

# Analytical Methods for System Dynamicists

هژیر رحمانداد

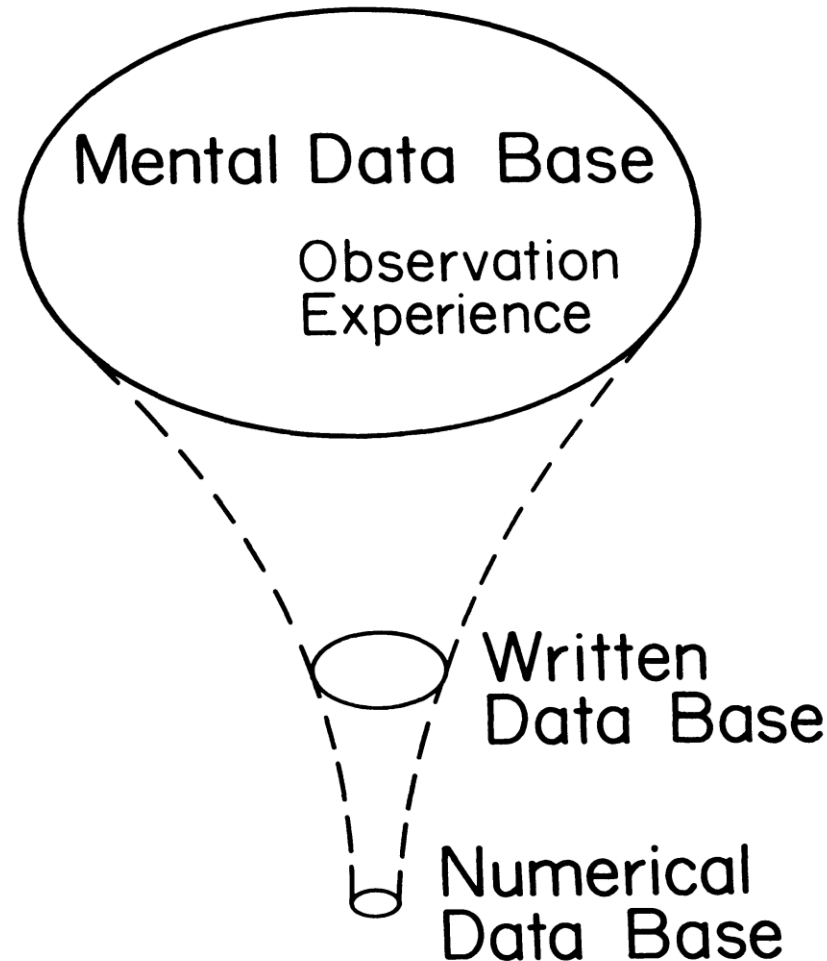
دانشیار پویاشناسی سامانه, ام.آی.تی

اولین کنفرانس ملی

انجمن ایرانی پویاشناسی سامانه ها

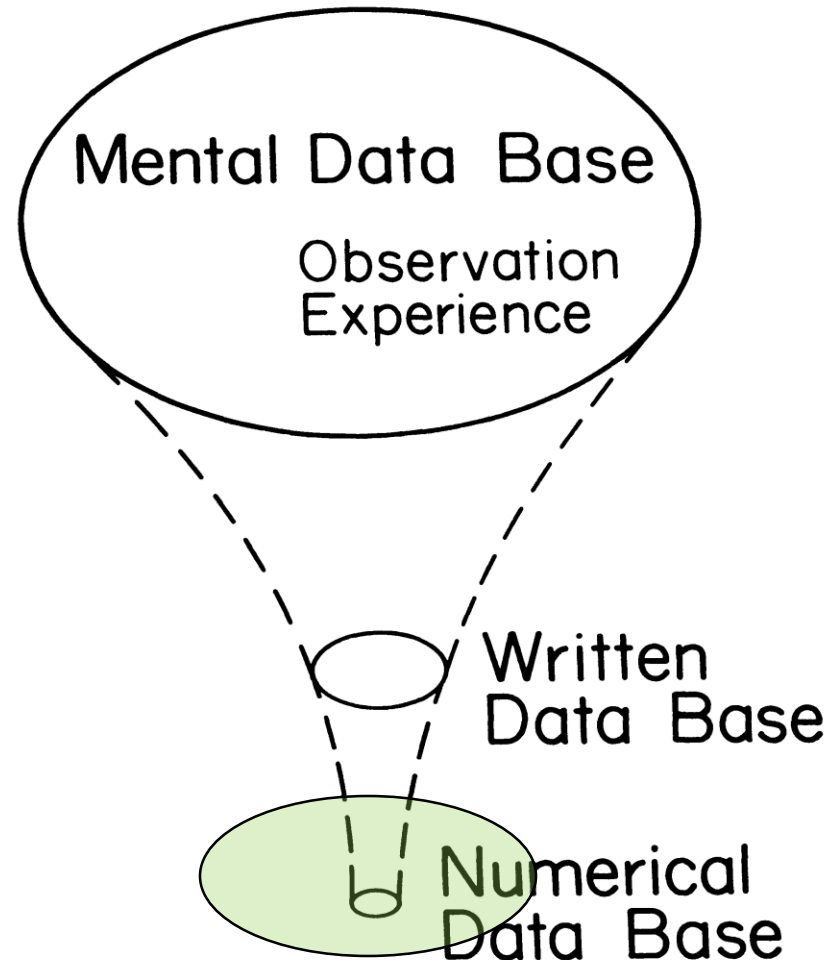
# Forrester's Representation of Three Data Bases

