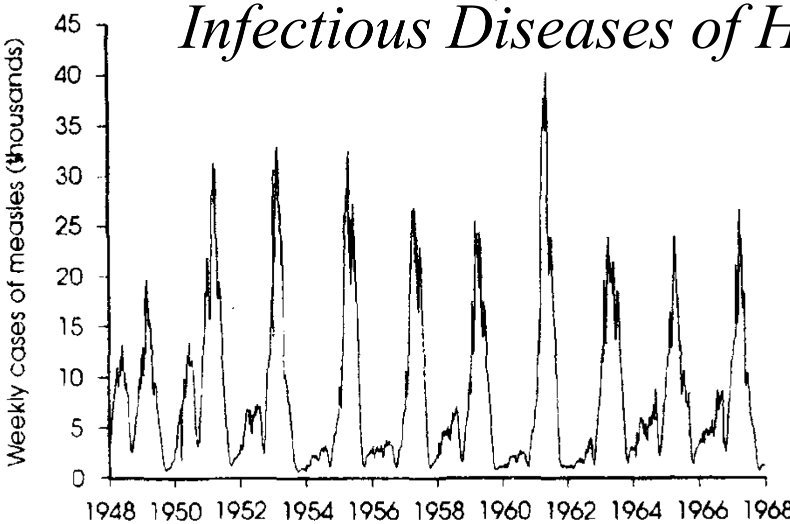
- Patterns are pieces of information which – if the model didn't match them – it would cast suspicion on the model
- Typically specific to the purpose (if goal were different, we'd use a different set of patterns)
- Try to rule out possible submodels using patterns
- Try to use very broad set of knowledge
  - Even if a given pattern is "weak" in constraining the model (e.g. Ng rates higher among women than men), a set of such weak patterns can collectively greatly constrain possible dynamic hypotheses (ABM structure)

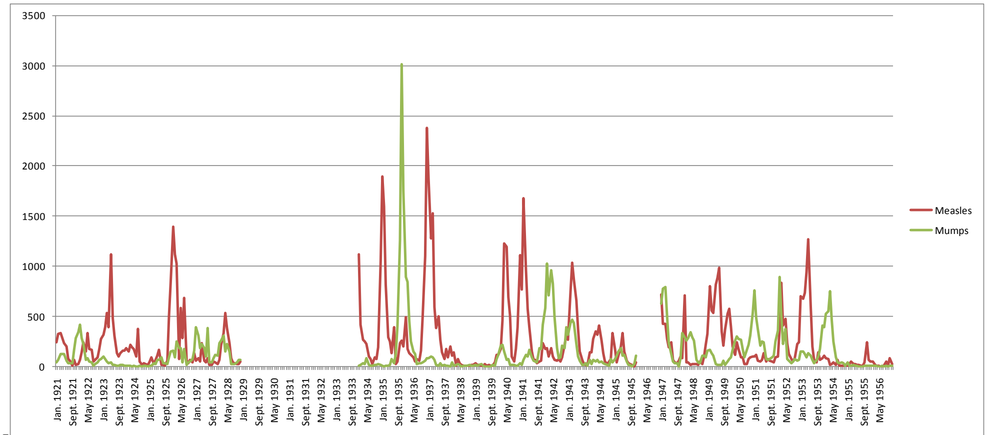
Patterns in Confidence Building

- When inspecting model results, we will seek to recognize other patterns (not built into model or used to judge it) & use them for cross-validation
- Example patterns from a model
  - Prior spatial distribution emerging from movement patterns
  - Contact patterns emerging from individual movement
  - Case-contact network structure emerging from contact tracing process

Example Patterns: Temporal Patterns

Source: Anderson & May
Infectious Diseases of Humans





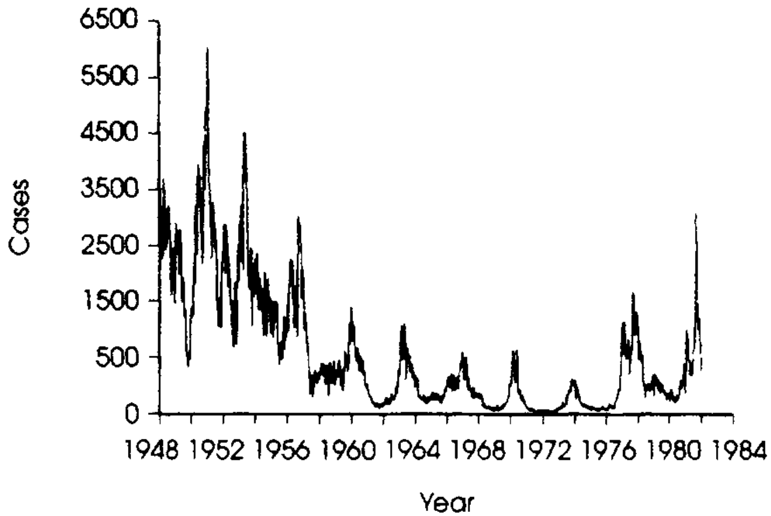


Fig. 6.8. Weekly case notifications of pertussis (whooping cough) in England and Wales for the time period 1948–82. Mass vaccination was introduced in 1956.

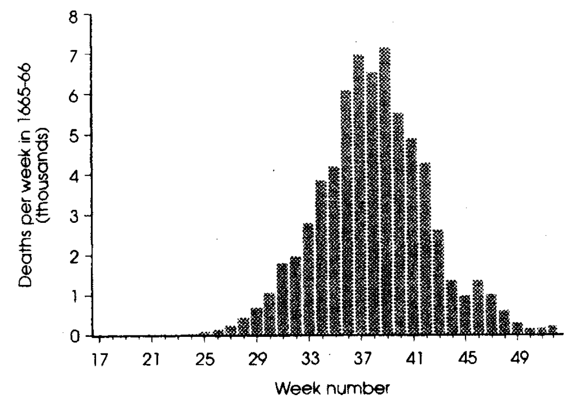
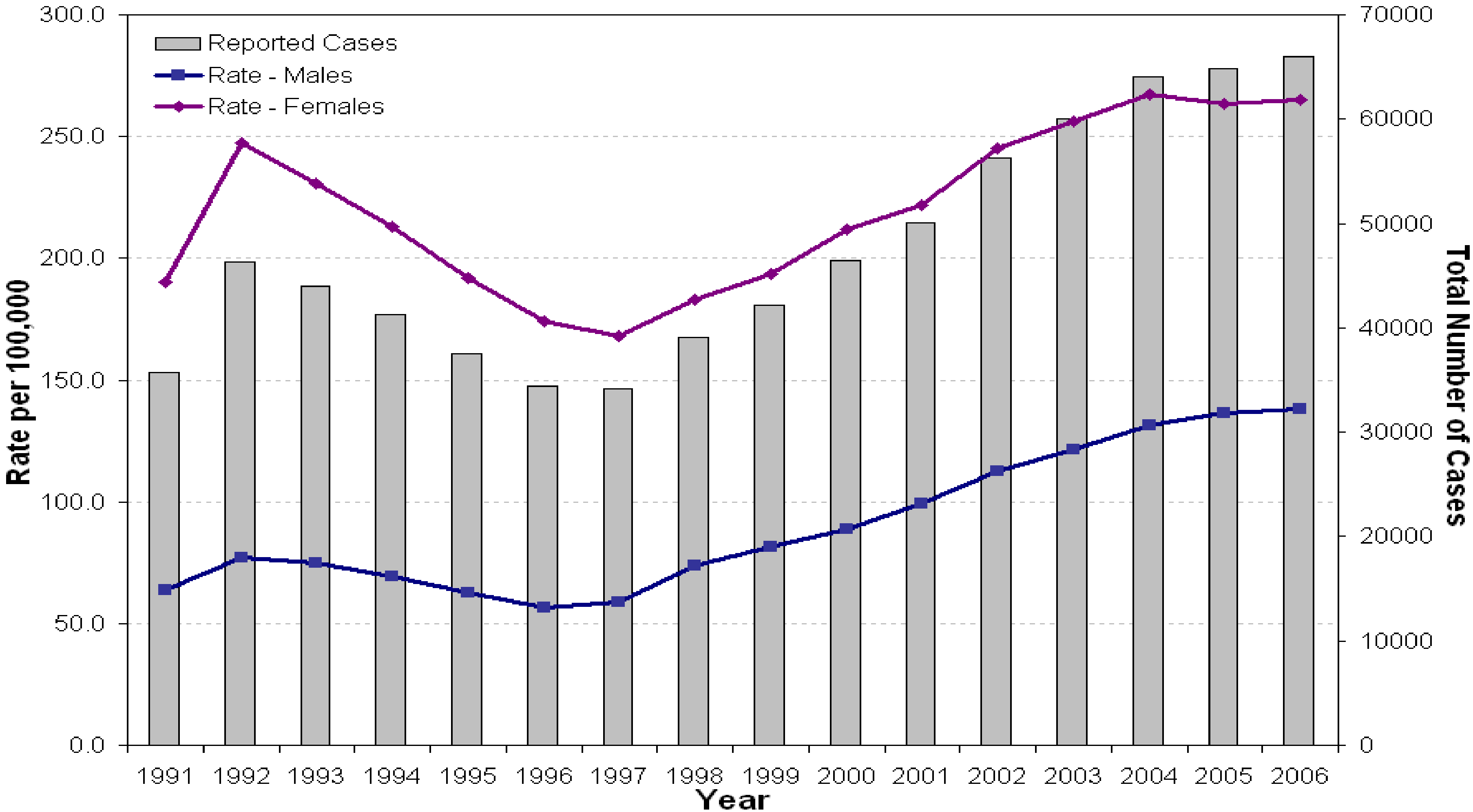


Fig. 3.11. Recorded deaths from the bubonic plague in London during the year (data from Brayley 1722).

Example Patterns: Temporal Patterns



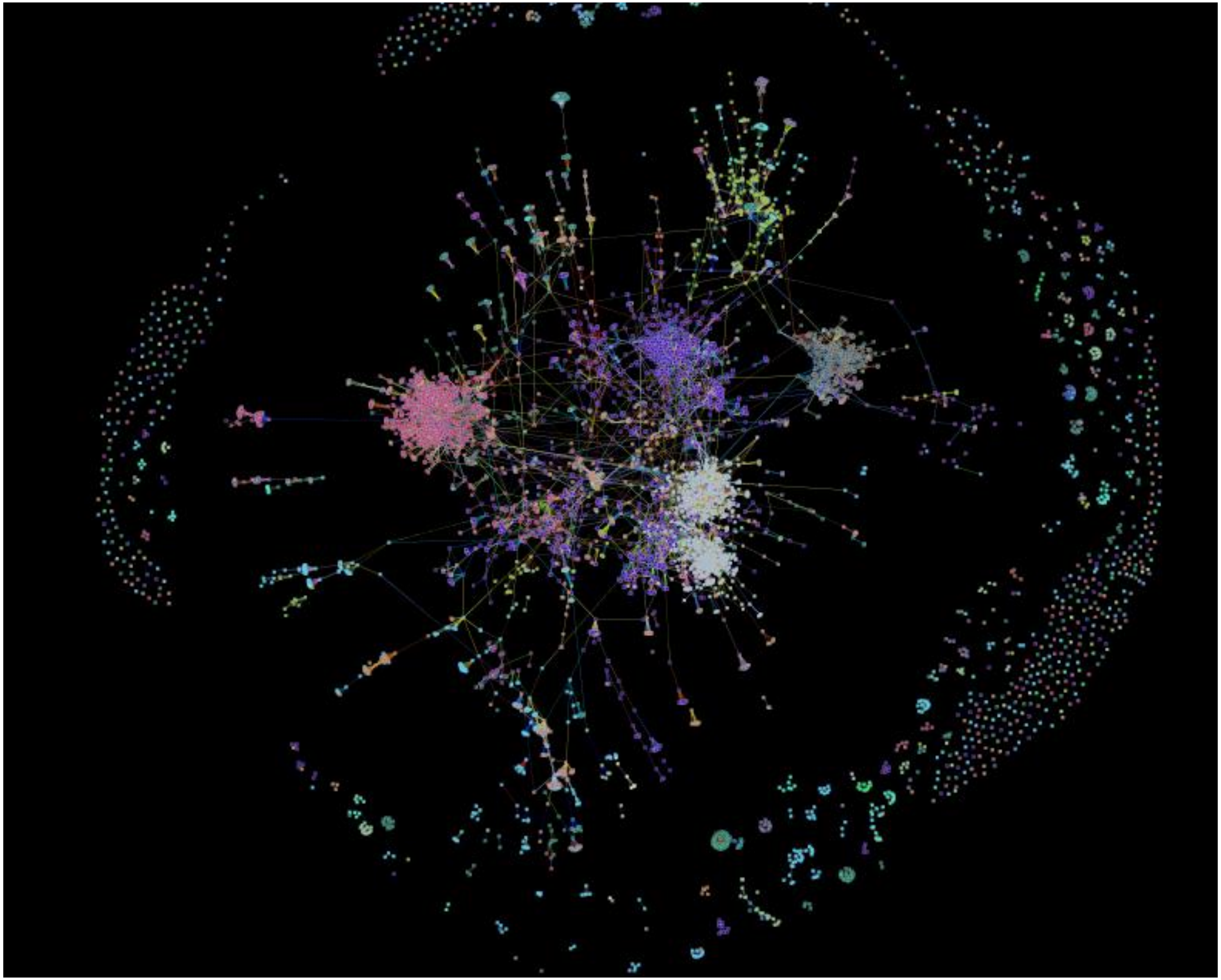
- Data for 2005 and 2006 are preliminary and are anticipated to change
- Source: Surveillance and Epidemiology Unit, Community Acquired Infections Division, PHAC

