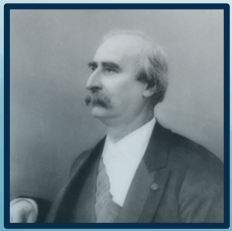
Video lecture: https://vimeo.com/album/3275353/video/120523455

How to combine observational data sources with first principles of physics to build stable and transportable models for power system design and control?

Louis Wehenkel
Systems and Modeling
Dept. of EE&CS – University of Liège, Belgium

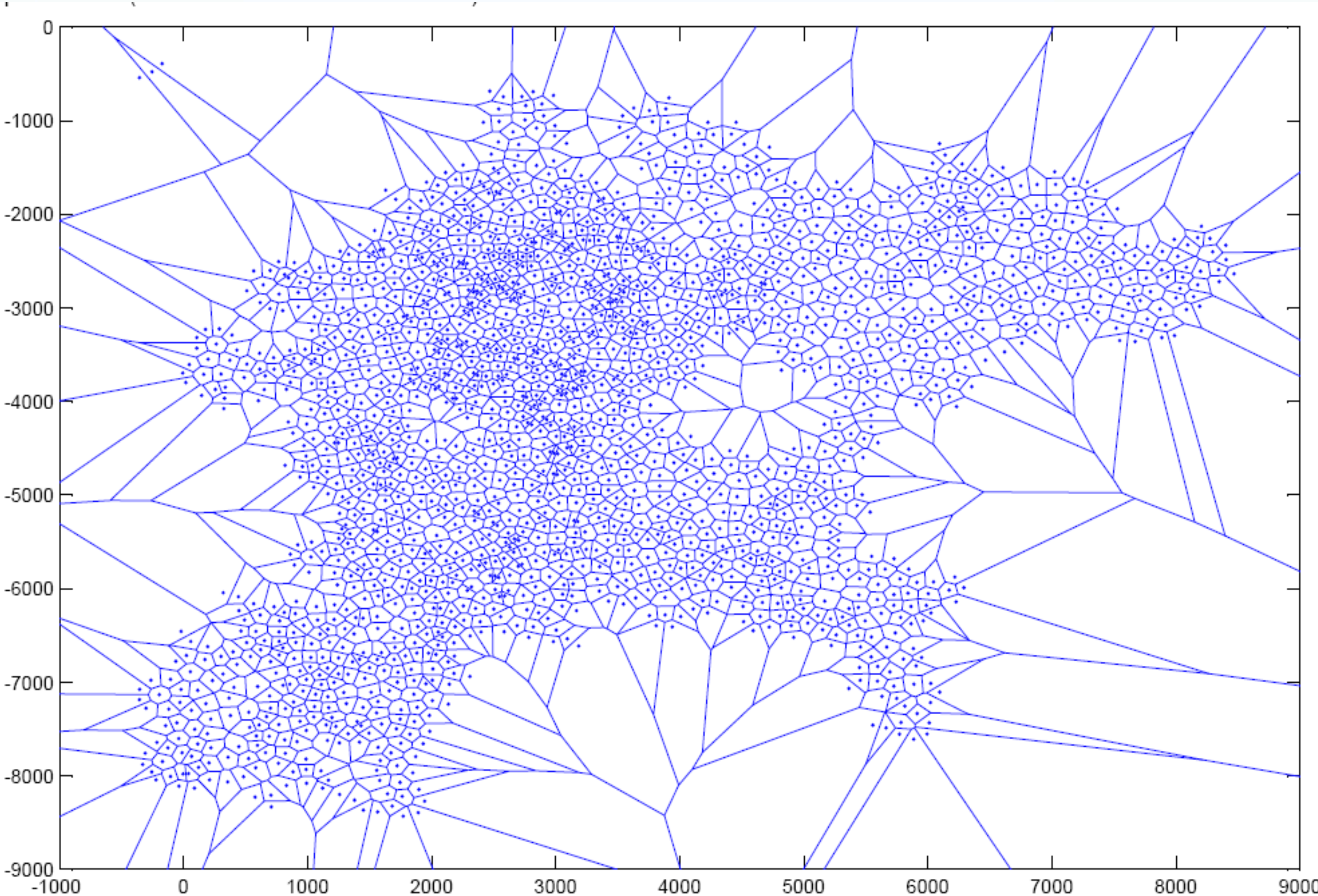
Analytic Research Foundations for the Next-Generation Grid. A workshop of the National Research Council of the National Academies, Irvine Feb. 11-12, 2015.