*A. Mental Data Base and Decreasing Content of Written and Numerical Data Bases*



# ... but numerical data is increasingly ubiquitous

*A. Mental Data Base and Decreasing Content of Written and Numerical Data Bases*

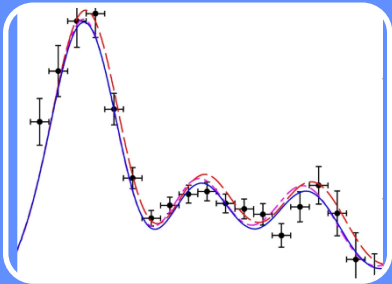


# The Modeling Process



## Model building

- Qualitative & archival data, expert opinion, prior theory
- Boundary & structure, dimensions, extreme conditions



## Calibration & Parameter Estimation

- Quantitative data from diverse sources
- Various methods depending on the data and model structure

Agency	Fluctuation	...	...	...	...	...	...
Marketing Affairs (Retailer)	PF012	-5.5	-	-	-15.5	15.5	-27.2
	PF011	-15.5	-	-	-15.5	15.5	-20.7
Japan Co. (China and Mexico National Government)	PF012	+2.2	-	-	+2.2	-1.3	+3.5
	PF011	0.5	0.5	-0.1	0.0	-	0.0
Administrative Agency (General Account for Meat Storage)	Fluctuation	+0.0	+0.0	+0.0	+0.0	-	+0.0
	PF012	0.0	0.0	-0.1	0.0	-	0.0
Account for Oil and Natural Gas	PF011	0.0	0.0	-0.1	0.0	-	0.0
	Fluctuation	+0.0	-	+0.0	+0.0	-	+0.0
Japan and National Service Agency Incorporated	PF012	0.0	0.0	-	-	-	0.0
	PF011	0.0	0.0	-	-	-	0.0
Administrative Agency (National Hospital)	Fluctuation	-0.0	-0.0	-	-	-	-
	PF012	56.4	17.7	-	15.8	15.3	1.8
Administrative Agency (National Hospital)	PF011	14.7	16.7	-	-2.0	2.2	-4.2
	Fluctuation	+15.7	+0.9	-	+14.8	+9.0	+5.8
Administrative Agency (National Hospital)	PF012	144.1	176.8	-	-22.8	77.6	-110.4
	PF011	103.2	106.0	-	-23.0	33.0	-116.0
Administrative Agency (National Hospital)	Fluctuation	-8.2	-8.2	-	-8.0	-4.2	-4.2
	PF012	11.0	10.7	-	0.2	27.0	-26.8
Administrative Agency (National Hospital)	PF011	-1.0	3.0	-	-20.0	20.2	-49.1
	Fluctuation	+11.9	+0.9	-	+11.1	-1.3	+12.3
Administrative Agency (National Hospital)	PF012	42.9	14.8	-	29.3	28.4	-0.1
	PF011	47.1	17.2	-	29.9	21.0	-1.1
Administrative Agency (National Hospital)	Fluctuation	-4.2	-2.7	-	-1.5	-2.5	+1.0

## Model and Policy Analysis

- Structure and behavior analysis
- Sensitivity & robustness analysis
- Various optimization methods

