Optimization and Simulation

Introduction

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

Outline

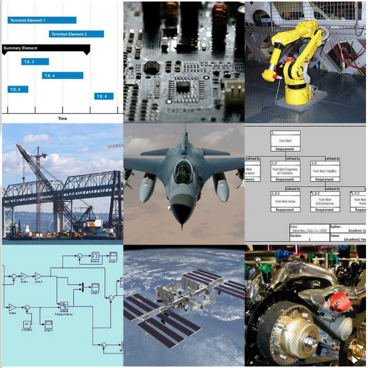
Engineering systems

Definition (Wikipedia)

Combination of components that work in synergy to collectively perform a useful function.

Properties

- Complex
- Large
- Designed
- Configurable
- Interactions with external world



Source: Wikipedia

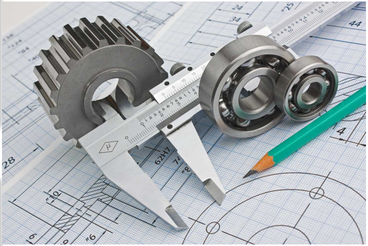
Engineering systems

Objectives

- Design
- Maintain
- Operate

Time horizon

- Long-term
- Medium-term
- Short-term

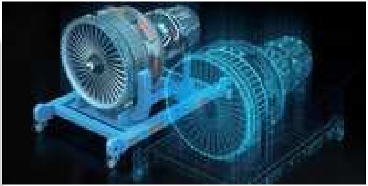


Source: Swiss Learning Exchange

Engineering systems

Mathematical and digital twins

- Modeling
- Simulation
- Optimization



Source: Konica Minolta

Engineering systems

	Modeling	Simulation	Optimization
Roles	Represent	Predict	Improve
How?	Capture causal effects	Capture the propagation of uncertainty	Investigate better configurations

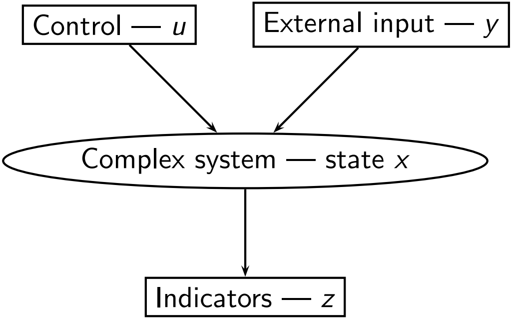
Outline

Modeling

System

A system can be seen as a black box, modeled by

$$z = h(x, y, u; \theta)$$



Modeling

$$z = h(x, y, u; \theta)$$

Example

A car:

- x captures the state of the system (e.g. speed, position of other vehicles)
- y captures external influences (e.g. wind)
- u captures possible human controls on the system (e.g. acceleration/deceleration)
- z represents indicators of performance (e.g. oil consumption).

Modeling

Decompose the complexity

- The model h is usually decomposed to reflect the interactions of the subsystems
- For example,
  - a car-following model captures the target speed of the driver,
  - an engine model derives the actual consumption as a function of the acceleration.

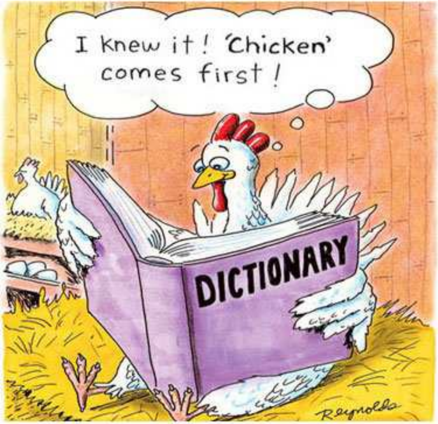
Modeling

Causal effects

- Very important to identify the causal effects
- Failure to do so may generate wrong conclusions

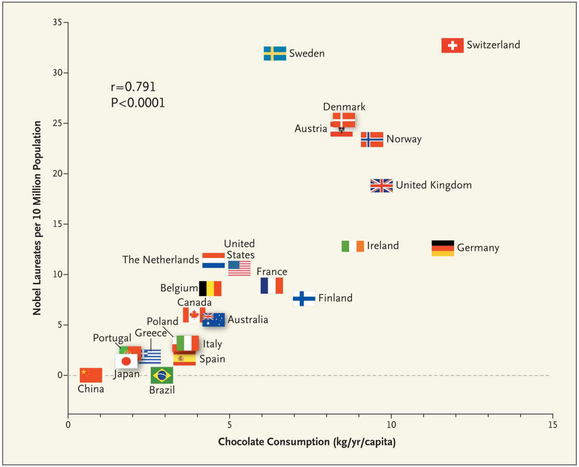
Forecasting

Assumption: causal effects are stable over time and configurations of the system.