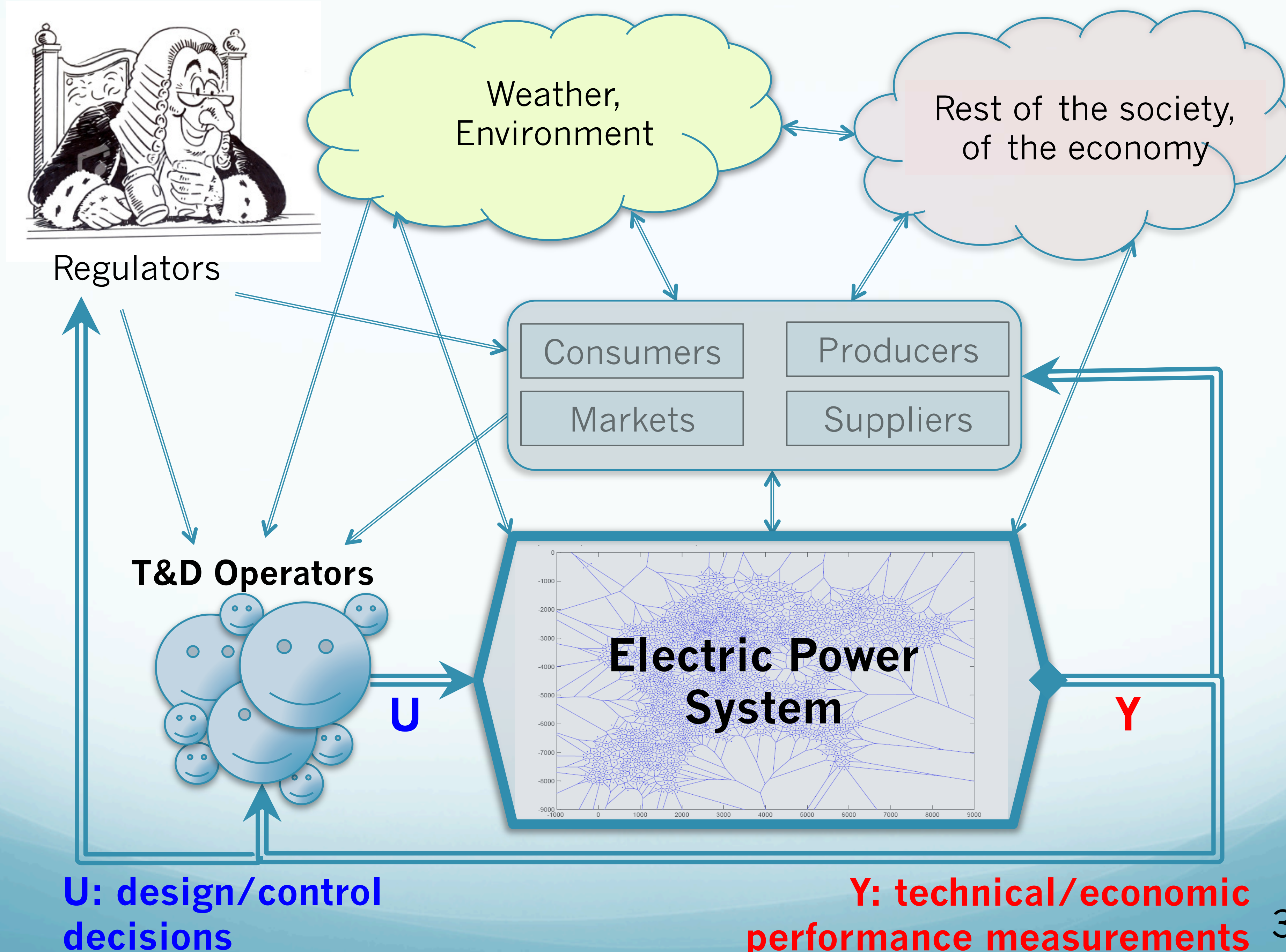


Institut Montefiore

Université de Liège

Electric Power Transmission




Regulators
Weather,
Environment
Rest of the society,
of the economy
Consumers
Producers
Markets
Suppliers
T&D Operators
U
Electric Power
System
Y
U: design/control
decisions
Y: technical/economic
performance measurements

3 different contexts of decision making

- Grid (re)design:
  - New technologies, New needs, More uncertainty
  - Towards an 'agile system design'?
- Asset management:
  - Aging infrastructure, can not be 'rebooted' nor rebuilt from scratch, time budget for maintenance and replacement
  - Towards a better modeling of ageing processes and a more effective prioritization of maintenance according to condition and criticality of asset?
- Operation and control:
  - Uncertainty, new dynamics, new control means