# Use of Data in Simulation Models

- Use multiple types of data for multiple purposes and different stages of modeling
  - Qualitative data is more appropriate for identification of the structure of problem and important dynamic hypotheses
  - Numerical and archival data is needed for parameter estimation and validation
  - Time series can feed into model the factors outside of model boundary

# From Qualitative to Quantitative Data

- Following the SD modeling process, you can get high structural reliability through extensive use of qualitative data
- The qualitative insights allow you to tell convincing stories
  - This is good: lots of insights; convincing.
  - This is bad: you can fool the client; and yourself.
- Hard data can keep you straight and allow for testing of hypotheses
- It also makes your work much stronger, both in persuasion and impact

# Two potential misconceptions

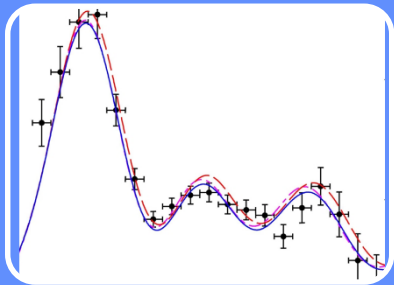
- Misconception 1: We can't build models without numerical data
  - If model structure is to follow real-world structures in charge of problem, qualitative data is needed for building it
  - Insights can be generated before starting to use numerical data
- Misconception 2: Numerical data adds little value to the modeling
  - Uncertainties in parameter values can only be set with empirical data
  - Many theoretical and practical implications depend on these parameters

# The Modeling Process



## Model building

- Qualitative & archival data, expert opinion, prior theory
- Boundary & structure, dimensions, extreme conditions



## Calibration & Parameter Estimation

- Quantitative data from diverse sources
- Various methods depending on the data and model structure

Agency	Fluctuation	...	...	...	...	...	...			
Education Affairs Ministry	PI002	-5.5	-	-	-15.5	15.5	-27.2	-12.8	0.8	
	PI001	-5.5	-	-	-15.5	15.5	-20.7	-10.7	-0.0	
Japan Co. Coe and Nippon National Government	Fluctuation	+2.2	-	-	+2.2	-1.3	+3.8	-	-	
	PI002	0.5	0.5	-0.1	0.0	-	0.0	-	0.5	0.5
Administration Agency	Fluctuation	+0.0	+0.0	+0.0	+0.0	-	+0.0	-	-	
	PI002	0.5	0.5	-0.1	0.0	-	0.0	-	0.5	0.5
Central Agency for Moral Education	PI001	0.5	0.5	-0.1	0.0	-	0.0	-	0.5	0.5
	Fluctuation	+0.0	-	-	+0.0	+0.0	-	+0.0	-	-
Ministry for Child and Family Affairs	PI002	0.5	0.5	-	-	-	-	-	0.5	0.5
	PI001	0.0	0.0	-	-	-	-	-	0.0	0.0
National Institute for Health	Fluctuation	-0.0	-0.0	-	-	-	-	-	-	-
	PI002	50.4	71.7	-	15.8	11.3	1.8	-	48.7	48.8
National Institute for Health Research	Fluctuation	34.7	16.2	-	-2.0	2.2	-4.2	-	17.7	18.1
	PI001	+18.7	+0.9	-	+14.8	+0.8	+5.8	-	-	-
National Institute for Health Research	Fluctuation	144.1	176.8	-	-22.8	77.6	-110.4	-	148.5	219.9
	PI001	103.3	106.0	-	-23.0	33.0	-91.6	-	134.7	204.8
National Institute for Health Research	Fluctuation	-8.2	-8.2	-	-8.8	-4.2	-4.2	-	-	-
	PI002	11.0	10.7	-	-0.2	17.0	-28.9	-	11.0	12.2
National Institute for Health Research	Fluctuation	+11.0	+0.9	-	+0.9	29.2	-49.1	-	-8.9	17.3
	PI001	-1.0	1.0	-	-11.1	-1.3	+12.3	-	-	-
National Institute for Health Research	Fluctuation	42.9	14.6	-	39.3	38.4	-0.1	-	-43.2	88.6
	PI001	47.1	17.0	-	39.9	31.9	-1.1	-	-47.7	88.8
National Institute for Health Research	Fluctuation	-4.2	-2.7	-	-1.5	-2.5	+1.0	-	-	-
	PI002	-	-	-	-	-	-	-	-	-

## Model and Policy Analysis

- Structure and behavior analysis
- Sensitivity & robustness analysis
- Various optimization methods



