Data Science

Smart Grid Applications

Laurent El Ghaoui
EECS Dept., UC Berkeley

Stephanie Jumel
EDF, Inc.

Agenda

1. Motivations
2. Overview of data science
3. Some smart grid applications
   – Optimization-friendly models of complex energy systems
   – Text analytics for maintenance & safety
   – Robust energy management

Acknowledgments

This work has been done in collaboration with

- Damien Faille, Hassan Bouia, Come Bissuel, Manon Fouquet, Mathieu Schuman and Benoit Pechiné (EDF).
- Soufiane Hallal, Jonathan Lacotte, Jean Pauphilet (Ecole Polytechnique).
- Andy Packard, Vu Pham, Chris Meissen (UC Berkeley).

Industry sponsors and collaborators:

EDF Research & Development

- +2,000 researchers
- Mission to improve safety, reliability & profitability of EDF Group assets:
  - Power Generation assets: Renewable Energy, Nuclear, Thermal..
  - Power Transmission and Distribution assets
  - 3rd Parties asset management
- Bring Data Science to EDF Business through applied research and demonstration projects.

Siebel Energy Institute

SIEBEL ENERGY INSTITUTE

Advancing the science of smart energy

SEI is a consortium of international universities focused on advancing data analytics research for energy systems in an open, collaborative and publicly-available manner.

SEI funds research in data analytics for the electrical smart grid, the oil and gas industry, and other modern energy systems.

http://www.siebelenergyinstitute.org/

IEEE SMARTGRID

IEEE Advancing Technology for Humanity

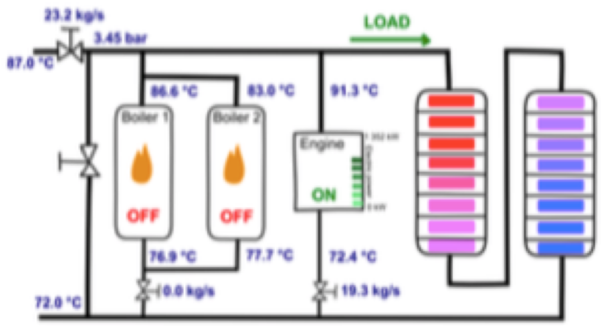
SumUp Analytics

The voice of people, markets, machines

SumUp is a startup specializing in predictive text intelligence, with solutions tailored for the energy markets.

Robust energy management

Of a Combined Heat and Power (CHP) Plant



CHP generation:

- Cheap.
- Environmentally friendly.

System has thousands of production variables.

Basic problem: adjust 24-hour production variables so as to minimize operational costs, under operational and demand constraints, with **demand not exactly known** in advance.

Design of a complex energy system

Using optimization-friendly models

Most energy systems have complex dynamics, which are highly dependent on design variables (shape of turbines, location of buildings, thickness of insulating material, etc).

Basic problem: adjust the design parameters so as to optimize the average performance of the system.

Predictive maintenance

Via text analytics

Predictive maintenance is based on installing sensors on all the relevant machinery, and collecting/analyzing petabytes of data.

- High setup costs.
- Capital intensive.
- ROI unclear.

Can we use text analytics on technician maintenance reports to help diagnose / predict maintenance or safety issues?

What is data science?

Data Science

=

Machine Learning, Statistics (Predict, diagnose)

+

Optimization, control (Act)

Analogy: driving

Outline of this brief tour…

- Representation of data
- Unsupervised learning
- Supervised learning
- Optimization models

Data sets

Today ``data'' covers many things:

- Numbers: physical measurements, prices, economic and index indicators, etc.
- Text: news, safety reports, SEC filings, PR documents, analyst reports, etc.
- Images & videos: satellite, TV broadcasts, interviews, transcripts, etc.