  - Towards probabilistic and/or robust optimization methods, exploiting more measurements, and based on better algorithms?

How can data, and data analytics, help out?

What kind of research efforts are needed?

Menu

1. About the types of models used by engineers
2. About the open modeling problems in power systems
3. Some topics for further research

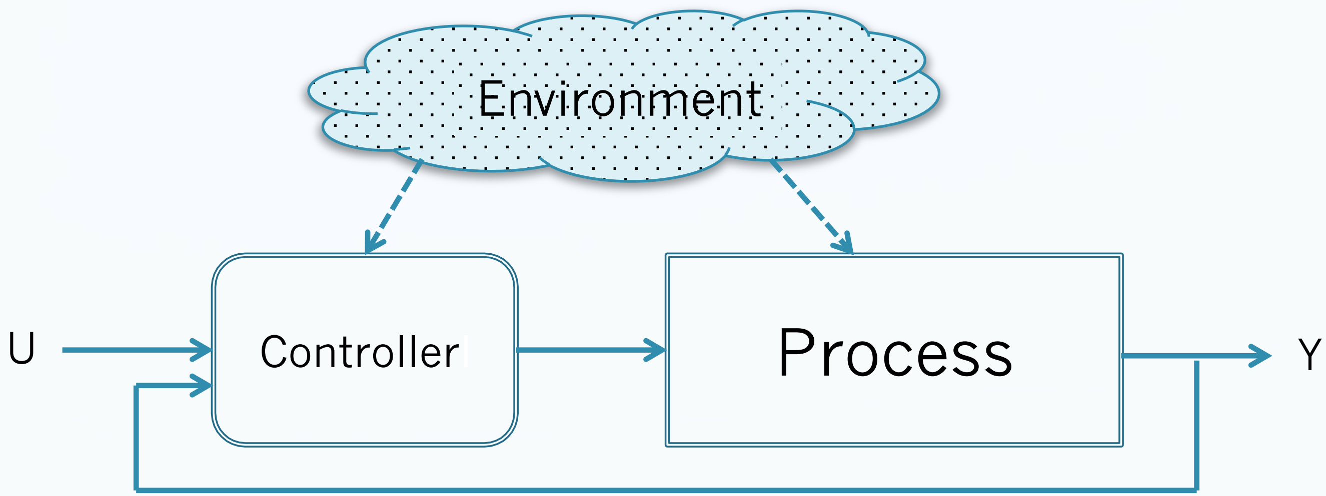
1. About the types of models

- Statistical models from observational data
  - I.e. representation of joint or conditional probability density over a set of random variables, induced from observational data
  - E.g. Gaussian processes, Markov chains/fields, etc.
  - E.g. Logistic models, Random Forests, SVMs
- Physical models from first principles
  - I.e. representation of deterministic constraints among physical quantities about a system
  - E.g. Algebraic / differential equations / PDE

Why do we need models?