# Calibration

- Calibration is the fitting of a model to data
  - We do this using an optimization
- The purpose of calibration
  - Rejection
    - Reject a model when it fails to fit data
    - We need to know the criteria for rejection
  - Parameter and structure estimation
    - Confidence intervals

# Calibration Process

- Assume the model is right
- Set up the model and run calibration
  - Specify the data items to match, the corresponding model variables, and the payoff (measure of closeness of data items to model outcomes)
  - Specify the parameter space for search
  - Specify the optimization algorithm
  - Run the calibration/optimization
- After finding the results
  - Measure the closeness, should model be rejected?
  - Find the confidence intervals for the estimated parameters

# Menu of Estimation Choices

- **Estimation scope:**
  - **Partial model/regression:** Estimate relationships separately
  - **System level:** Estimate the full model (or larger chunks with multiple items to match)
- **Calibration payoff function:**
  - Ad hoc, e.g. weighted least squares
  - Maximum likelihood
  - MSM/Indirect inference
- **Use of state resetting and filtering:**
  - Simple simulation
  - Simple regressions
  - Filtering (e.g. Extended Kalman Filter (EKF) or Particle filter)
- **Confidence intervals:**
  - Likelihood based
  - MSM based
  - Bootstrapping and subsampling

# Model Calibration Tips

- Include in calibration problem ALL knowledge available about system parameters
  - known parameters
  - physical constraints on parameters
  - likely uncertainty range
- Use the smallest calibration problems possible
  - Breakdown the model into separate pieces and estimate the parameters for each piece
  - Estimate the next pieces based on external data and input from already estimated pieces
    - More manual, but more robust results due to immediacy of parameters to independent and dependent variables
- Test the hypotheses “The estimated parameter matches the observable structure of the system”
  - Does the model match the historical behavior
  - Does the model match the structure

# Tips for Calibration and Optimization

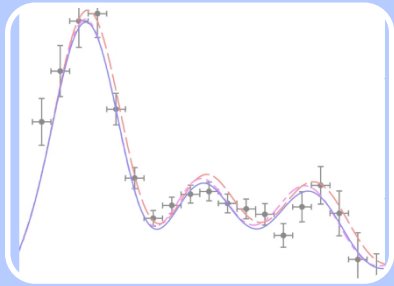
- Creating the payoff function
  - Audience matters: engineering and applied communities prefer simple metrics such as MSE, MAPE. Statistically oriented communities want confidence intervals and that requires MLE or Method of Simulated Moments.
- Optimization: No single algorithm dominates; order of magnitude differences in efficiency; thus explore:
  - Quadratic methods feasible or payoff landscape noisy?
  - Starting from different points, how varied are the local peaks found by gradient-based methods?
  - How effective are gradient-free methods (e.g. Powell set method)?
- Stop calibration only when the majority of new searches converge to a peak already identified.

# The Modeling Process



## Model building

- Qualitative & archival data, expert opinion, prior theory
- Boundary & structure, dimensions, extreme conditions