Example Patterns: Longitudinal Individual Data

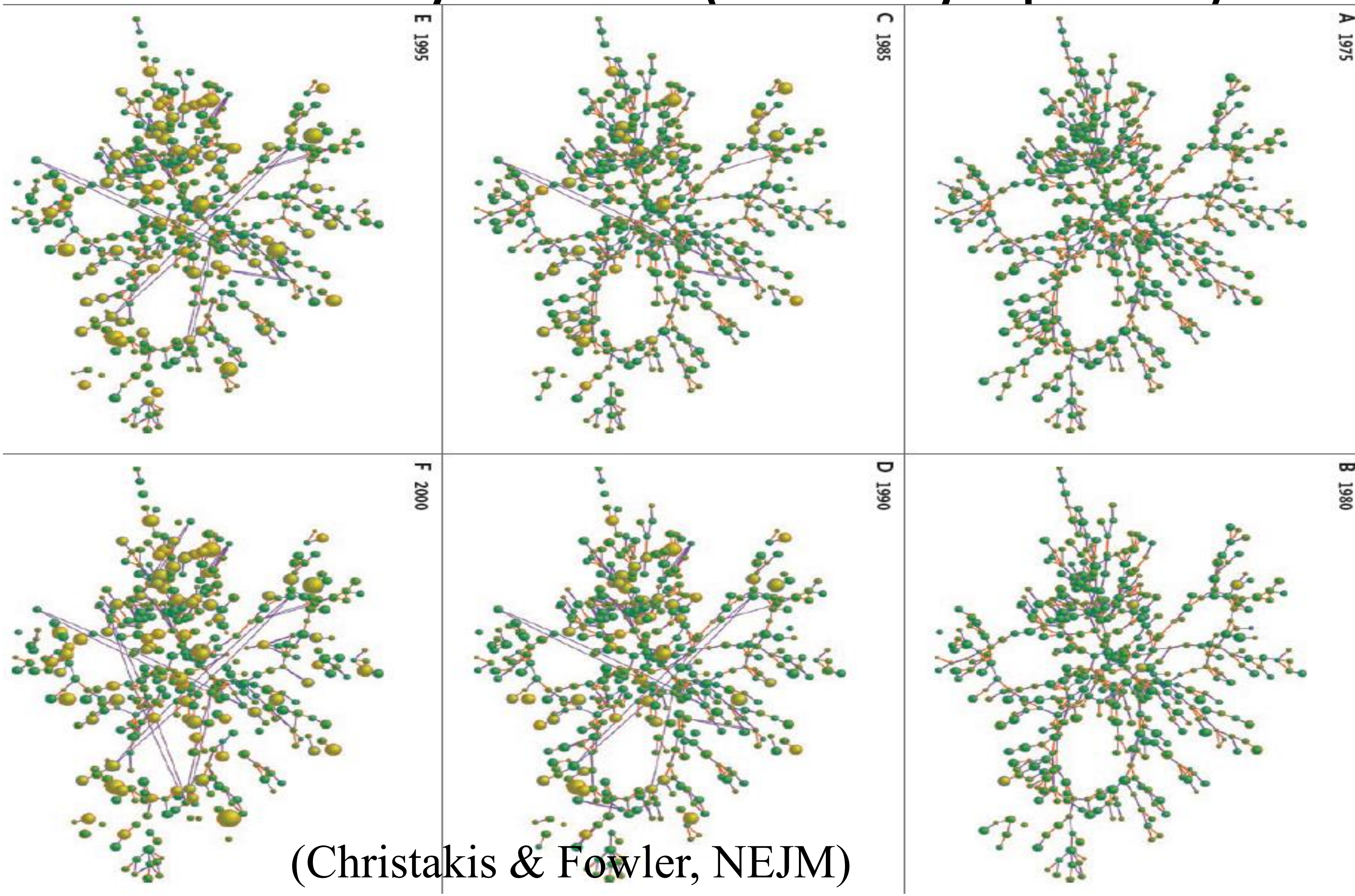
Such data can be very awkward to capture or to compare to in an aggregate model

Interval From Latent to TB	Cumulative Percentage in 1972 for Non-Indian (%)	Cumulative Percentage in 1972 for Indian (%)
<0.5 year	31.8	36.4
0.5-1 year	42.1	50
1-2 years	52.6	63.6
2-3 years	78.9	68.2
3-4 year	84.2	72.2
4-5 year	89.5	86.4
5-6 year	-	90.9
6-7 year	-	95.5
7-8 year	-	1
8-9 year	-	1
>9 year	1	1

Example Patterns: Emergent Network Structure



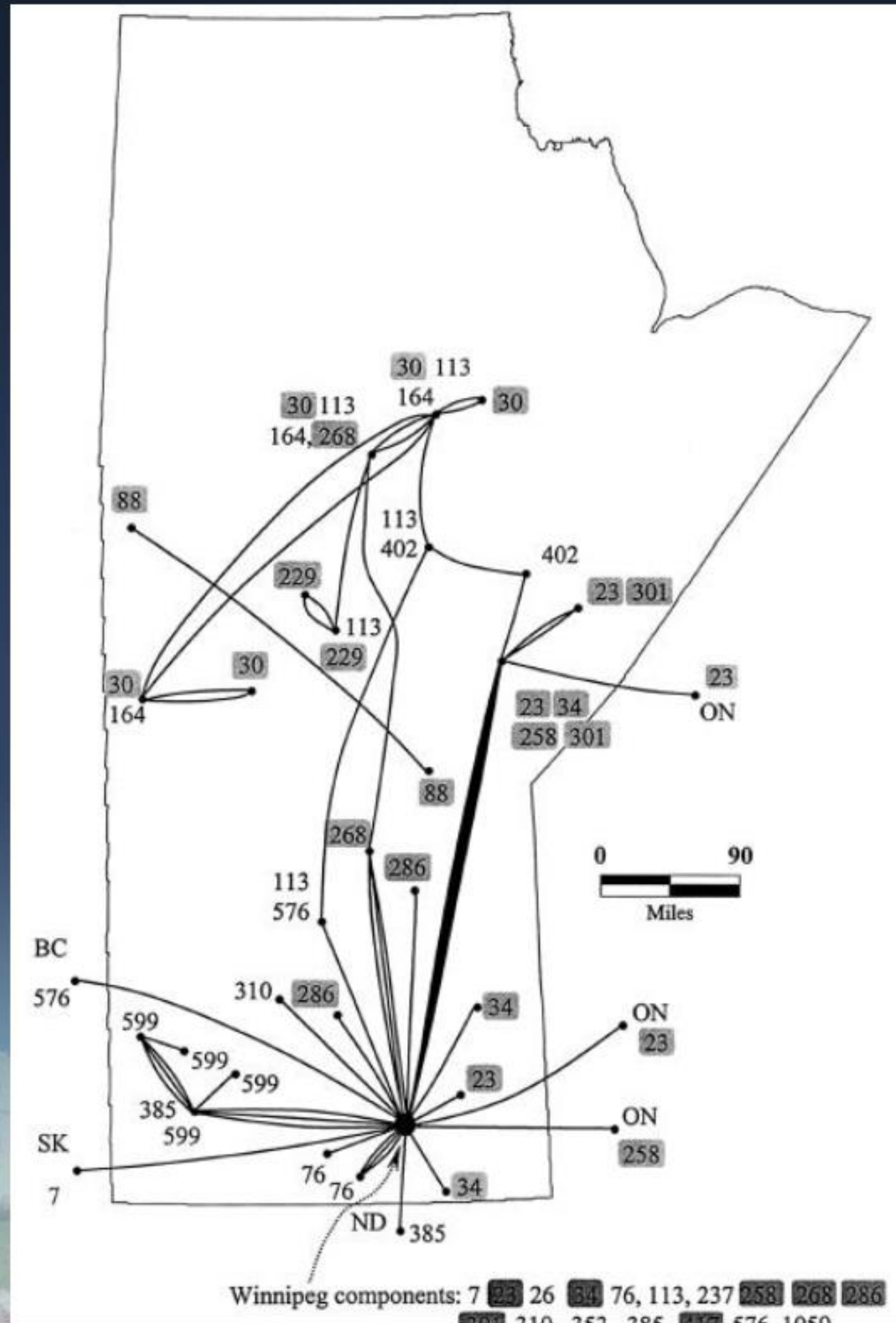
Example Patterns: Network Dynamics (Obesity Spread)



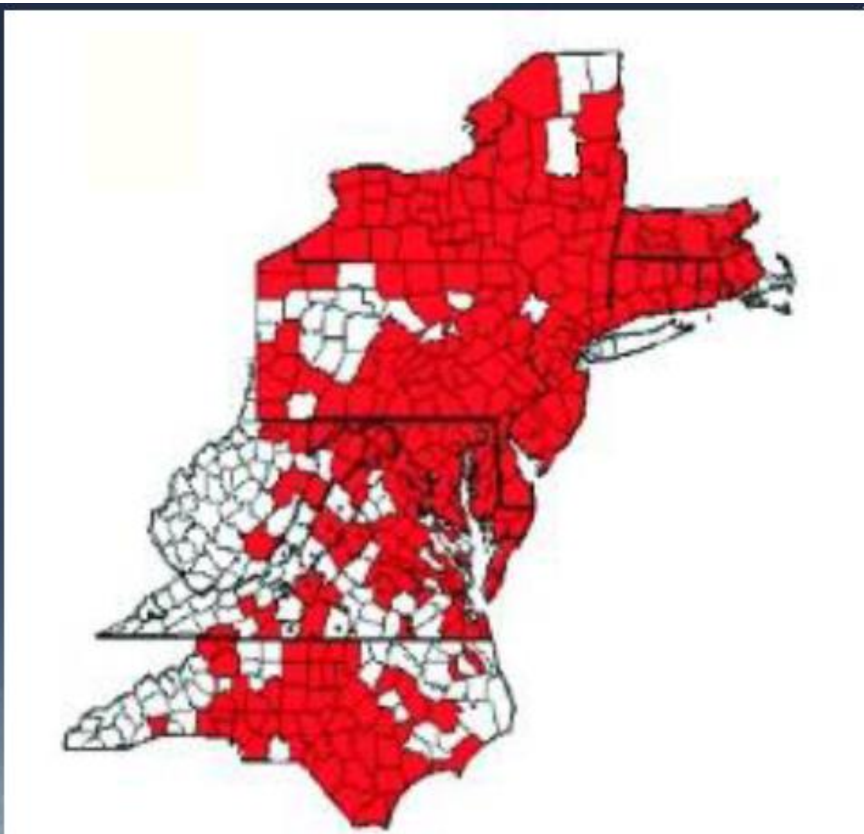
(Christakis & Fowler, NEJM)

Chlamydia & Gonorrhea in Manitoba

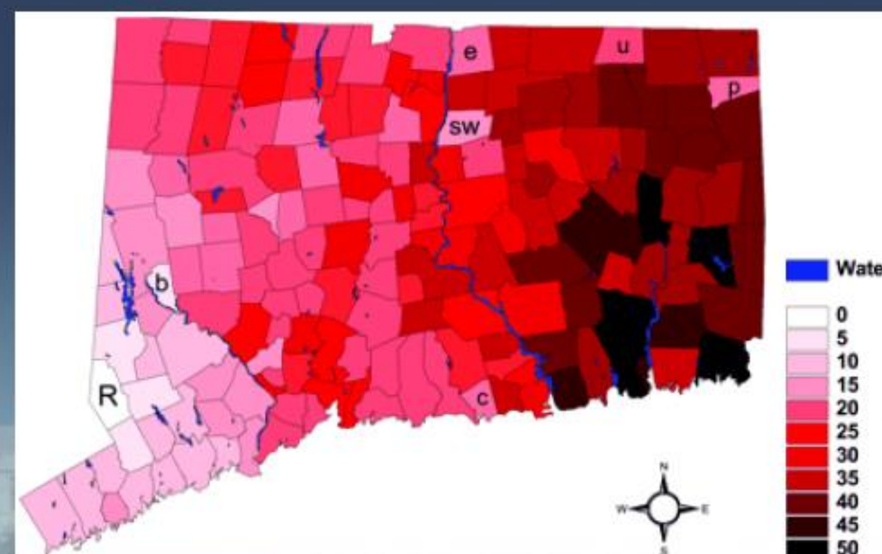
Wylie, J.L, Jolly, A. Patterns of Chlamydia and Gonorrhea Infection in Sexual Networks in Manitoba, Canada Sex Transm Dis. 2001 Jan;28(1):14-24.



Example Patterns: Spatial Spread (Rabies)

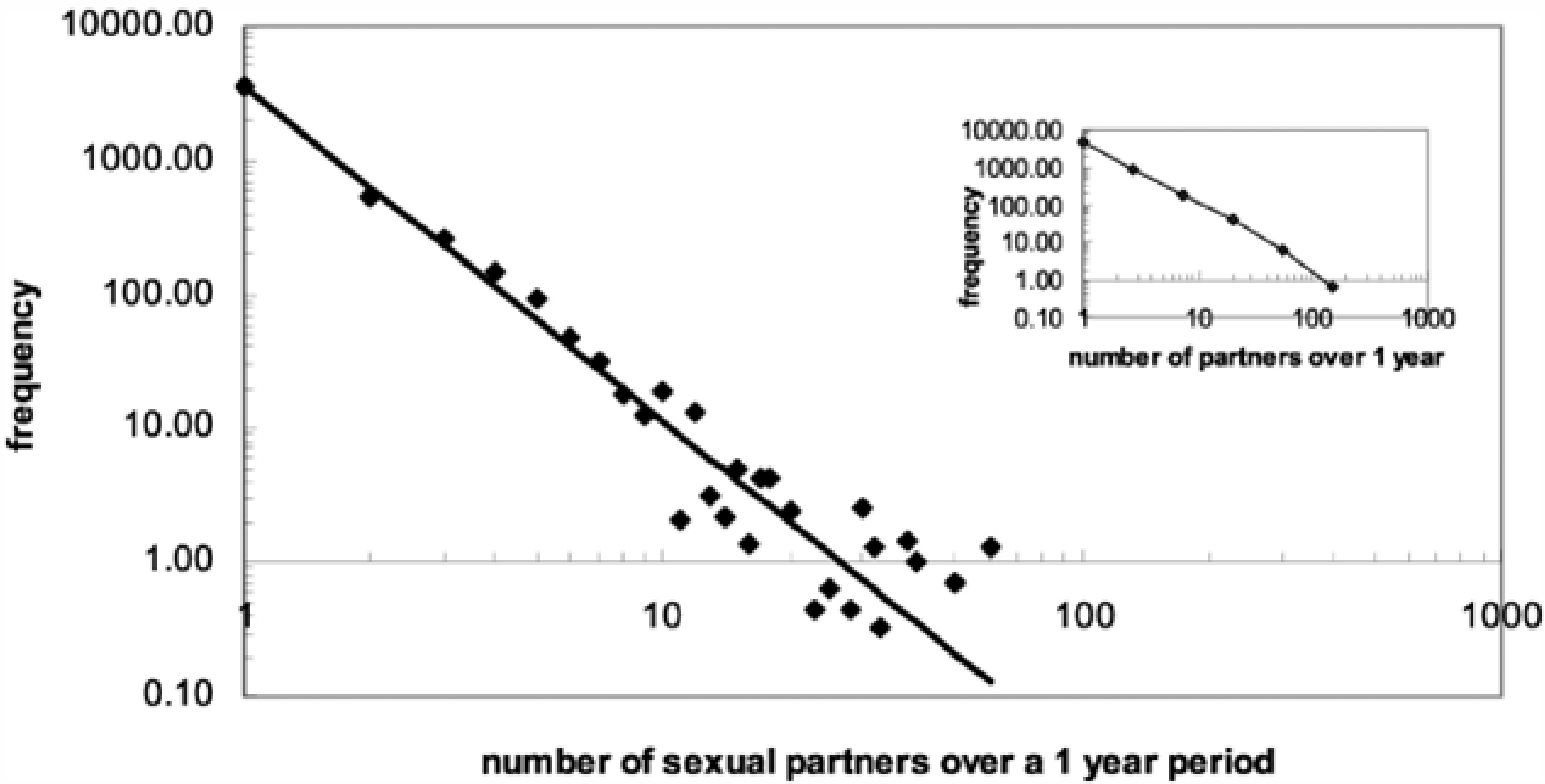


Marta A. Guerra,* Aaron T. Curns,* Charles E. Rupprecht,* Cathleen A. Hanlon,* John W. Krebs,* and James E. Childs*
Skunk and Raccoon Rabies in the Eastern United States: Temporal and Spatial Analysis.
Emerg Infect Dis. 2003 September; 9(9): 1143–1150.
doi: 10.3201/eid0909.020608