- To describe testable hypotheses about the real-world behavior
- To understand and communicate knowledge about the real-world
- To take decisions on how to act in the real-world
  - To gather further information in order to validate/invalidate/refine models of a sub-system
  - To define control/optimization policies so as to modify the behavior of a sub-system when operating in the real-world

Our models should enable us to answer 3 types of inference questions



- Observational questions:
  - ✓What if we see A? (What is?) ➔ What is the probability of blackout, given that weather conditions are bad?
- Action questions:
  - ✓What if we do A? (What if?) ➔What would the probability of blackout be, if we decide to curtail some load?
- Counterfactual questions:
  - ✓What if we had done things differently? (Why?) ➔ Would there have been a 2003 blackout, if MISO & FirstEnergy computers had been working correctly?

Other desirable properties of models

- Simplicity: easy to understand, no superfluous parts
- Falsifiability: possible to verify through experiments
- Tractability: can be exploited efficiently
- Modularity: can be combined with other models
- Scalability: can be used in the real-world
- Stability: can be smoothly updated over time
- Transportability: can be used in different contexts

2. Some modeling problems that need further work

- Environment:
  - Model impact of weather, climate on generation and grid subsystems and on end-users
  - E.g. from wind/cloud forecast, to joint renewable energy and load forecast
- Socio-economic factors:
  - Model behavior, preferences of end-users, markets, societies
  - E.g. demand side response, market response
- Power system physics:
  - Understand behavior of components, refresh/revisit dynamic system models
  - E.g. ageing and failure modes of devices, dynamics of distribution subsystems

Main characteristics of these modeling problems

- Available data comes as a mixture of large sets of
  - Internal vs external data
  - Pure observational data vs information about physical structure
- In most cases the aspects to be modeled are coupled
  - confounding by weather or by other external factors
  - spatio-temporal correlations among load, generation, faults
- Sometimes there is a lack of appropriate data
  - Due to the rare nature of some extreme events
  - Due to data "censoring", as a result the past and current power system operation and maintenance policies

3. Needs for further work

- Some specific examples
  - Wind forecast, cloud forecast : can we do better than current practice? What are the fundamental limits of predictability?
  - How to build tractable Macro-models of larger subsystems (e.g. markets, overall demand response) from micro-models of their parts?
- Stability and transportability of models:
  - How to ensure that generation and load models can be easily adapted over time?
  - How to easily 'transport' such models built for one place to another place?
- Exploit data pooling in a more effective way:
  - How to combine information about different devices, about different areas of the system to build more accurate models?

A. Forecasting problems

- NB: they come in different flavors for different prediction horizons, but share some main features:
- In general, we need forecasting methods that estimate not only conditional expectations of future values, but as well their conditional distributions, so as to quantify uncertainty (needed for risk assessment, and decision making)
- Weather conditions are one of the main influencing factors, and they yield correlations among load, generation, and outage rates, both in space and in time
- How to build tractable models, from available data?
- Ideas: build on sparse or hierarchical models, tree-structured graphs, or chordal graphs ?

B. Load/demand modeling

- Better dynamic models (response to voltage/frequency variations over seconds, minutes): needed for stability analysis
- Better forecasts (over hours to years, cf preceding slide): needed for planning and operation
- Better estimation of value of lost load (i.e. end-user utility functions): needed to formulate probabilistic reliability management criteria
- What kind of models are needed to enable demand side response ?
- Ideas: consider the possibilities of novel data acquisition channels, together with optimal experiment design; active learning and reinforcement learning approaches might offer solutions. DSO – TSO collaboration is necessary to progress.