## Calibration & Parameter Estimation

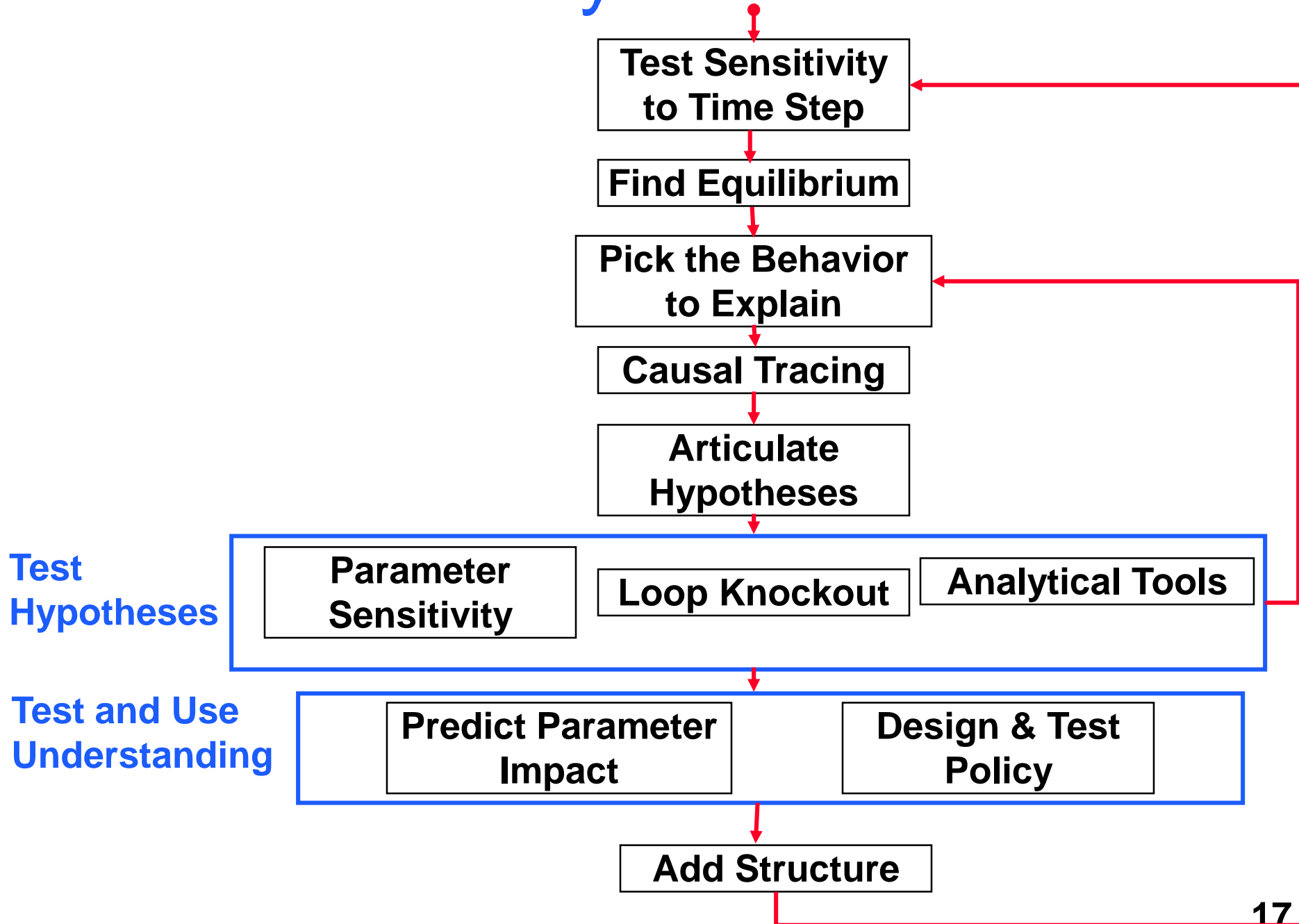
- Quantitative data from diverse sources
- Various methods depending on the data and model structure

Agency	Fluctuation	...	...	...	...	...	...		
Marketing Affairs (Retailer)	PF012	-15.5	-	-	-15.5	15.5	-27.2	-12.8	0.8
	PF011	-15.6	-	-	-15.6	15.0	-20.7	-10.7	-0.0
Japan Co. (Japan and Mexico National Government)	PF012	+2.2	-	-	+2.2	-1.3	+3.5	-	-
	PF011	0.5	0.5	-0.1	0.0	-	0.0	-	0.5
Administrative Agency (Japan)	Fluctuation	+0.0	+0.0	+0.0	+0.0	-	+0.0	-	-
	PF012	0.0	0.0	-0.1	0.0	-	0.0	-	0.5
General Account for Meat Storage	PF011	0.0	0.0	-0.1	0.0	-	0.0	-	0.5
	Fluctuation	+0.0	-	+0.0	+0.0	-	+0.0	-	-
Account for Oil and Natural Gas	PF012	0.0	0.0	-	-	-	-	-	0.0
	PF011	0.0	0.0	-	-	-	-	-	0.0
Fluctuation	Fluctuation	-0.0	-0.0	-	-	-	-	-	-
	PF012	56.4	11.7	-	15.8	15.3	1.8	-	62.7
Japan Post National Service Agency Incorporated	PF011	14.7	16.7	-	-2.0	2.2	-4.2	-	17.7
	Fluctuation	+15.7	+0.9	-	+14.8	+9.0	+5.8	-	-
National Hospital Organization	PF012	144.1	176.8	-	-22.8	77.6	-110.4	-	148.1
	PF011	103.2	106.0	-	-22.0	33.0	-116.0	-	134.7
National Hospital Organization	Fluctuation	-8.2	-8.2	-	-8.0	-4.2	-4.2	-	-
	PF012	11.0	10.7	-	0.2	27.0	-26.8	-	11.0
National Cancer Center	PF011	-1.0	3.0	-	-0.0	29.2	-49.1	-	-6.9
	Fluctuation	+11.9	+0.9	-	+11.1	-1.3	+12.3	-	-
National Center for Global Health and Research	PF012	42.9	14.8	-	29.3	28.4	-0.1	-	43.2
	PF011	47.1	17.2	-	29.9	21.0	-1.1	-	47.7
Fluctuation	Fluctuation	-4.2	-2.7	-	-1.5	-2.5	+1.0	-	-