Many of these data sets can be put in numerical, tabular form

Example: representing text collections

Sentence: *Gold drops as China tightens, down 2 percent on week.*

Dictionary: gold, silver, china, u.s., bernanke, tightens

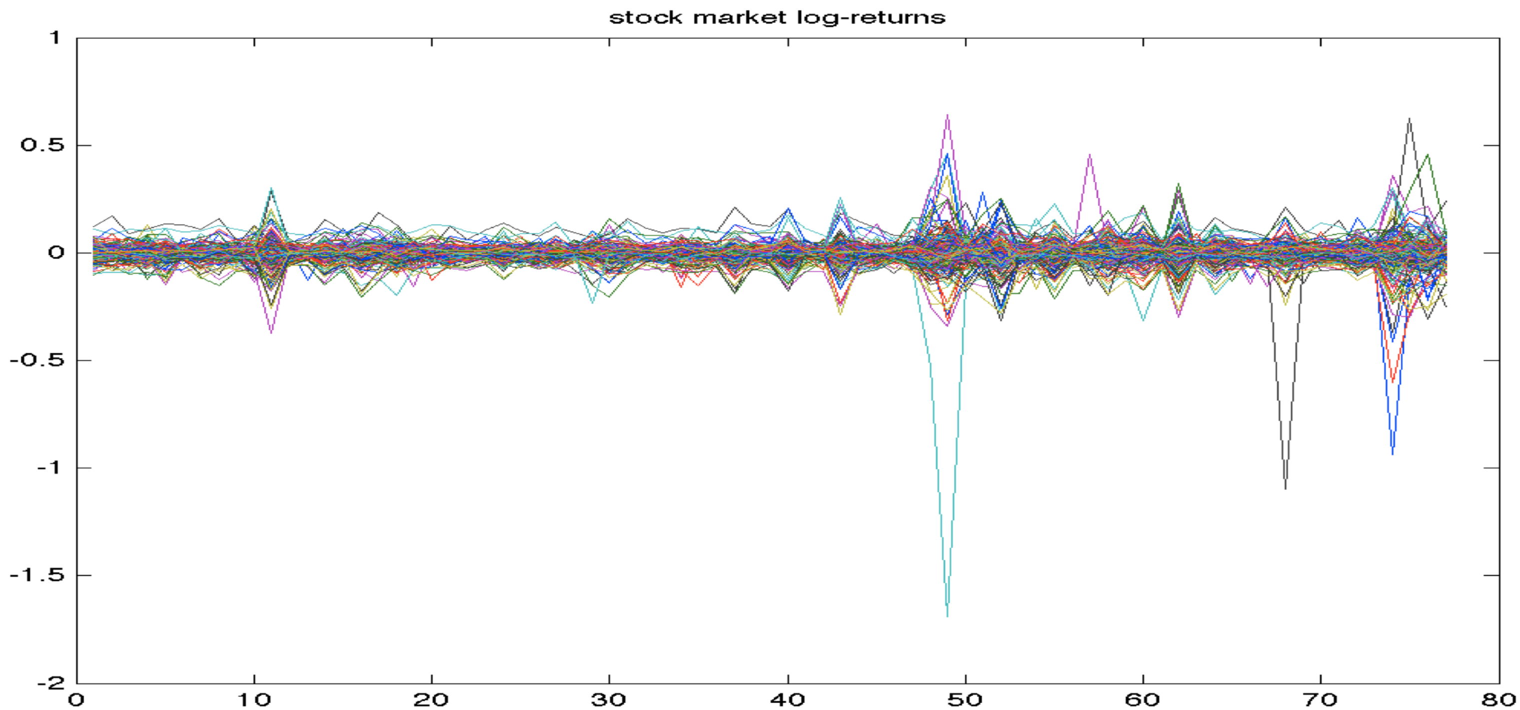
Numerical form of sentence: x = (1,0,1,0,0,1)

Any collection of documents can be represented in tabular form:

- A column represents a single document.
- A row represents the “score” of a particular term across documents.