C. Problems of scarce data

- I will develop two examples
  - The estimation of remaining life-time of transmission system assets
  - The estimation of joint probabilities of multiple faults
- Both are in principle required in order to develop risk-based reliability management strategies
- NB: they are currently under investigation, in the context of the European FP7 project GARPUR
  - See http://www.garpur-project.eu

Remaining lifetime assessment (1)

- Reliability centered maintenance needs to quantify both asset health condition and its criticality for system reliability
  - Health condition of a given device depends on the history of stress (climate based, flow based, on/of cycles) and on past maintenance operations
  - EHV equipment come in technology groups, but individual elements may have quite different ‘life-styles’.
- In EHV systems, past maintenance policies have basically led to very few, if any, equipment being in its ‘terminal’ state. They did not really take ‘life-style’ of criticality into account, but are rather based on technical sheets from manufacturers.
- So, this means that better models should allow one to reduce maintenance budget at fixed reliability level, but we have a problem with data censoring.

Remaining lifetime assessment (2)

- To solve this problem requires a combination of physical models of degradation processes, of additional experimental data, and of ad hoc statistical estimation techniques.
- Furthermore, data sharing and experience sharing among TSOs should be encouraged.
- NB: The same problem also exists in distribution systems, but it seems that observations of real failures is less rare in these latter systems. Maybe we can transport some models from there to transmission systems?

Probabilities of N-k events

- Needed to develop probabilistic reliability assessment and management techniques.
- Two aspects to consider:
  - Model the impact of weather conditions and equipment health-state, on the probability of single outages
  - Find out under which conditions individual events may be treated as independent (conditionally on the weather and the health-states), and if not how to quantify then the joint probabilities of multiple events.
- NB: it may be needed to consider jointly 2, 3 or even more events, to assess correctly the actual threats.

D. Using Machine Learning to build ‘proxies’

Notion of 'proxy'

- Suppose that we are able to quite well model in detail real-time operation : typically this would be in the form of some (maybe stochastic) SCOPF formulation + some algorithmic solution heuristics.

- When taking day-ahead decisions, we will need to 'simulate' next-day real-time operation over many possible scenarios and over many different time-steps.

- This means that day ahead decision making carries the complexity of real-time decision making raised by several orders of magnitude.

- When moving to asset management and further to system planning, we talk about optimization horizons of one to several years: obviously, complexity is again raised by several orders of magnitude.

Notion of 'proxy'

- A 'proxy' for real-time operation, is a function taking as input a representation of the information state used in real-time operation, and computing as output an estimate of the result of the real-time decision making process.
- In principle, such 'proxies' could be built by using state-of-the-art machine learning algos, combined with Monte-Carlo simulation and optimization tools.
- If a good proxy is obtained, it can be used in day ahead in place of the cumbersome detailed real-time decision making model.

Backwards propagation

- Similarly, a proxy for day-ahead decision making maybe built for use in the longer horizons (asset management, system planning).
- Such a 'day-ahead proxy' would also integrate the effects modeled by the real-time proxy.
- Etc…

Further benefits

- TSOs could share such 'proxies', so as to allow each-other to take into account the needed information from other areas when taking decisions
  - Leads so some kind of horizontal coordination approach