Data can be misleading

Chocolate Consumption, Cognitive Function, and Nobel Laureates



Source: [Messerli, 2012]

Inference

Data collection

On an existing system, collect N observations of x_n, y_n, u_n, z_n, $n = 1, \ldots, N$.

Goodness of fit

For a given value of θ, "distance" $d_n(\theta)$ between

- the predicted value $h(x_n, y_n, u_n; \theta)$, and
- the observed value z_n.

Inference

Find $\widehat{\theta}$ that minimizes the total distance:

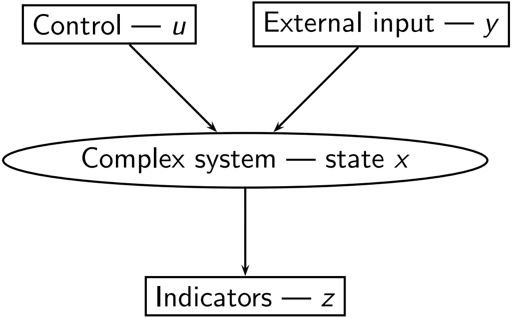
$$\widehat{\theta} = \text{argmin}_\theta \sum_{n=1}^{N} d_n(\theta).$$

Outline

Simulation

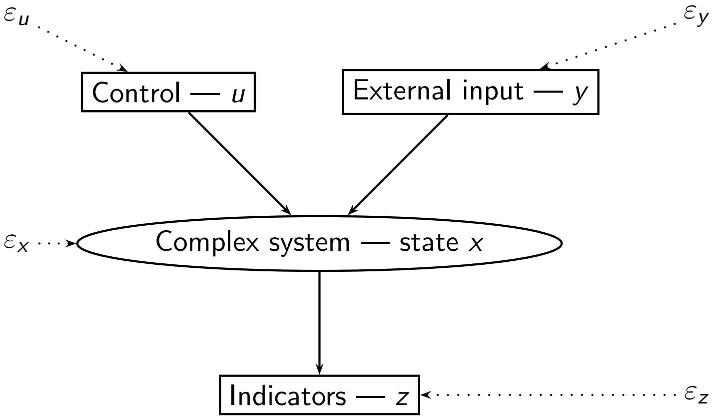
Simulation is more than simply applying the model.

$$z = h(x, y, u; \widehat{\theta})$$



Simulation

$$Z = h(X, Y, U; \widehat{\theta}) + \varepsilon_z$$



Simulation

Propagation of uncertainty

$$Z = h(X, Y, U; \widehat{\theta}) + \varepsilon_z$$

- Given the distribution of X, Y, U and ε_z
- what is the distribution of Z?

Simulation

Sampling

- Draw realizations of X, Y, U, ε_z
- Call them $x^r, y^r, u^r, \varepsilon_z^r$
- For each r, compute

$$z^r = h(x^r, y^r, u^r; \widehat{\theta}) + \varepsilon_z^r$$

- z^r are draws from the random variable Z

Analysis

- Generate many draws from Z.
- Analyze their empirical distribution.

Importance of number of draws

Theory vs. practice

- Theory: true distribution of Z when $r \to \infty$.
- Practice: finite number R of draws.
- If R is too small, simulator output is just noise.

Analogy with real world

- Nature also generates instances of a complex random variable.
- Experiments must be repeated in order to reach conclusions.

Example: policy analysis

- The real impact of a policy is difficult to analyze.
- Incomplete results that are consistent with expectations may lead to erroneous conclusions.

Example: improving safety

Accidents in Kid City

- The mayor of Kid City has commissioned a consulting company
- Objective: assess the effectiveness of safety campaigns
- They propose to use simulation

Example: improving safety

Accidents in Kid City

Example: improving safety

Accidents in Kid City: 