Unsupervised learning

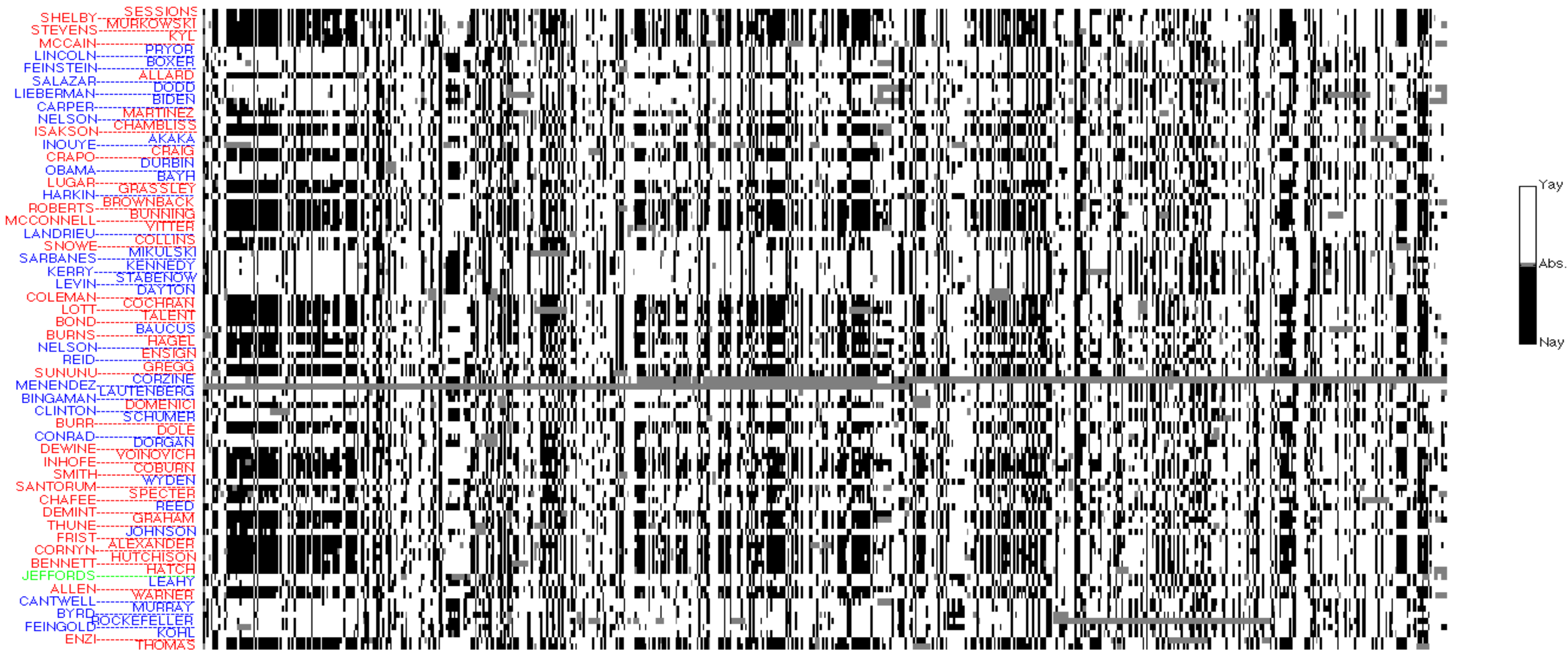
Market data



Can we make sense of this market data?