## Model and Policy Analysis

- Structure and behavior analysis
- Sensitivity & robustness analysis
- Various optimization methods

# Analysis Process

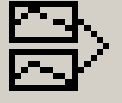


# Pick the Behavior to Explain

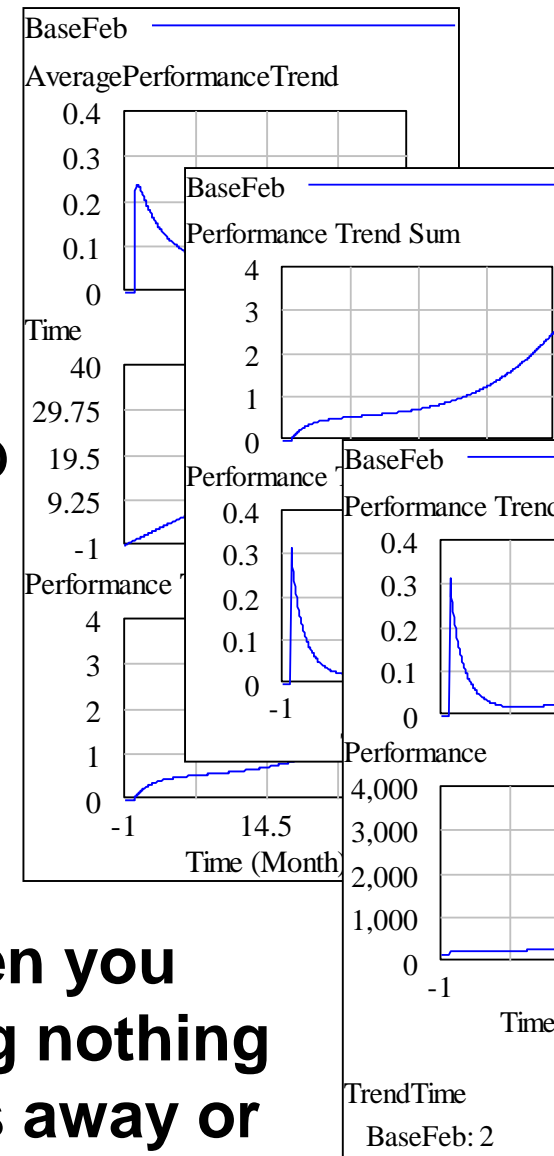
- Different variables may have different modes of behavior, and the sources of dynamics are not always the same
- Pick one variable's behavior and understand that
- Usually by doing this for 2-3 variables, you will know the model's behavior completely



# Causal Tracing and Hypotheses



- Look at the causes for each variable and trace back until:
  - Single source is found (e.g. a draining, isolated stock)
  - A loop hypothesis is found (you loop back to the original variable)
  - You form other hypothesis why the behavior happens
- Design an experiment to test your hypothesis: **if hypothesis is  $X \rightarrow Y$ , then you remove  $X$  (while changing nothing else) to see if  $Y$  also goes away or not (e.g. loop knock-out analysis)**



# Structural Dominance Analysis

- Objectives of SDA
  - Articulate structural explanations for behavior
  - Support policy design
- Three flavors of SDA
  - Exploratory analysis of dominant structure
    - Simulation-based and/or manual approaches; e.g. loop-knockout, sensitivity analysis and statistical screening
  - Formal assessment of dominant structure
    - Eigenvalue/eigenvector analysis of linearized model
    - Pathway Participation Method