Example: improving safety

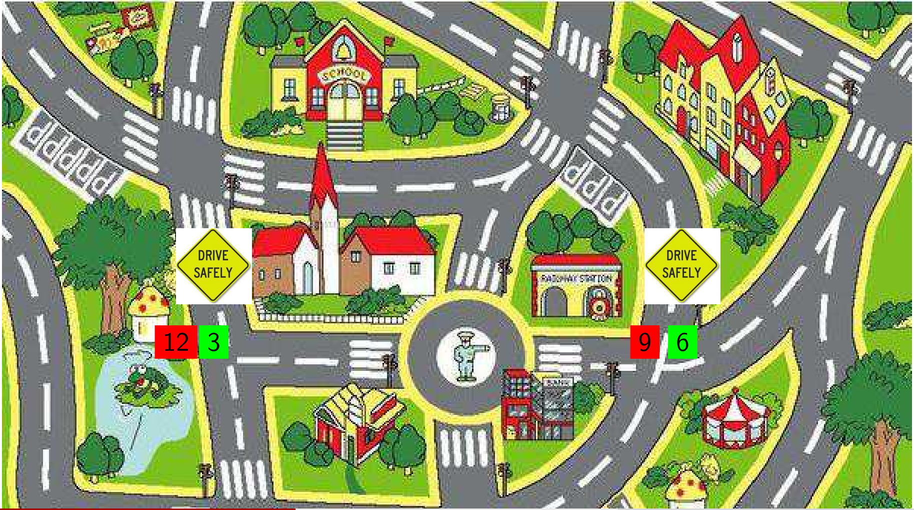
Accidents in Kid City

Example: improving safety

Accidents in Kid City

Example: improving safety

Accidents in Kid City:



Example: improving safety

Two major flaws

- Causal effects are not modeled
- Simulation performed with only one draw

What should have been done

- Simulate the number of accidents many times.
- If so, the average number of accidents is around 7, everywhere, with or without the sticker.
- A formal statistical test would not reject the null hypothesis that the sticker has no effect.

Outline

Simulation

Derivation of indicators from the distribution

- Mean
- Variance
- Modes
- Quantiles

Statistics

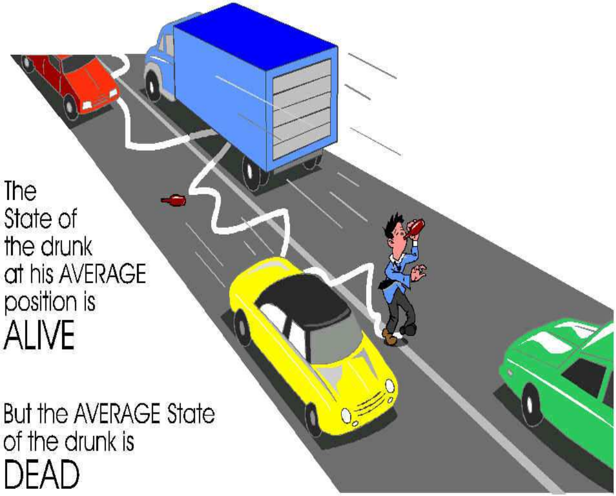
"Numbers don't lie. That's where we come in."

Indicators

- Mean: $\mathrm{E}[Z] \approx \bar{Z}_R = \frac{1}{R}\sum_{r=1}^{R} z^r$
- Variance: $\mathrm{Var}(Z) \approx \frac{1}{R}\sum_{r=1}^{R}\left(z^r - \bar{Z}_R\right)^2$.
- Modes: based on the histogram
- Quantiles: sort and select

Important: there is more than the mean

The mean



[Savage et al., 2012]

The mean

The flaw of averages

[Savage et al., 2012]

$$\mathrm{E}[Z] = \mathrm{E}[h(X, Y, U; \widehat{\theta}) + \varepsilon_z] \neq h(\mathrm{E}[X], \mathrm{E}[Y], \mathrm{E}[U]; \widehat{\theta}) + \mathrm{E}[\varepsilon_z]$$

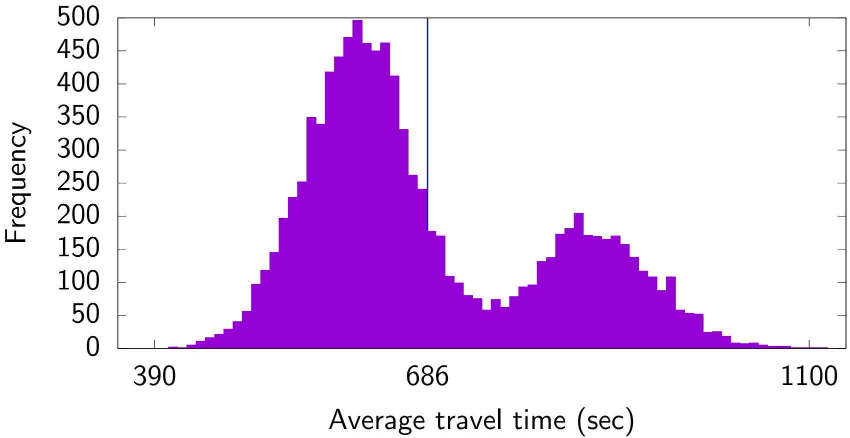
... except if h is linear.