Unsupervised learning

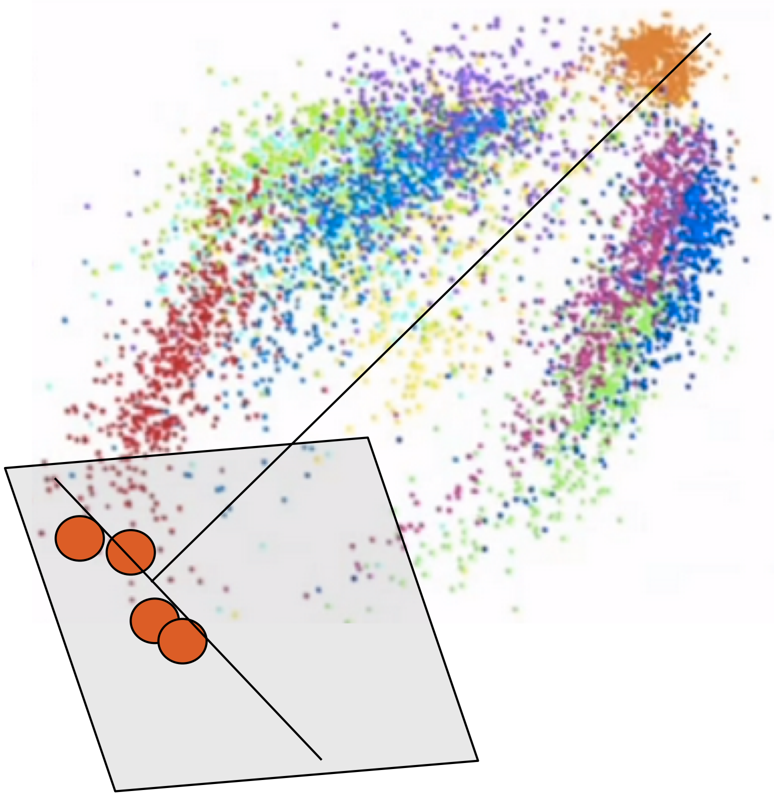
Senate voting data (2004-2006)



Are there any patterns you see?

Unsupervised learning

Principal component analysis

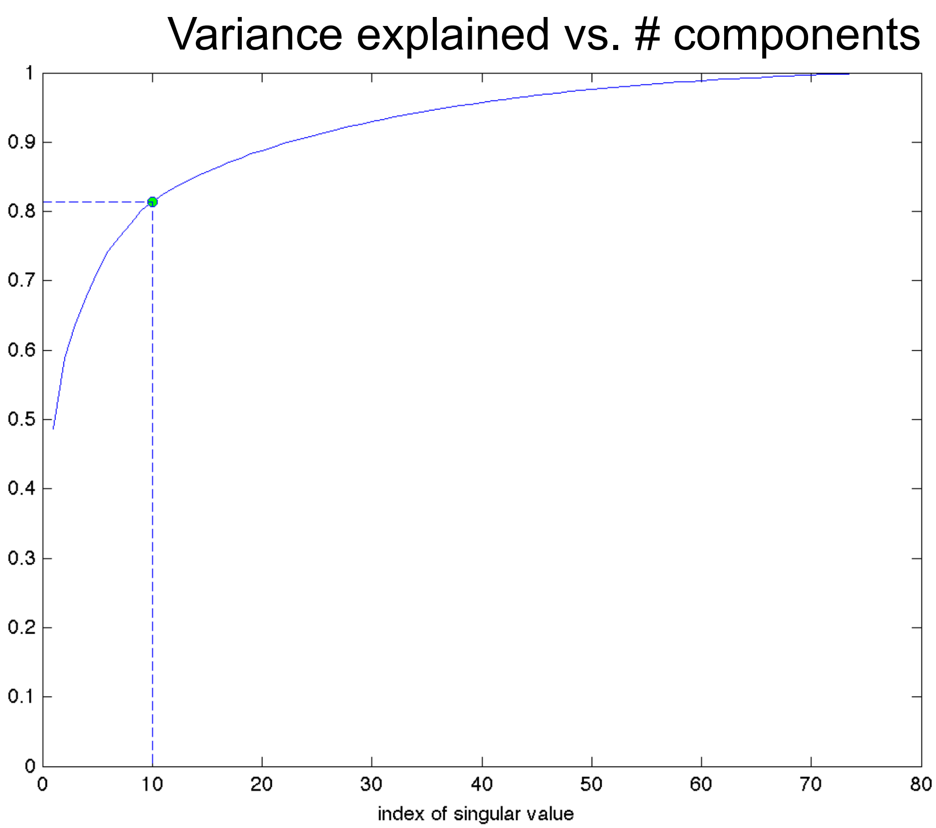


PCA algorithm:

- Find direction of highest variance
- Project data orthogonal to that direction
- Repeat on projected points
- Stop until satisfactory level of cumulative variance