# Test and Use Your Understanding

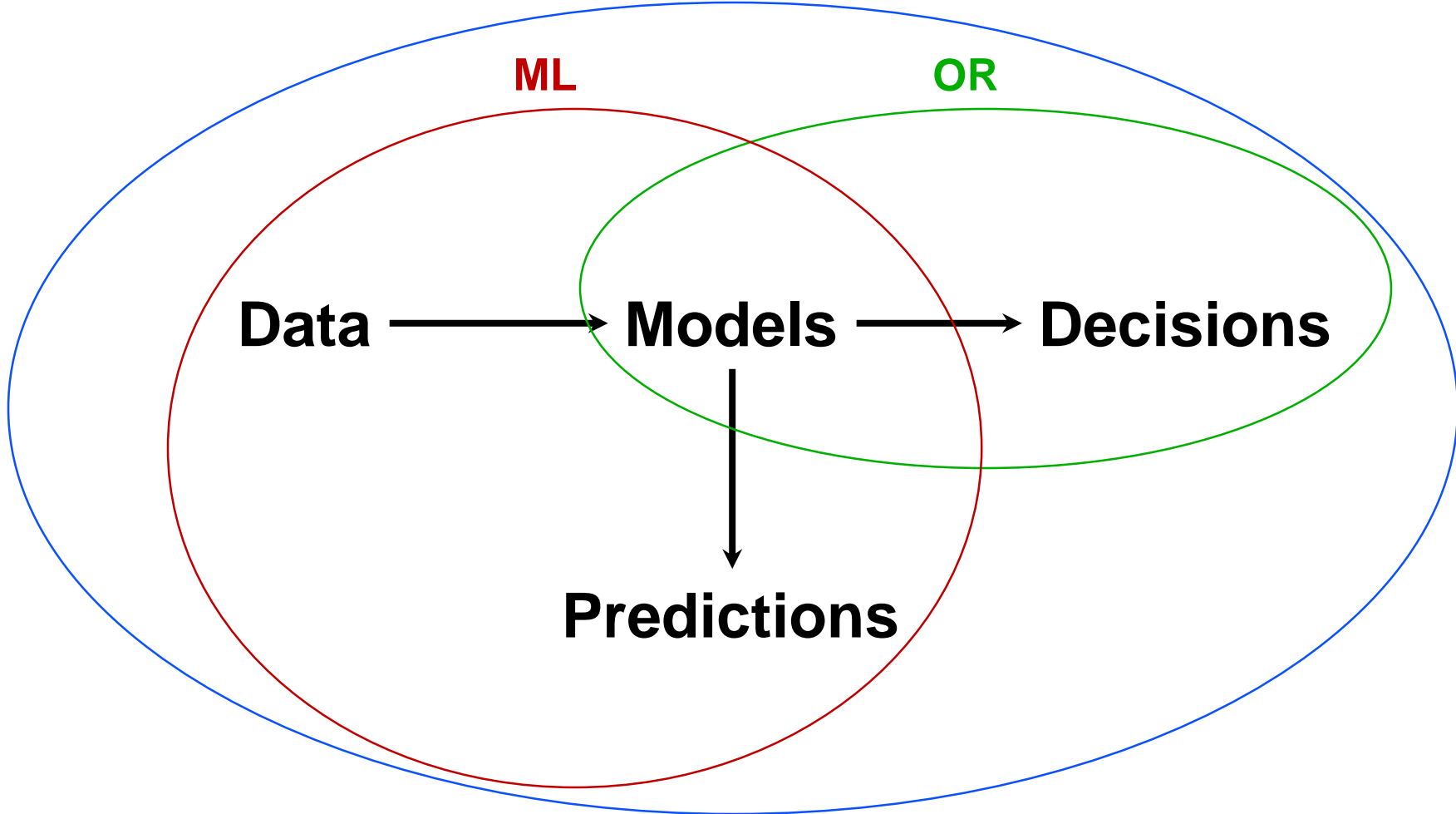
- Predict parameter impact
  - List the model parameters in a table
  - Predict what will happen if you change each (e.g. the mode increases/decreases)
  - Test and see if your prediction was correct
  - When wrong, find out what you missed
- Help build client's mental models
  - Design simple simulations to communicate the basic insights to your client
  - Be ready to explain in simple terms why the model behaves the way it does
  - Go through likely scenarios and their mechanisms
  - Engage client in explaining dynamics