There is more than the mean

Example

- Intersection with capacity 2000 veh/hour
- Traffic light: 30 sec green / 30 sec red
- Constant arrival rate: 2000 veh/hour during 30 minutes
- With 30% probability, capacity at 80%.
- Indicator: Average time spent by travelers

There is more than the mean



Pitfalls of simulation

Few number of runs

- Run time is prohibitive
- Tempting to generate partial results rather than no result

Focus solely on the mean

- The mean is useful, but not sufficient.
- For complex distributions, it may be misleading.
- Intuition from normal distribution (mode = mean, symmetry) do not hold in general.
- Important to investigate the whole distribution.
- Simulation allows to do it easily.

Challenges

- How to generate draws from Z?
- How to represent complex systems? (specification of h)
- How large R should be?
- How good is the approximation?

Pseudo-random numbers

Definition

- Deterministic sequence of numbers
- which have the appearance of draws from a $U(0,1)$ distribution

Typical sequence

$$x_n = a x_{n-1} \text{ modulo } m$$

- This has a period of the order of m
- So, m should be a large prime number
- For instance: $m = 2^{31} - 1$ and $a = 7^5$
- x_n/m lies in the $[0, 1[$ interval

Outline of the lectures

- Drawing from distributions
- Discrete event simulation
- Data analysis
- Variance reduction
- Markov Chain Monte Carlo

Reference

[Ross, 2012]

Outline

Optimization

Assumptions

- Control U is deterministic.

$$Z(u) = h(X, Y, u) + \varepsilon_z$$

- Various features of Z are considered: mean, variance, quantile, etc.

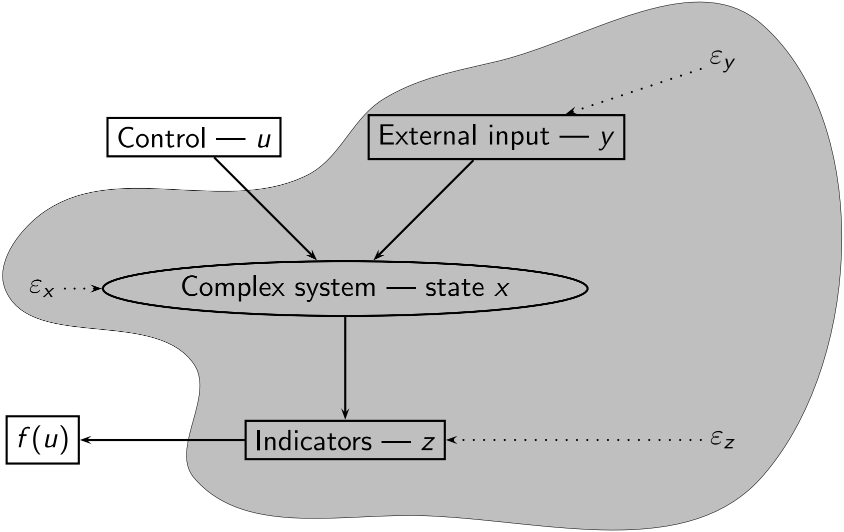
$$(z_1(u), \ldots, z_m(u))$$

- They are combined in a single indicator:

$$f(u) = g(z_1(u), \ldots, z_m(u))$$

- If not, it is called *multi-objective* optimization.

General framework: the black box



Optimization problem

$$\min_{u \in \mathbb{R}^n} f(u)$$

subject to

$$u \in \mathcal{U} \subseteq \mathbb{R}^n$$

- u: decision variables
- $f(u)$: objective function
- $u \in \mathcal{U}$: constraints
- $\mathcal{U}$: feasible set

In this course. . .

- Classical optimization problems
- Heuristics
- Multi-objective optimization

Summary

Modeling

- Decomposition of the complexity.
- Causal effects.

Simulation

- Propagation of uncertainty.
- Requires many draws.
- Analysis of the entire empirical distribution.
- There is more than the mean.

Optimization

- Identify the control that improves a function of the indicators.
- Optional: multi-objective optimization.

Bibliography

Messerli, F. (2012).
Chocolate consumption, cognitive function, and nobel laureates.
The New England Journal of Medicine, 367:1562–1564.

Ross, S. M. (2012).
Simulation.
Elsevier, fifth edition.

Savage, S., Danziger, J., and Markowitz, H. (2012).
The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty.
Wiley.