Unsupervised learning

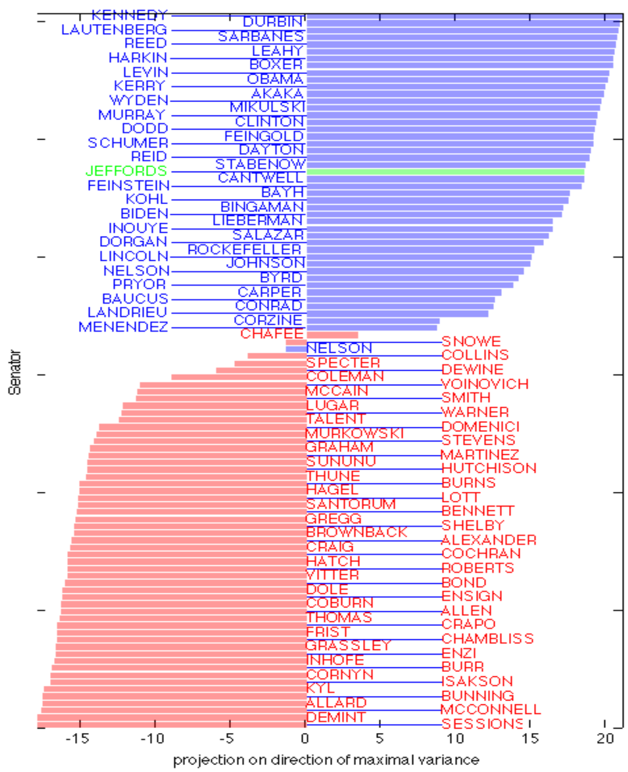
PCA of market data



- First ten components explain 80% of the variance.
- Highest component all involves troubled financials (ABC, FTU, MER, AIG, MS).

Unsupervised learning

Senate voting data: projecting data

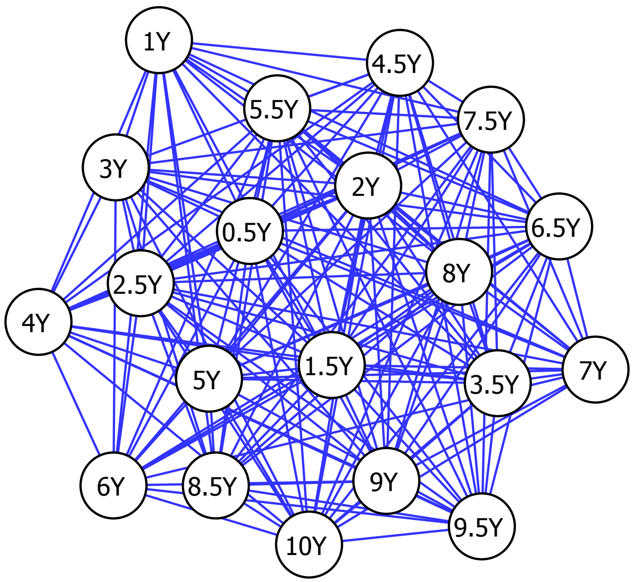


Highest-variance direction recovers party line!

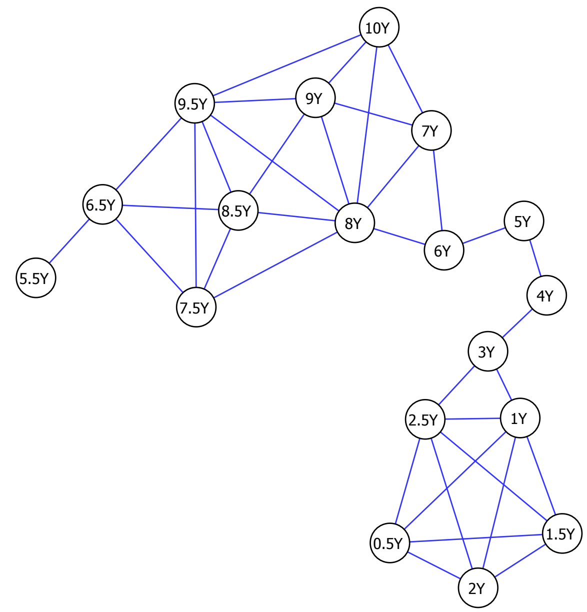
Allows to score Senators.