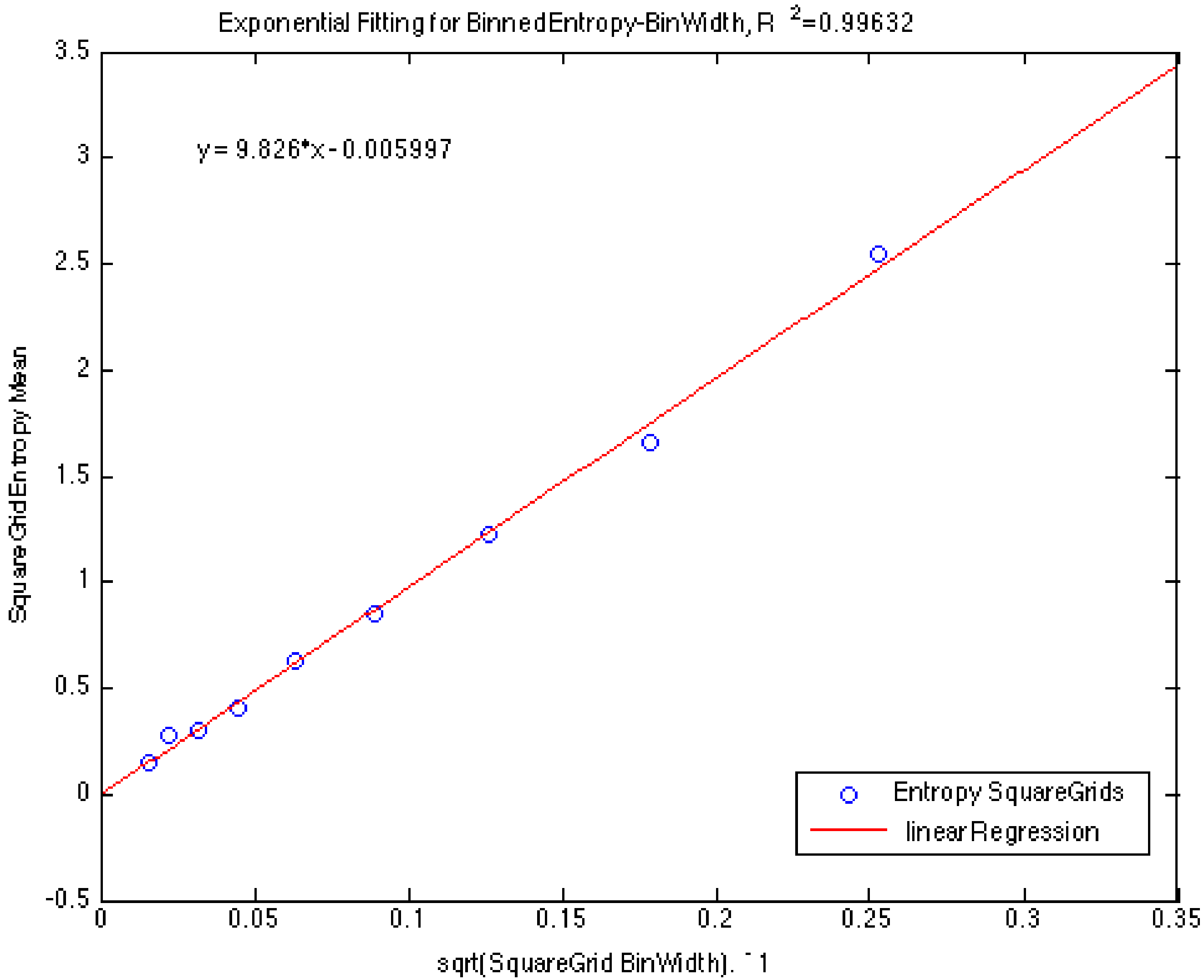
David L. Smith*†, Brendan Lucey‡, Lance A. Waller § , James E. Childs¶, and Leslie A. Real
Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. PNAS March 19, 2002 vol. 99 no. 6 3668-3672

Example Patterns: Distributions&Scaling Relationships



Schneeberger et al., Scale-Free Networks and Sexually Transmitted Diseases: A Description of Observed Patterns of Sexual Contacts in Britain and Zimbabwe , Sexually Transmitted Diseases, June 2004, Volume 31, Issue 6, pp 380-387

Example Patterns: Distributions&Scaling Relationships

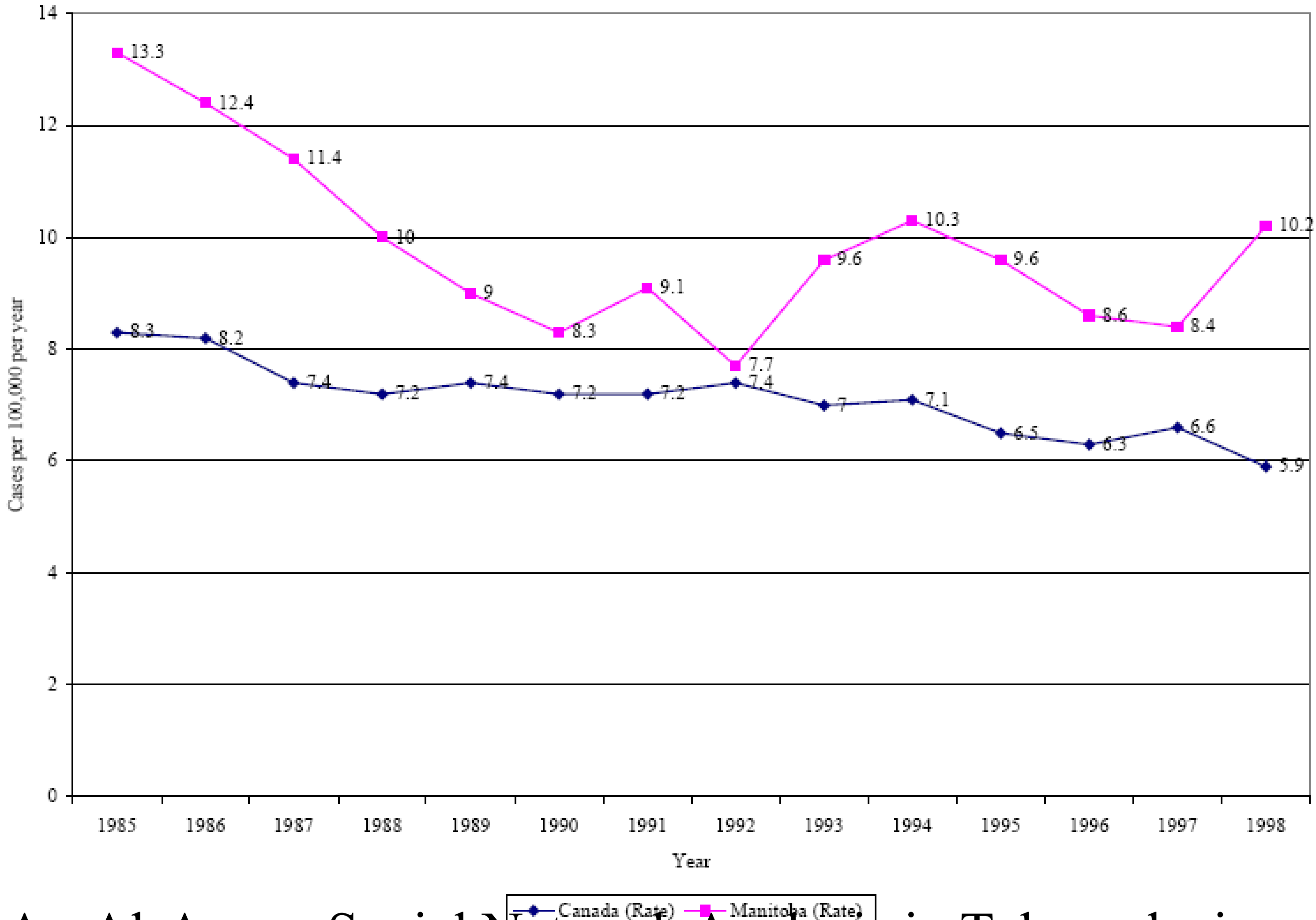


Example Patterns: Empirical Distributions

Network Degree	TB		Mantoux Positive		Mantoux Negative		Total
	(N)	(%)	(N)	(%)	(N)	(%)	
All	68	13.5	109	21.6	327	64.9	504
2	45	36.8	35	28.7	42	34.4	122
3	28	62.2	10	22.2	7	15.6	45
4	15	68.2	7	31.8	0	0	22
5	14	77.8	4	22.2	0	0	18
6	9	90	1	10	0	0	10
7	7	100	0	0	0	0	7
8	7	100	0	0	0	0	7

A. Al-Azem, Social Network Analysis in Tuberculosis Control Among the Aboriginal Population of Manitoba Doctoral Dissertation, University of Manitoba, 2006

Example Patterns: Multi-Scale Models: TB



A. Al-Azem, Social Network Analysis in Tuberculosis Control Among the Aboriginal Population of Manitoba Doctoral Dissertation, University of Manitoba, 2006

Example Patterns: Multi-Scale

Figure 38: Degree Centrality Graph of Community 1 TB group network (people and places)

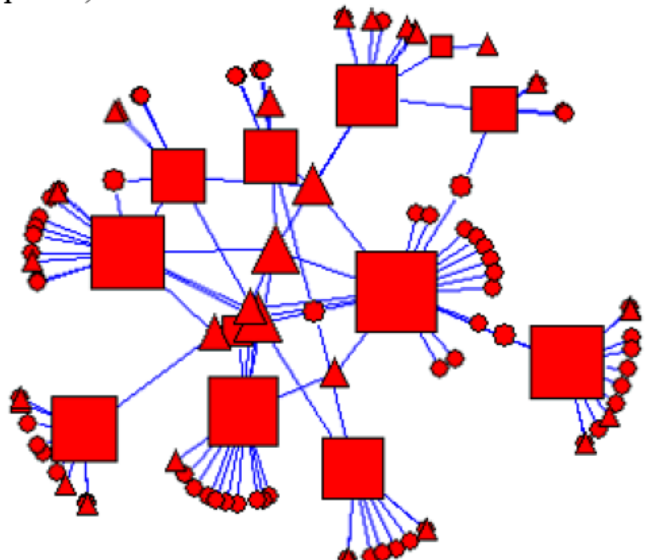
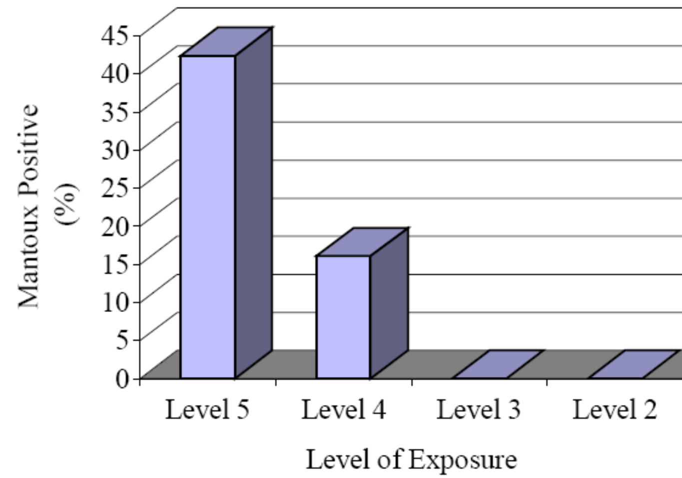


Figure 4: Trend of Mantoux positivity (M+v) by degree of exposure in Community 1



A. Al-Azem, Social Network Analysis in Tuberculosis Control Among the Aboriginal Population of Manitoba Doctoral Dissertation, University of Manitoba, 2006