# Controlling Dynamic Systems

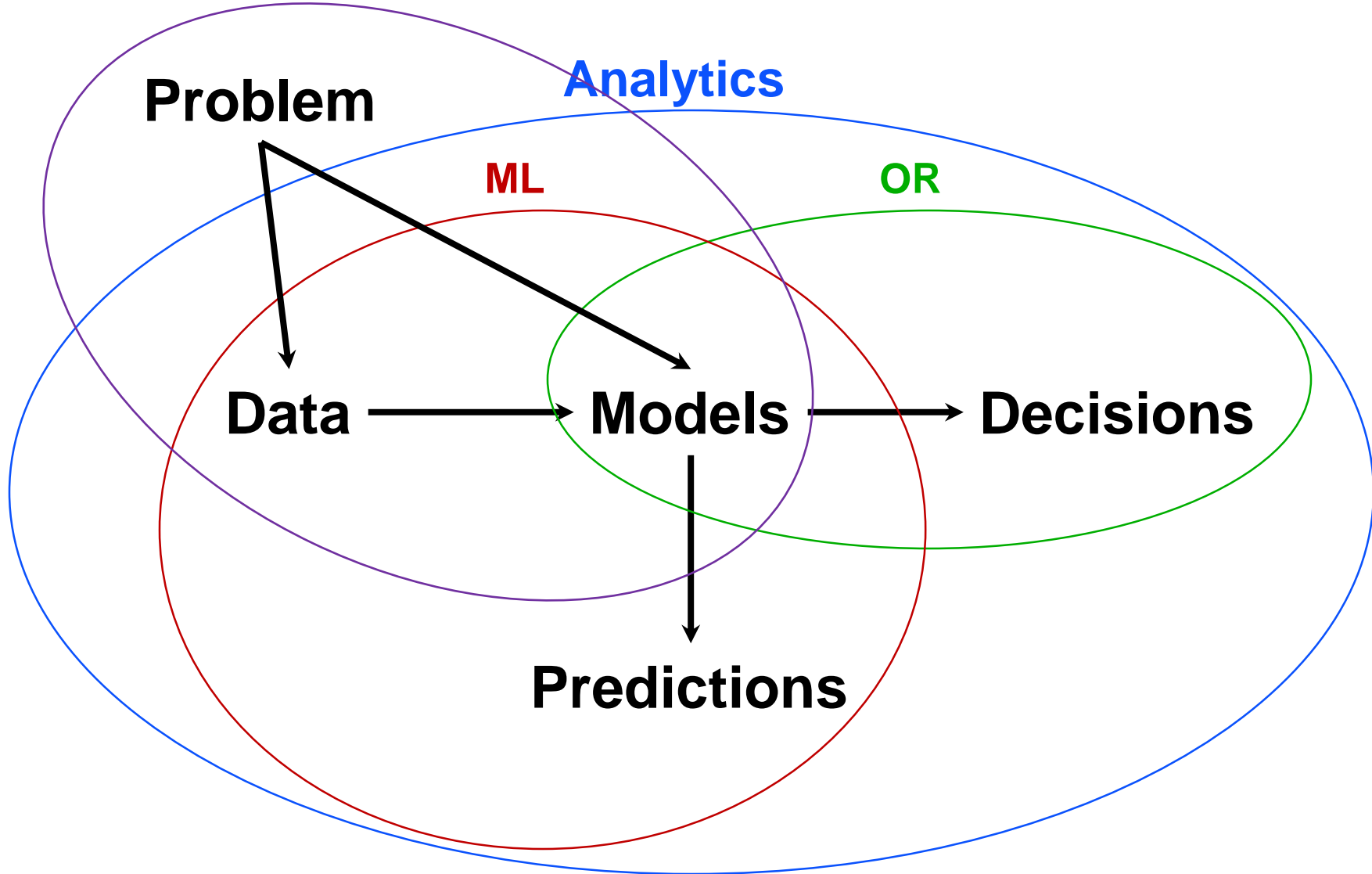
- **Overview:** you have (a model of) a dynamic system, how can you use this model to best manage (control) the system?
  - You are the main actor → e.g. Managing your inventory, vehicle control, epidemic planning
    - optimal control, (approximate) dynamic programming and policy optimization
    - Decision analysis, decision trees, stochastic optimization
  - There are other (rational) actors with different or opposing goals → e.g. Pricing in competition, missile defense, Market entry decisions
    - Dynamical games

# Analytics

## Analytics



# SD and Analytics



# SD models, Big Data, and Analytics

- The model is build from data beyond the numerical base
- Emphasis on operational explanations means the focus of SD is in describing the system that generates the data
  - Question the source of the data
  - Question the quality of the data
    - Violations of conservation of matter
    - Biases on data collection of classification
- Use OR, ML, Econometric methods for estimation, optimization, and decision analysis