Unsupervised learning

Beyond PCA: graphical model for interest rate data



Correlation graph:
All assets are correlated



Conditional independence graph:
Discovers structure

IEEE SMARTGRID

IEEE Advancing Technology for Humanity

Supervised learning

Overview

In supervised learning, data points come with “side” information:

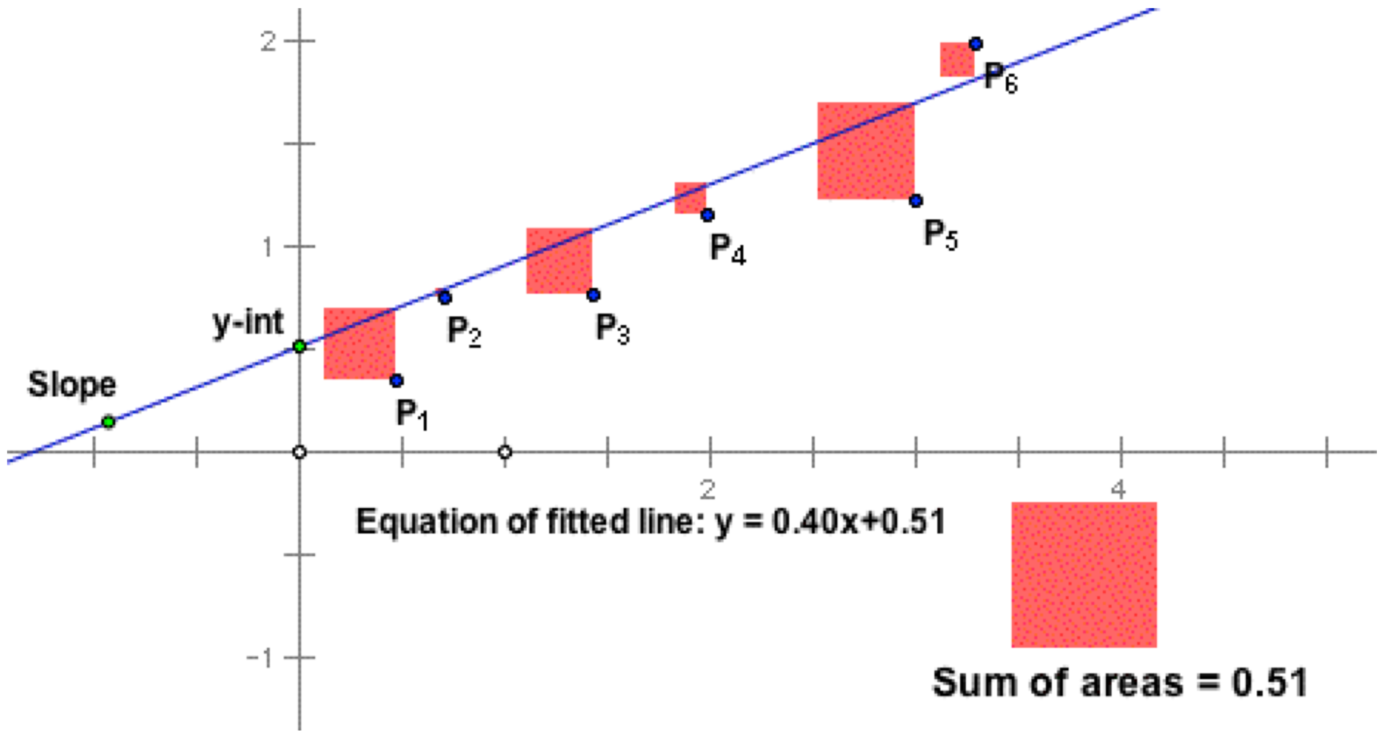
- Real numbers
- Binary
- Other

Goal is to predict the “side information” for a new data point.

Example: based on time-series data, predict failure of an equipment.