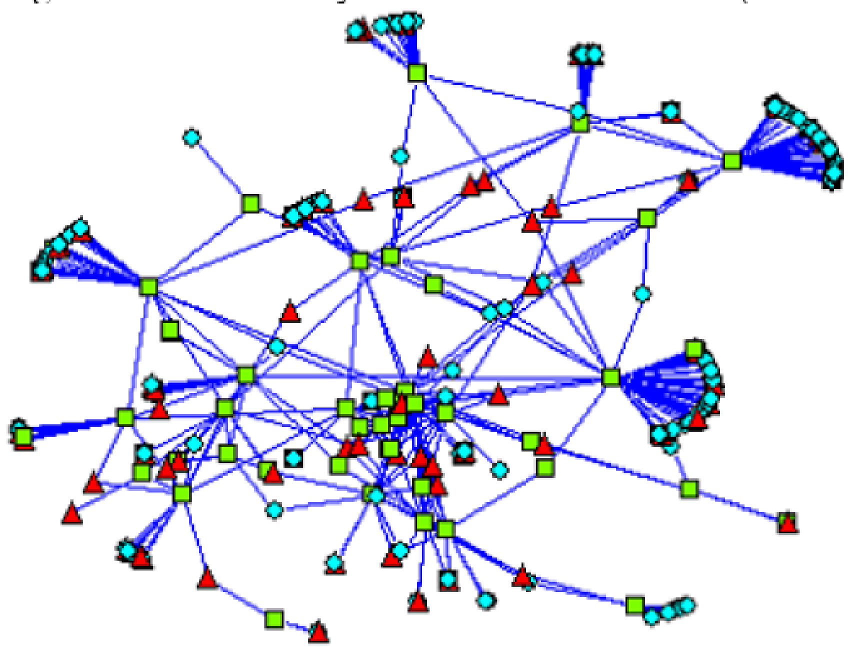
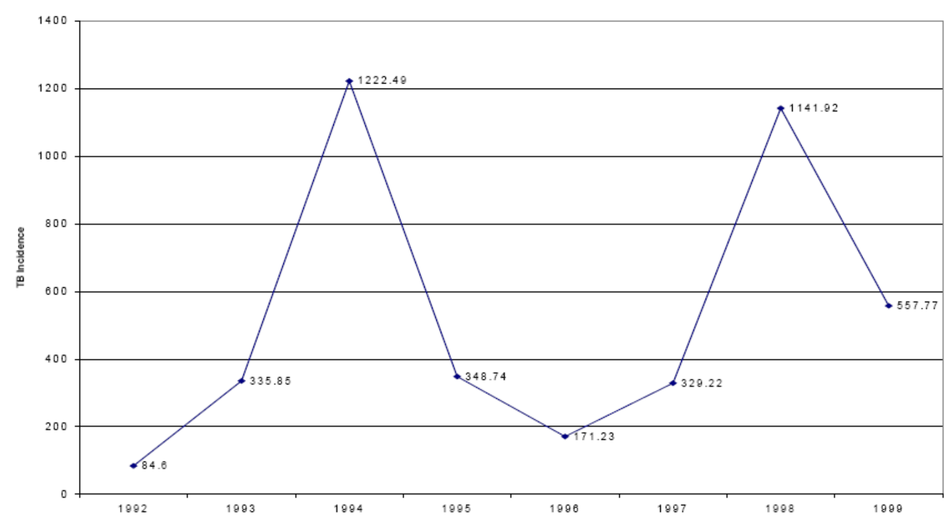


Figure 14: TB incidence in Community 1, 1992-1999 (Cases per 100,000 by year)



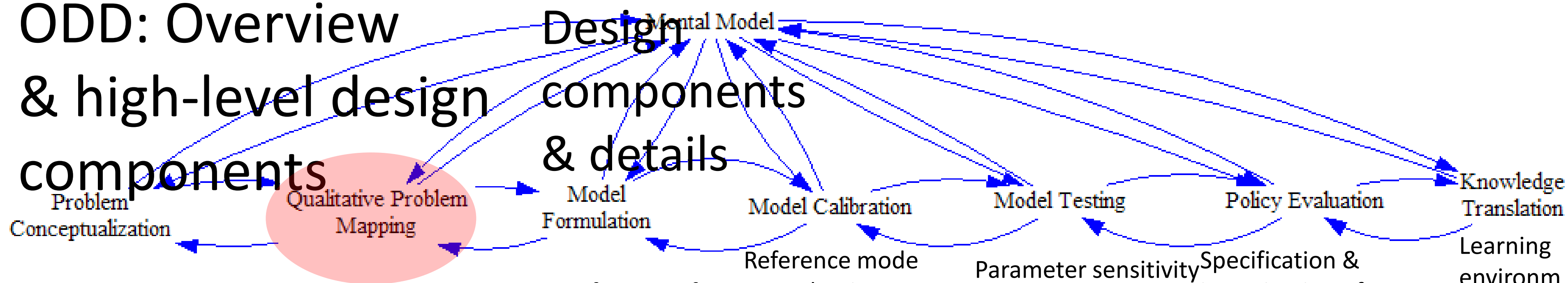
Example Phenomenological Patterns

- Flocking
- Oscillation
- Gradients
- Waves
- Cascaded transitions over time
- Phase change phenomena
- Clustering
- “Waves” of topological spread/precolation
- Punctuated equilibria

ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem Conceptualization

Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

Qualitative Problem Mapping

State charts

Parameter & state variables

Qualitative State charts

Influence & Causal loop diagrams

Multi-agent interaction diagrams

Process flow structure

Key events

Rough model time & spatial extent

Observers to measure outputs

Model Formulation

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Model Calibration

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Model Testing

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered
  - Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Policy Evaluation

Specification & investigation of intervention scenarios

Investigation of hypothetical external conditions

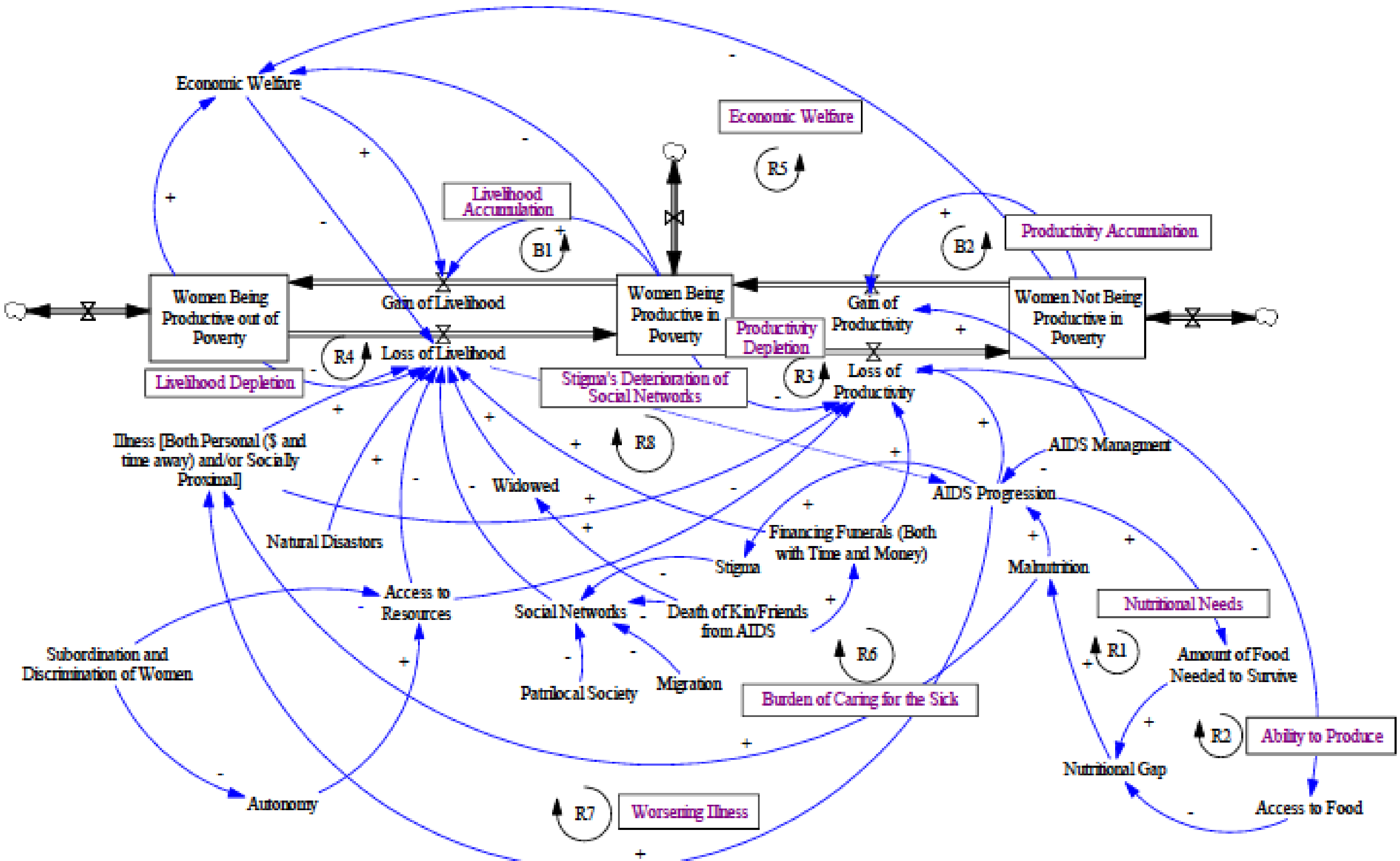
Cross-scenario comparisons (e.g. CEA)

Knowledge Translation

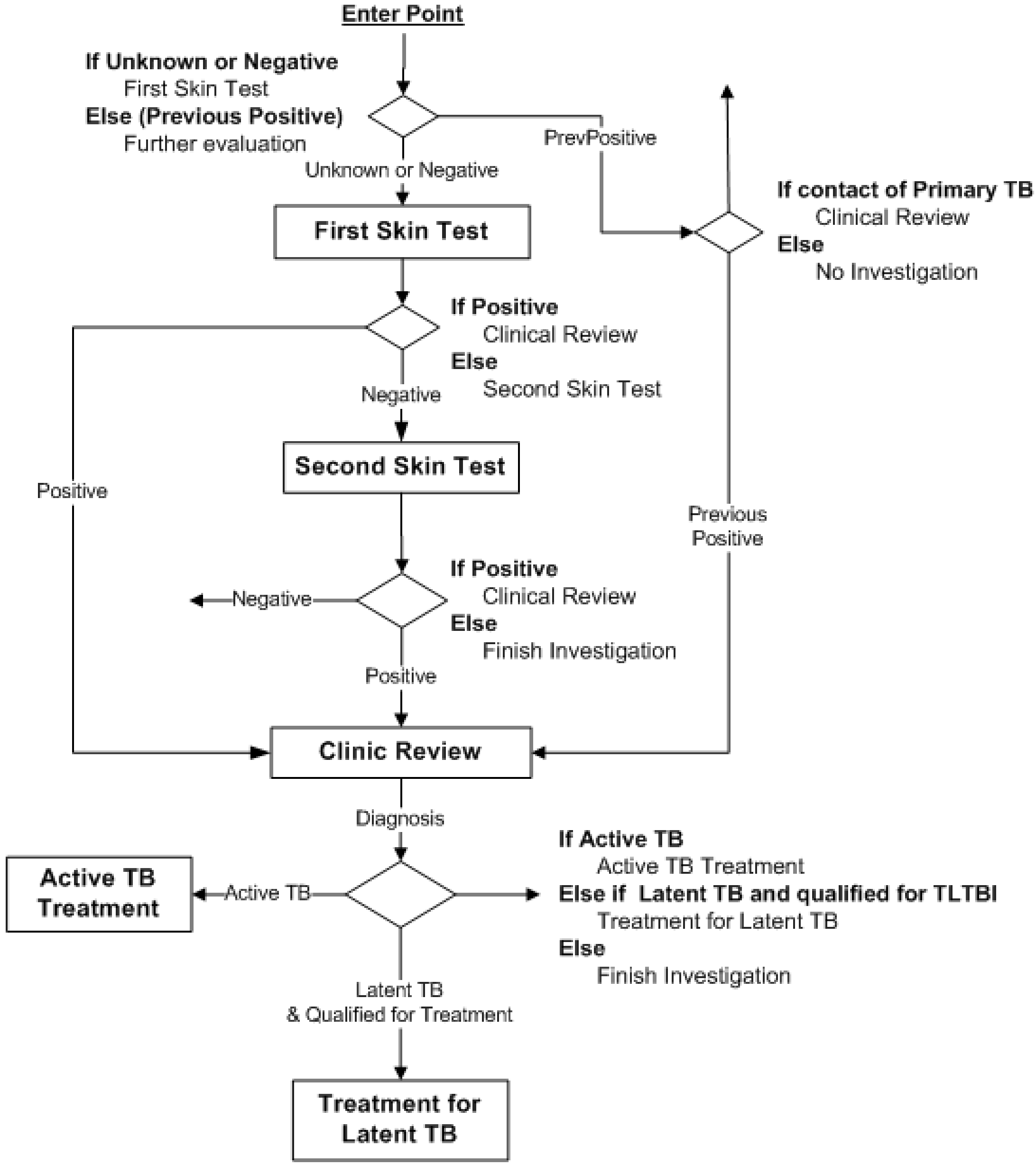
Learning environments (e.g. DISimS)

Visualizations

Problem Mapping: Qualitative Models (System Structure Diagram)