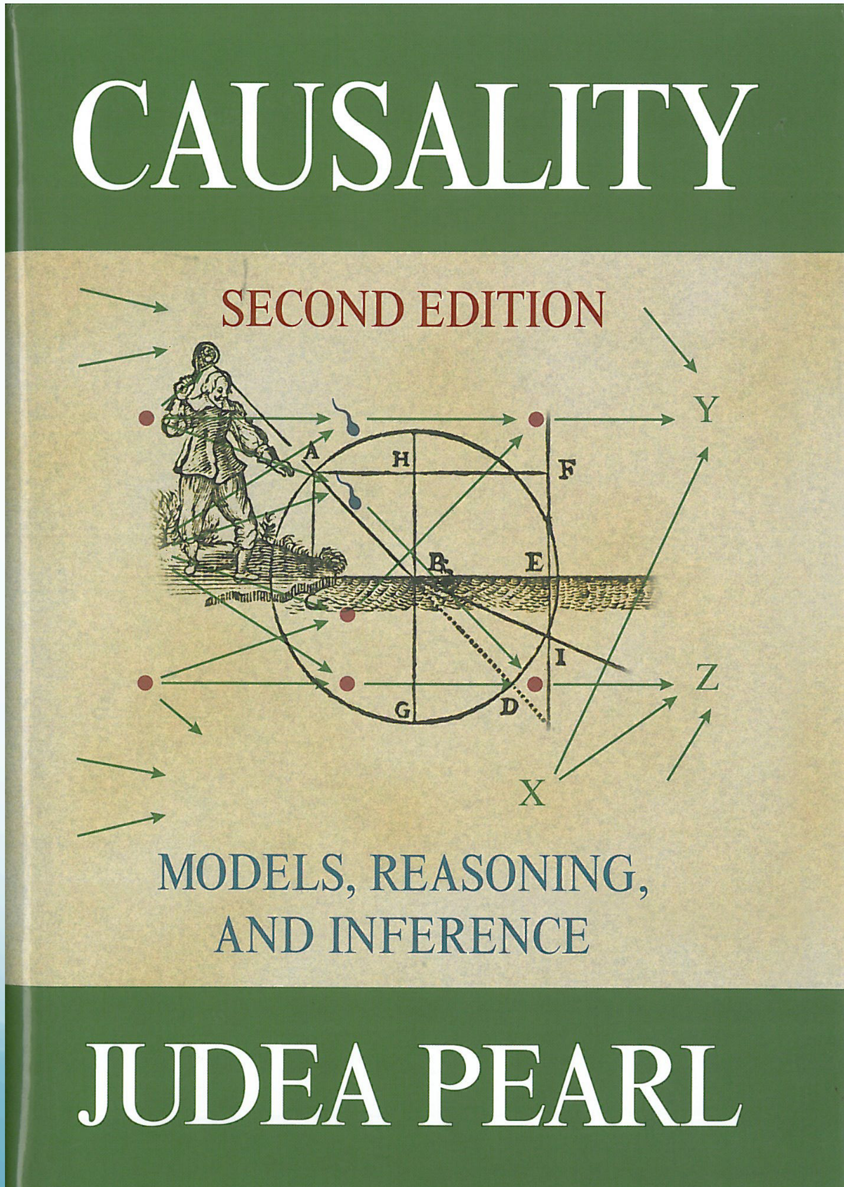
- TSOs and DSOs could as well share such 'proxies'
  - could lead to some kind of vertical coordination

- NB: such ideas are currently under investigation, in the context of the European FP7 project iTesla
  - See http://www.itesla-project.eu

Concluding comments

- We highlighted the need to have causal models, i.e. beyond what can be provided by pure statistics. To construct such models needs to hybridize physics and statistics in the proper way.
- We want to stress the need to integrate modeling, simulation and control into a single overarching activity. In this context, causal models may help to guide the exploration exploitation tradeoff.
- Machine learning may be used to build tractable 'proxies' of subsystems and of subtasks; these latter may be reused in many different contexts. How to do this in the best way requires further research.

Some further readings



Cambridge University Press, 2009

Judea Pearl, UCLA

IEEE Spectrum 1978, Fred Schweppe (MIT)

LARGE SYSTEMS

Power

Power systems '2000': hierarchical control strategies

Multilevel controls and home minis will enable utilities to buy and sell power at 'real time' rates determined by supply and demand

Because more devices for customer generation and storage of energy will be in operation by the year 2000, the customer—residential, commercial, or industrial—will be considered a vital part of the electric power systems of the future. New types of central-station generation, storage, transmission, and distribution will be available, and there will be basic changes in the total energy picture as

parallels the needs of electric power systems. Future control systems will exploit this technology extensively.

This writer's prediction of the control systems of 2000 is based on the foregoing predictions of influencing factors. The implications are that the future will see more sophisticated control systems involving many sensors and