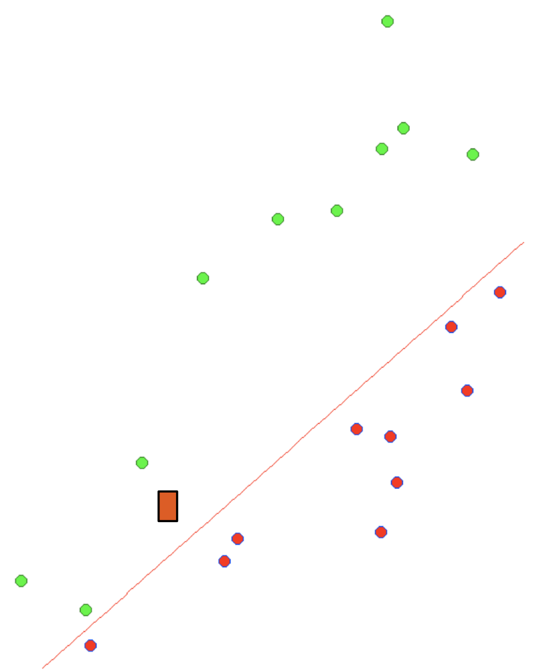
Supervised learning

Model fitting



Supervised learning

Classification



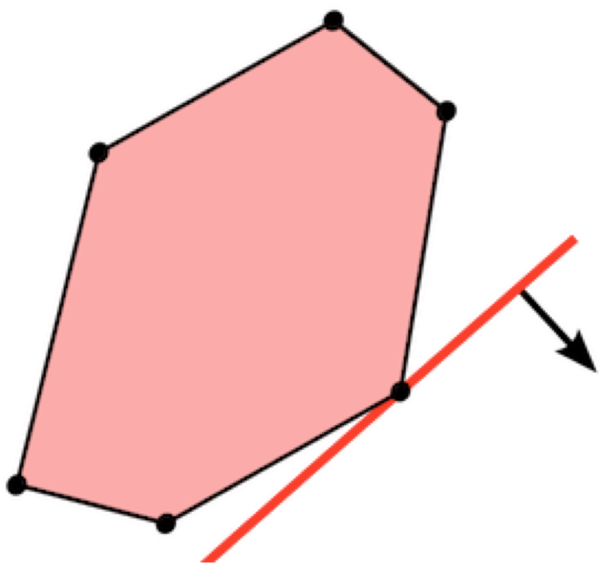
Data set comes with labels

- Task: predict the label of a new point
- Method: separate training data with a plane
- Result: predict label of a new point based on which side it falls

In **sparse learning**, we also need to identify the key features that are involved in the prediction.

Optimization

Linear programming model



Linear program:

$$\min_{x} \; c^T x \text{ subject to } Ax \geq b$$

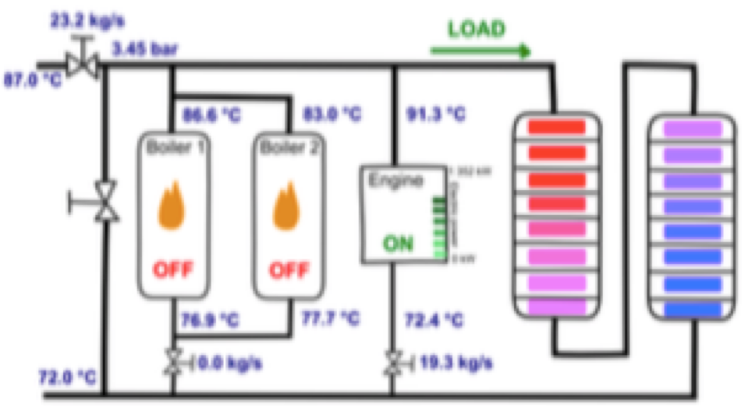
- x is a vector of "decision variables"
- Constraints are linear on x

LPs and variants can be used to describe many decision problems, e.g. energy management or optimal design of engineering systems.