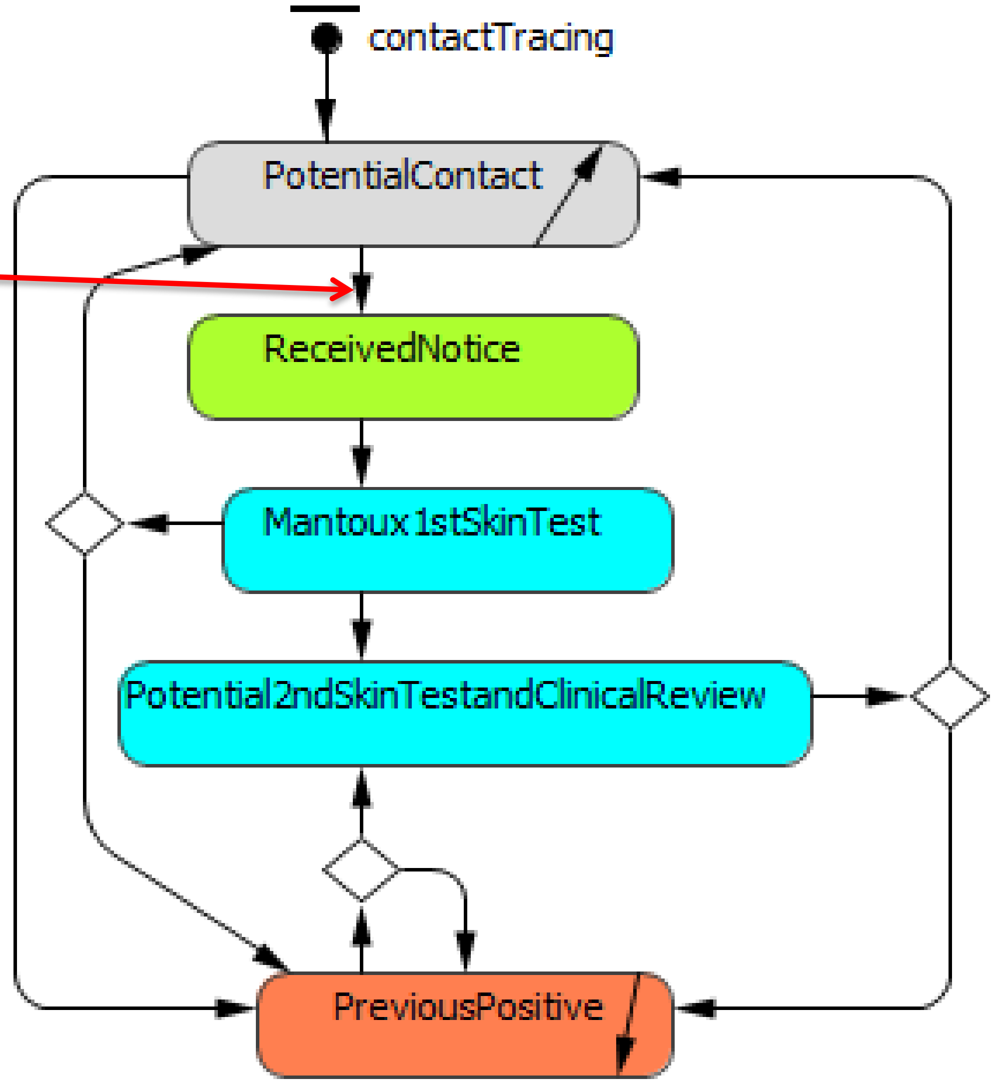
Headley, J., Rockweiler, H., Jogee, A. 2008. Women with HIV/AIDS in Malawi: The Impact of Antiretroviral Therapy on Economic Welfare, Proceedings of the 2008 International Conference of the System Dynamics Society, Athens, Greece, July 2008.



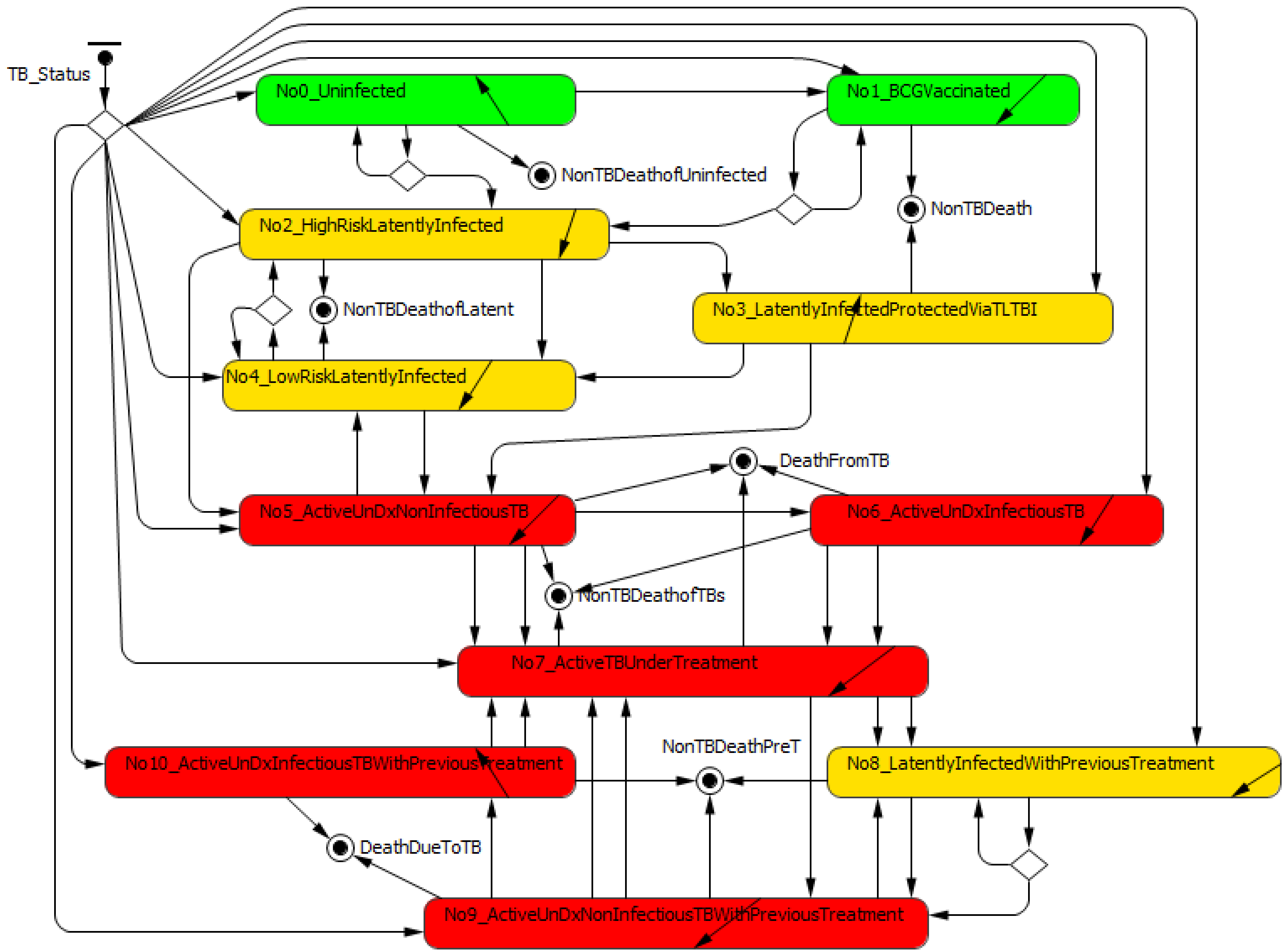
Note: Contact tracing for primary case is different. Since the target is to find the source of infection, once the presumed source is found, contact tracing is discontinued. If the contact is not positive in a TST test, then no need for the 2nd skin test.

State Diagrams to Document State Progression

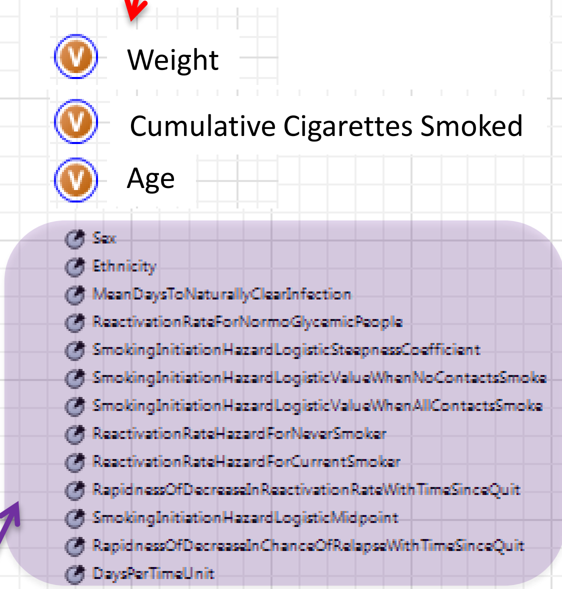
Qualitative Transitions (no likelihood or governing rules yet unambiguously specified)

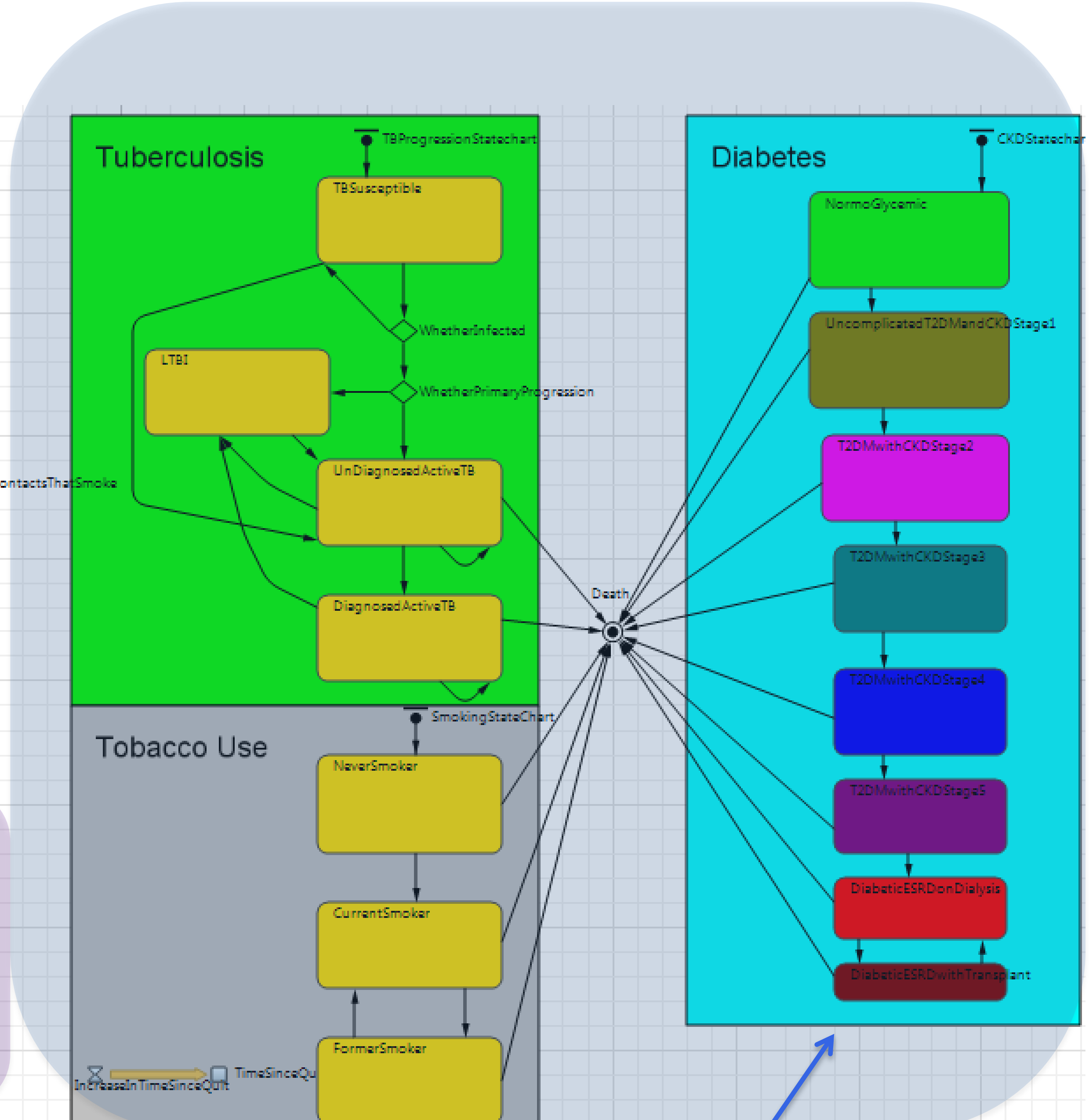


State Diagram 2



These variables are aspects of *state*.

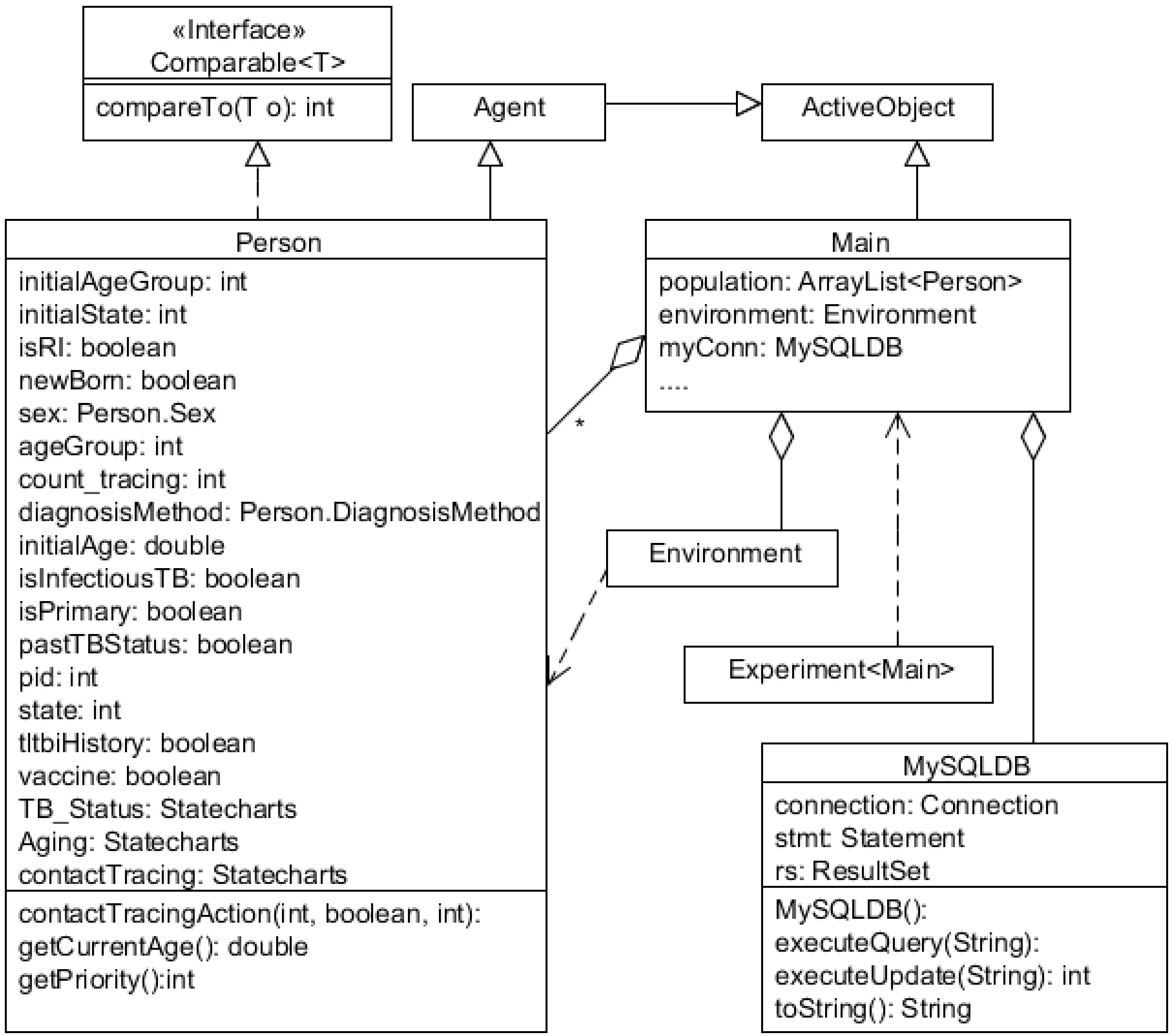




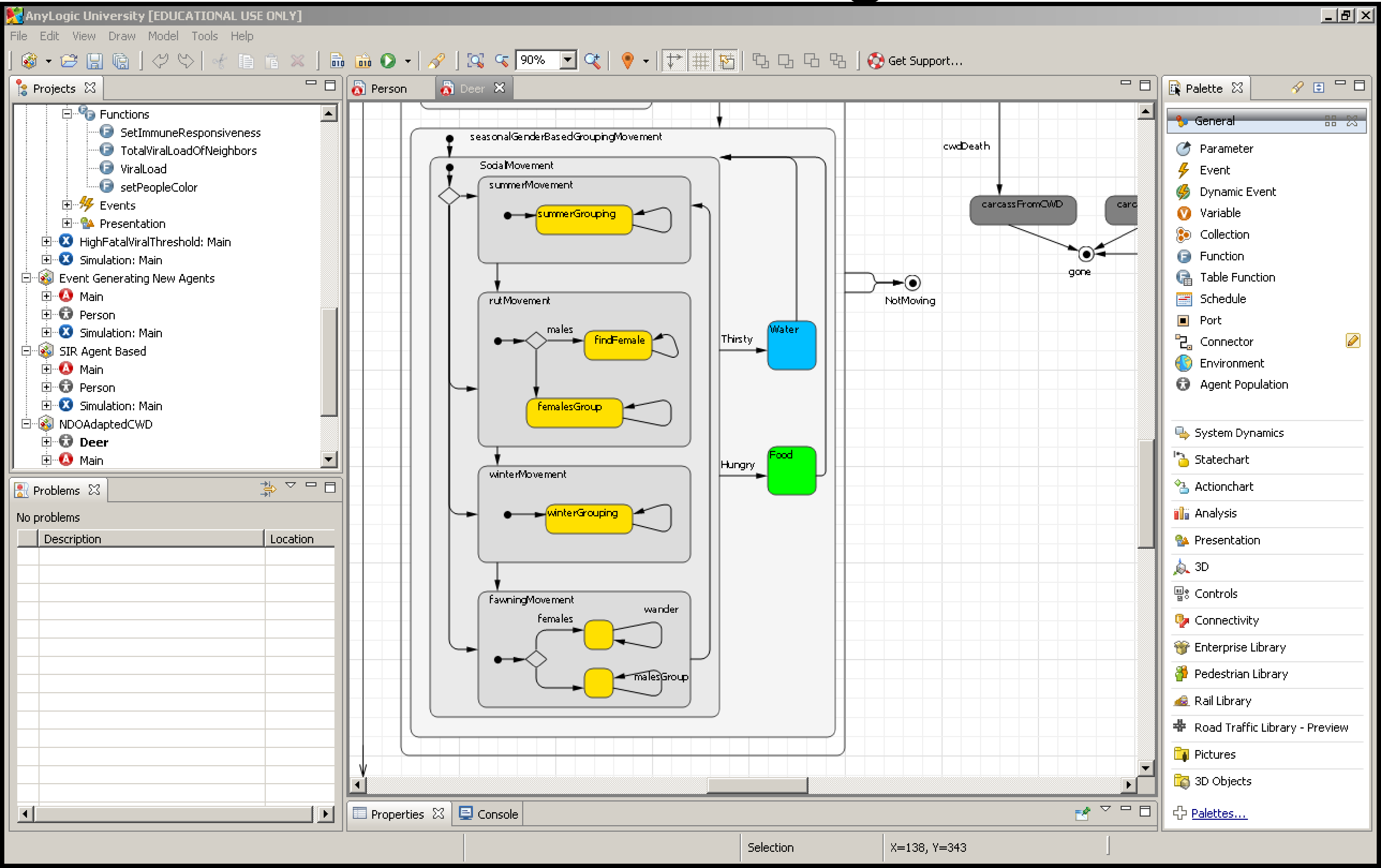
These "parameters" give static characteristics of the agent

These describe the "behaviours" – the mechanisms that will govern agent dynamics

Documenting Agent Characteristics in UML
UML Class Diagram



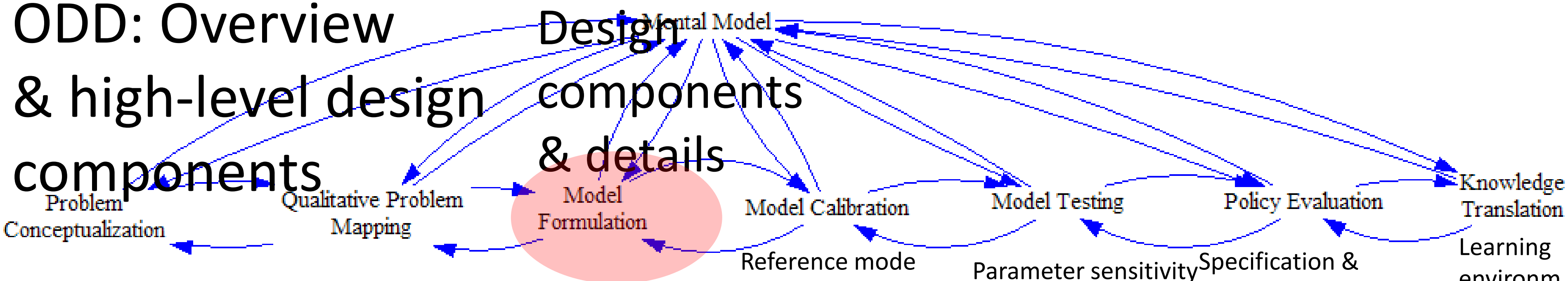
UML State Diagram



ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Specification & investigation of intervention scenarios

Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Learning environments (e.g. DISimS)

Visualizations

Model Formulation

- Model formulation elaborates on problem mapping to yield a fully specified, quantitative model
- Key missing ingredients: Specifying unambiguous specification for
    - Statechart transitions
    - Flows (in terms of other variables)
    - Observer processes
    - Intermediate variables
  - Parameter values

Model Specification Mechanisms

Stock & Flow Models: "Hedgehog Knowledge"

- Small modeling vocabulary
- Power lies in combination of a few elements
- Analysis conducted predominantly in terms of elements of model vocabulary

Agent-Based Modeling: "Fox Knowledge"

- Large modeling vocabulary
- Different subsets of vocabulary used for different models
- Power in flexibility & combination of elements
- Variety in analysis focus

ABMs: Larger Model Vocabulary & Needs

- Events
- Multiple mechanisms for describing dynamics
  - State diagrams
  - Stock and flow
  - Custom update code
- Inter-Agent communication (sending & receiving)
- Multiple types of transitions
- Diverse types of agents
- Spatial & topological connectivity & patterning
- Subtyping
- Mobility & movement
- Graphical interfaces
- Data output mechanisms
- Stochastics complicated
  - Scenario result interpretation
  - Calibration
  - Sensitivity analysis
- Synchronous & asynchronous distinction, concurrency

ODD Design Concepts to Consciously Consider

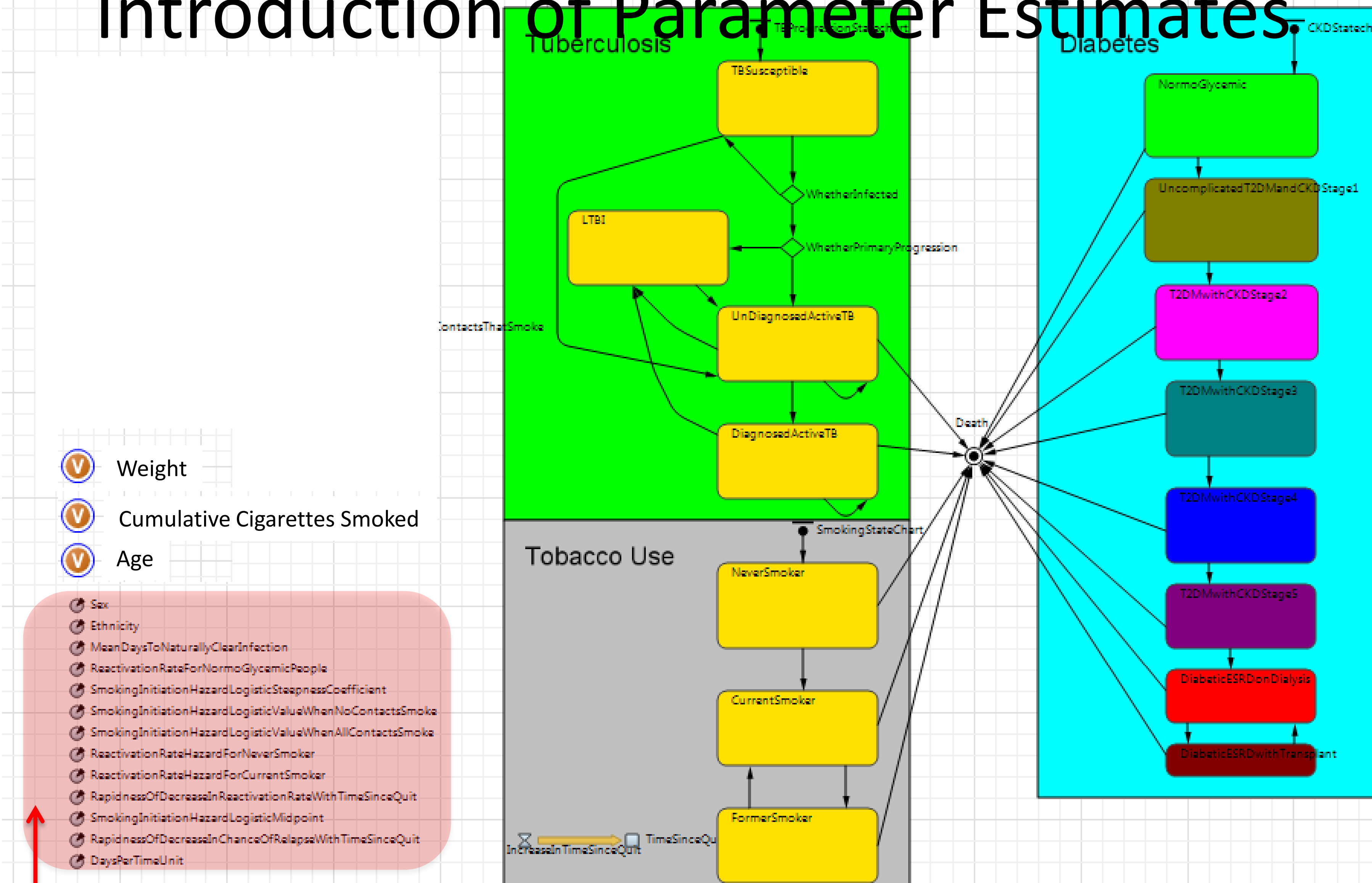
- Origin & character of basic principles underlying model
- Emergence: To what degree are results pre-programmed vs. arising naturally out of a myriad of interactions
- Adaptation: How does system evolution lead to entity behavior change?
- Sensing: What information do entities retrieve from world?
- Objectives: Any goal seeking behavior? How interacts w/state?
- Learning: How does experience drive change in strategies?
- Prediction: How do entities anticipate the future?
- Interaction: How do entities interact directly & indirectly?
- Stochastics: Character of & motivation for stochastic effects
- Observation: What information & associated processes are required for operational use or for testing & confidence bldg

Common Sources for Parameter Estimates (Health)

- Surveillance data
- Controlled trials
- Outbreak data
- Clinical reports data
- Intervention outcomes studies
- Calibration to historic data
- Expert judgement
- Systematic reviews

Parameter*	Description	Baseline value (units)	Reference
μ	Entry/exit of sexual activity	0.0056 ($years^{-1}$)	Garnett and Bowden, 2000
c	Partner change rate per Susceptible	16.08 ($years^{-1}$)	Approximated from Garnett and Bowden, 2000
β	Probability of infection per sexual contact	0.70	Garnett and Bowden, 2000
φ	Fraction of Infectives who are symptomatic	0.20	Garnett and Bowden, 2000
$1/\gamma$	Latent period	0.038 (years)	Brunham et. al., 2005
$1/\sigma$	Duration of infection	0.25 (years)	Brunham et. al., 2005
θ	Asymptomatic recovery coefficient	1.5	Garnett and Bowden, 2000
$1/\pi$	Duration of naturally-acquired immunity	1 (year)	Approximated from Brunham et. al., 2005

Introduction of Parameter Estimates

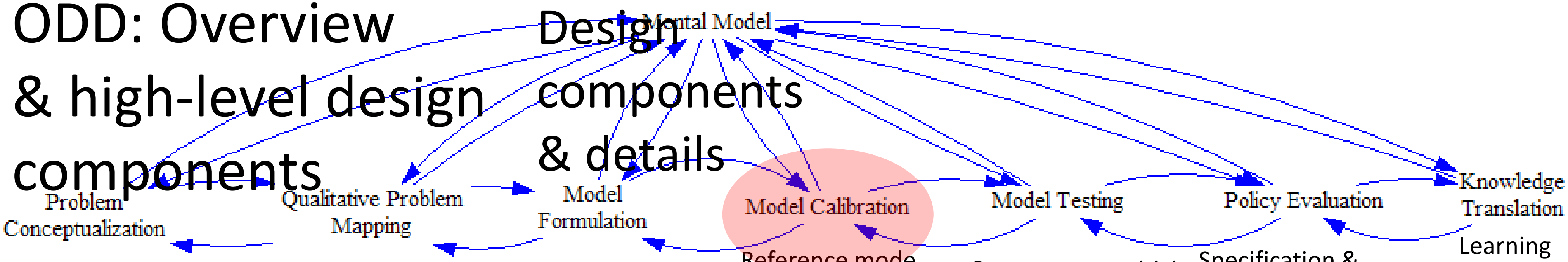


These parameters must have constants specified

ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem Conceptualization

Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

Qualitative Problem Mapping

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Model Formulation

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Model Calibration

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Model Testing

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Policy Evaluation

Specification & investigation of intervention scenarios

Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Knowledge Translation

Learning environments (e.g. DISimS)

Visualizations

Calibration

- Often we don’t have reliable information on *some* parameters
  - Some parameters may not even be observable!
- Some parameters may implicitly capture a large set of factors not explicitly represented in model
- Often we will calibrate less well known parameters to match observed data
  - “Analytic triangulation”: Often try to match against *many* time series or pieces of data at once
- Sometimes we learn from this that our model structure just can’t produce the patterns!

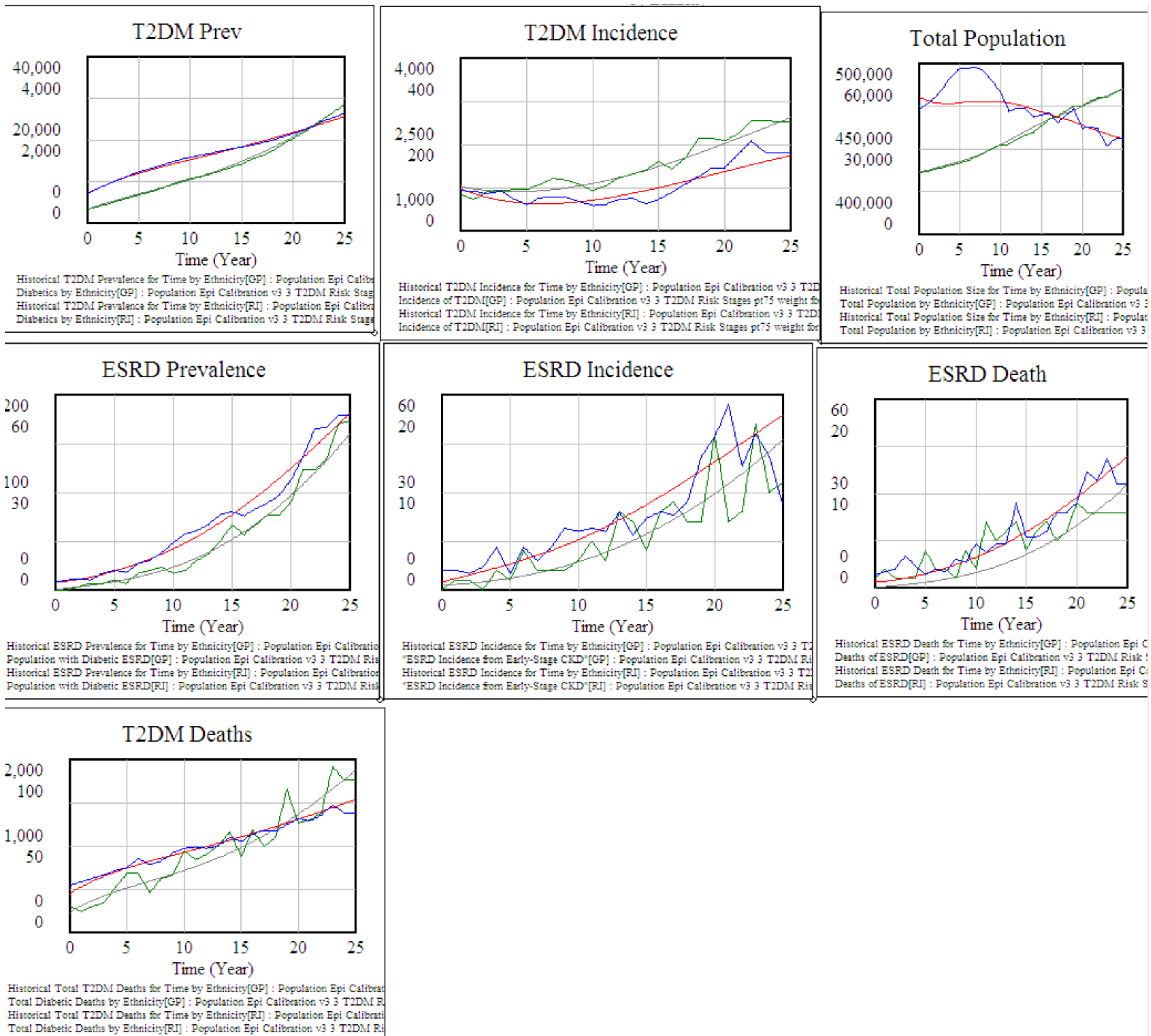
Calibration: "Triangulating" from Diverse Data Sources

- Calibration involves "tuning" values of less well known parameters to best match observed data
  - Often try to match against *many* time series or pieces of data at once
  - Idea is trying to get the software to answer the question: "What must these (less known) parameters be in order to explain all these different sources of data I see"
- Observed data can correspond to complex combination of model variables, and exhibit "emergence"
- Frequently we learn from this that our model structure just can't produce the patterns!

Calibration

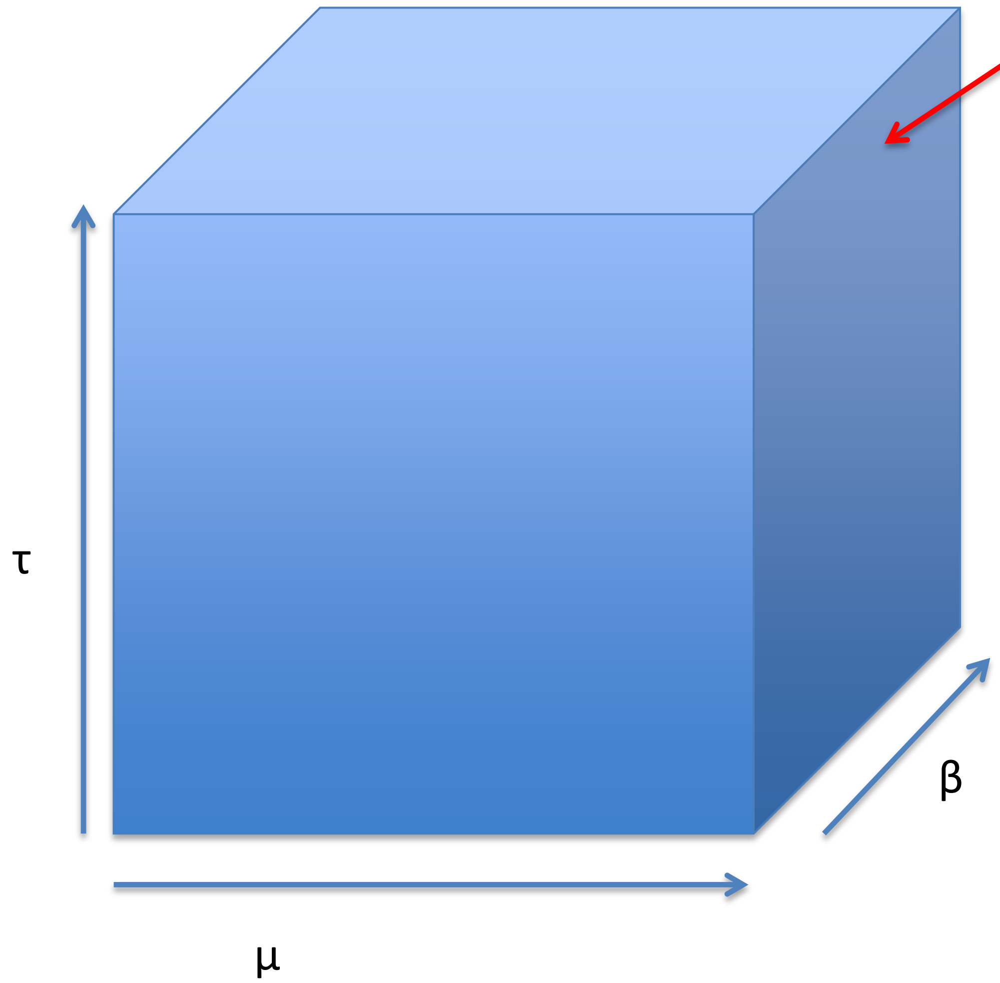
- Calibration helps us find a reasonable (specifics for) “dynamic hypothesis” that explains the observed data
  - Not necessarily the truth, but probably a reasonably good guess – at the least, a consistent guess
- Calibration helps us leverage the large amounts of diffuse information we may have at our disposal, but which cannot be used to directly parameterize the model
- Calibration helps us falsify models

Single Model Matches Many Data Sources



Envisioning “Parameter Space”

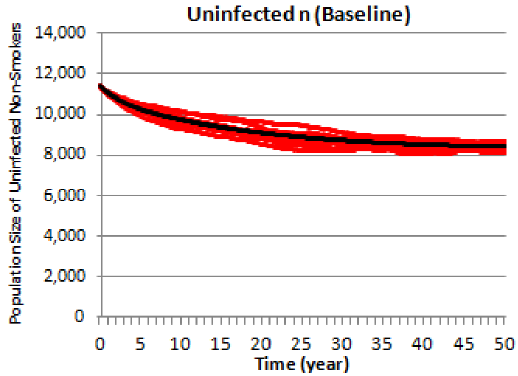
For each point in this space, there will be a certain “goodness of fit” of the model to the collective data

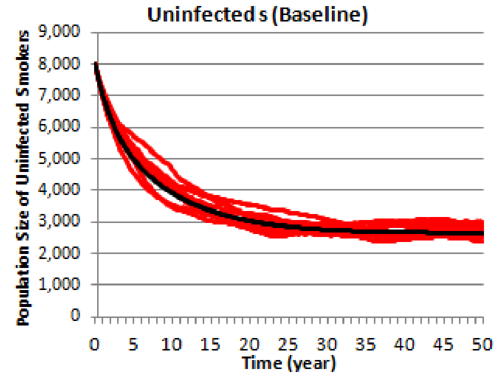


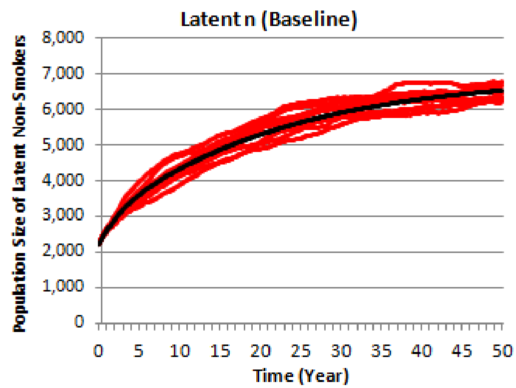
Stochastics in Agent-Based Models

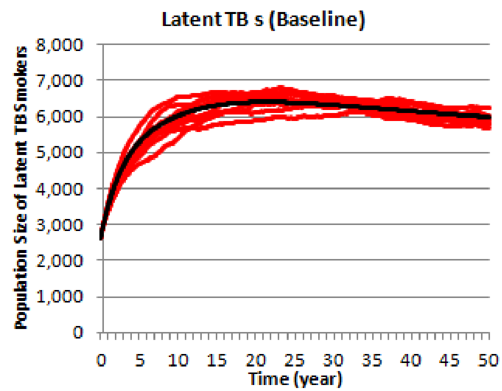
- Recall that ABMs typically exhibit significant stochastics
  - Event timing within & outside of agents
  - Inter-agent interactions
- Can have a pronounced impact on system evolution
- Such stochastics can account for observed patterns that are otherwise hard to explain
- When calibrating an ABM, we wish to avoid attributing a good match to a particular set of parameter values simply due to chance
- To reliably assess fit of a given set of parameters, we need to repeatedly run model realizations
  - We can take the mean fit of these realizations
- Often best to match not time series, but summaries

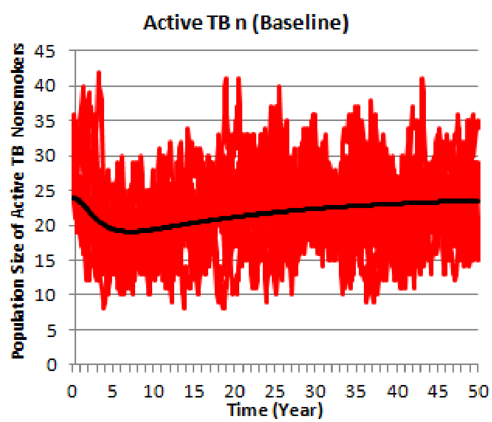
Examples of Stochastics (Compared to Mean Field Deterministic Model)

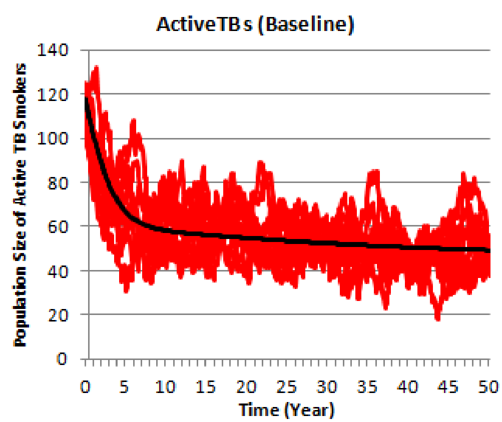








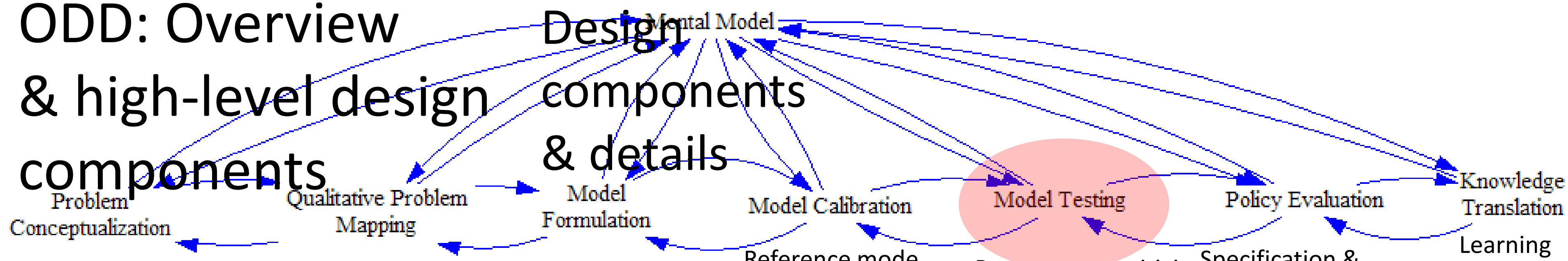




ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered
  - Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Specification & investigation of intervention scenarios

Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Learning environments (e.g. DISimS)

Visualizations

Units & Dimensions

- Distance
  - Dimension: Length
  - Units: Meters/Fathoms/Li/Parsecs
- Frequency (Growth Rate, etc.)
  - Dimension:1/Time
  - Units: 1/Year, 1/sec, etc.
- Fractions
  - Dimension: “Dimensionless” (“Unit”, 1)
  - Units: 1

Dimensional Analysis

- DA exploits structure of dimensional quantities to facilitate insight into the external world
- Uses
  - Cross-checking dimensional homogeneity of model
  - Deducing form of conjectured relationship
    (including showing independence of particular factors)
  - Sanity check on validation of closed-form model analysis
  - Checks on simulation results
  - Derivation of scaling laws
  * Construction of scale models
  - Reducing dimensionality of model calibration, parameter estimation

Uncertainty Analyses

- Same relative or absolute uncertainty in different parameters may have hugely different effect on outcomes or decisions
- Help identify parameters that strongly affect
  - Key model results
  - Choice between policies
- We place more emphasis in parameter estimation into parameters exhibiting high sensitivity

Uncertainty Analysis: Initial Value

- Frequently we don’t know the exact state of the system at a certain point in time
- A very useful type of sensitivity analysis is to vary the initial value of model stocks
- In Vensim, this can be accomplished by
  - Indicating a parameter name within the “initial value” area for a stock
  - Varying the parameter value

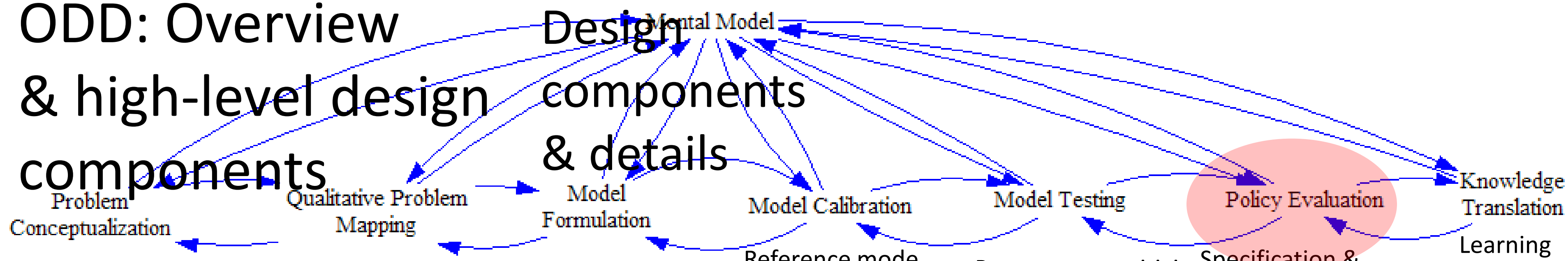
Robustness Analysis

- To what degree are model conclusions robust under changing model structural and other large assumptions?
  - Distinguish cases where
    - Results depends on something essential about the model
    - Results depend on happenstance of simplifying assumptions
      - e.g. spatial neighborhood assumption, size or granularity of space, convenient assumptions regarding rules or what is known
- We want to rule out cases where getting "right result for wrong reasons"!
- Seek to find whether conclusions change radically when just a few assumptions are changed?
- Process is similar to what used for submodel testing, but done for entire model
- Robustness under extreme values

ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered
  - Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Specification & investigation of intervention scenarios

Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Learning environments (e.g. DISimS)

Visualizations

Contact Tracing Simulation

We can make it better!

Run the model and switch to Main view

Network type

- (•) Random
- ( ) Small world
- ( ) Scale free

Network Settings

Connect Per Agent [ ]

Notes: Connects Per Agent is for Random and Small World Networks

Neighbourhood Link Prob [ ]

Notes: Link Prob is for Small World Networks

ScaleFreeM [ ]

Note: ScaleFreeM is for Scale Free Networks

Parameter Settings

Simulation Fraction of RI [ ]

Simulation Fraction of NonRI [ ]

[x] **Enable Database**

Contact Tracing Policy Selection

- (•) No Contact Tracing Program
- ( ) Contact Tracing With Priority

Contact Tracing Priority Settings (Weight)

[x] Age Priority [x] Ethnicity Priority [x] RR of Count Priority

Contact Tracing Targets

- (•) Tracing Infectious Active TB Cases ONLY
- ( ) Tracing All Active TB Cases
- ( ) Tracing Infectious Active TB Cases and Primary TB

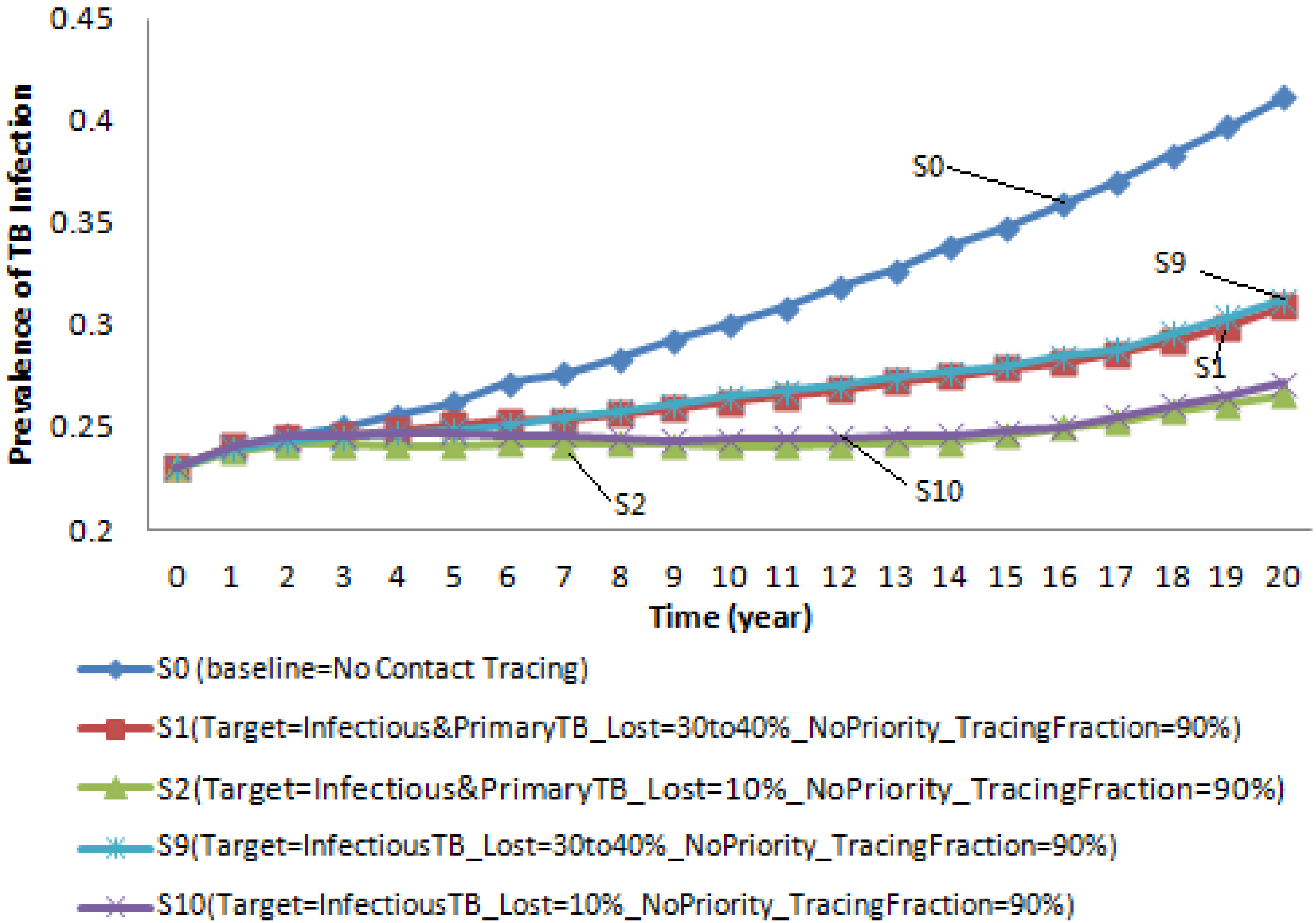
Contact Tracing Percentage on Average

Average Percentage of Contacts to Investigate: [ ]

Scenario Information

Description [ ]

Scenario Results (Means)



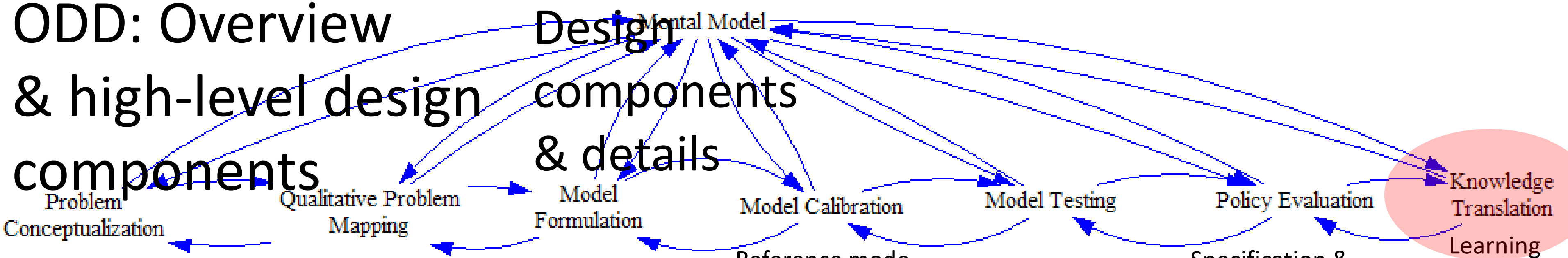
Variability in Results

Scenario Id	Cumulative Incident Cases (Active TB)				
	Mean	*Max*	*Min*	*Std. Deviation*	*C.V*
S_0	425.633	614	289	74.659	0.175
S_1	311.767	429	217	49.646	0.159
S_2	279.1	392	211	49.682	0.178
S_3	318.667	403	207	48.093	0.151
S_4	283	364	193	40.403	0.142
S_5	302.233	486	194	64.917	0.215
S_6	363.2	508	239	70.19	0.193
S_7	291	383	190	53.018	0.182
S_8	265.5	400	185	44	0.166
S_9	315	438	184	49.2	0.156
S_{10}	271.6	387	192	41.57	0.153

ABM Modeling Process Overview

ODD: Overview & high-level design components

ODD: Design components & details



Problem/research question articulation

Patterns for explanation

Model scope/boundary selection (endogeous, exogenous,ignored)

Key entities & their relationships
- Agents (&collectives)
- Environment
- Nesting hierarchy or network diagrams

Output of interest

State charts
Parameter & state variables
Qualitative State charts
Influence & Causal loop diagrams
Multi-agent interaction diagrams
Process flow structure
Key events
Rough model time & spatial extent
Observers to measure outputs

Specification of
- Parameters
- Quantitative causal relations
- Decision/behavior rules
  - Transitions
- Interactions
  - Messaging &handlers
  - Resources
- Relationship dynamics
- Mobility dynamics
- Initial conditions

Reference mode reproduction

Matching of intermediate time series

Matching of observed data points

Constrain to sensible bounds

Structural sensitivity analysis

Parameter sensitivity analysis

Cross-validation
- Formal
- w/Discovered
  - Patterns

Robustness tests

Extreme value tests

Unit checking

Problem domain tests

Specification & investigation of intervention scenarios

Investigation of hypothetical external conditions

Cross-scenario comparisons (e.g. CEA)

Learning environments (e.g. DISimS)

Visualizations

Contact Tracing Simulation

We can make it better!

Run the model and switch to Main view

Network type

- (•) Random
- ( ) Small world
- ( ) Scale free

Network Settings

Connect Per Agent [ ]

Notes: Connects Per Agent is for Random and Small World Networks

Neighbourhood Link Prob [ ]

Notes: Link Prob is for Small World Networks

ScaleFreeM [ ]

Note: ScaleFreeM is for Scale Free Networks

Parameter Settings

Simulation Fraction of RI [ ]

Simulation Fraction of NonRI [ ]

[x] **Enable Database**

Contact Tracing Policy Selection

- (•) No Contact Tracing Program
- ( ) Contact Tracing With Priority

Contact Tracing Priority Settings (Weight)

[x] Age Priority [x] Ethnicity Priority [x] RR of Count Priority

Contact Tracing Targets

- (•) Tracing Infectious Active TB Cases ONLY
- ( ) Tracing All Active TB Cases
- ( ) Tracing Infectious Active TB Cases and Primary TB

Contact Tracing Percentage on Average

Average Percentage of Contacts to Investigate: [ ]

Scenario Information

Description [ ]

Key Take-Home Messages from this Lecture

- Models express dynamic hypotheses about processes underlying observed behavior
- Frequently observed behavior is "emergent" – it is qualitatively different than the behavior of any one piece of the system, or a simple combination of behavior of those pieces
- Models help understanding how diverse pieces of system work together
- ABM focus on agent interactions as the fundamental shapers of dynamics
- Models are specific to purpose