Example

Model of energy production

Linear program: $\min_{x} \; c^T x \text{ subject to } Ax \geq b$



In energy production applications:

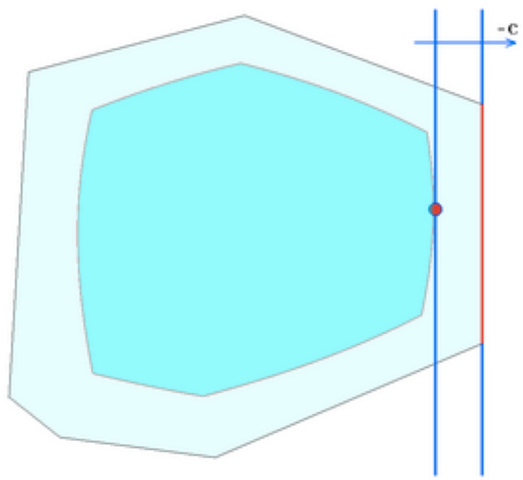
- x is a vector of production variables.
- Ax represented produced items.
- Vector b represents demand.
- Vector c represents cost.

Some variables can be binary, leading to a more complex model.

Robust optimization

Robust LP

In practice, data in LPs may not be exactly known at solution time.



Robust linear program:

$$\min_x \; c^T x \text{ subject to } Ax \geq b \text{ for all } b \in \mathcal{B}$$

- Vector b is unknown, but bounded in a set ***B***
- Constraints are satisfied for every b in set ***B***

IEEE SMARTGRID

IEEE Advancing Technology for Humanity

Summary

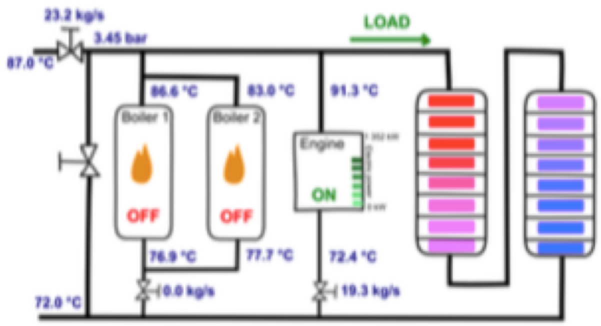
Which data science for which case?

To apply data science in practice:

Task	Model to use
Diagnose, understand big picture Prepare data for other tasks	Unsupervised learning
Predict	Supervised learning
Automate decisions	Optimization

Robust energy management

Of a Combined Heat and Power (CHP) Plant



CHP generation:

- Cheap.
- Environmentally friendly.

System has thousands of production variables.

Basic problem: adjust 24-hour production variables so as to minimize operational costs, under operational and demand constraints, with **demand not exactly known** in advance.

Robust energy management

Variables and constraints

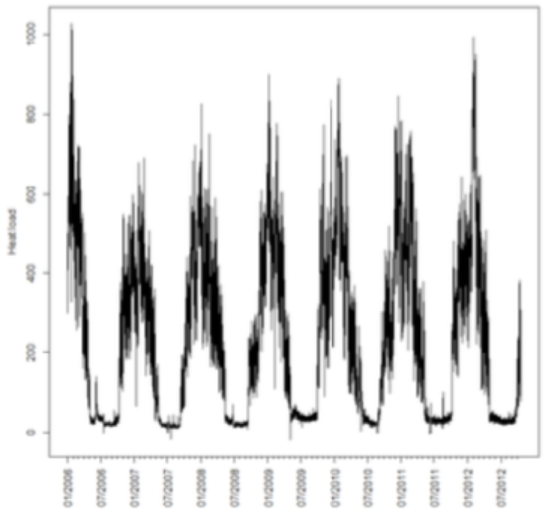
The production vector x contains

- 840 binary variables which represent all-or-nothing states (On/Off mode),
- 672 continuous variables which represent the amount of gas consumed by the equipment or the level of energy produced, etc.
- Among the 2,518 constraints, only 24 of them are subject to uncertainty: Every hour, we impose that the heat delivered to the network satisfies the heat demand.

Building the model

Fitting uncertainty model to historical data

Simple uncertainty model: $\mathcal{B} = \left\{ \hat{b} + Bu : u \in \mathbb{R}^p, \|u\|_\infty \leqslant \rho \right\}$



- b now corresponds to the estimated demand vector (over the next 24-hour period).
- Matrix B allows to account for “ripple effects” in uncertainty.

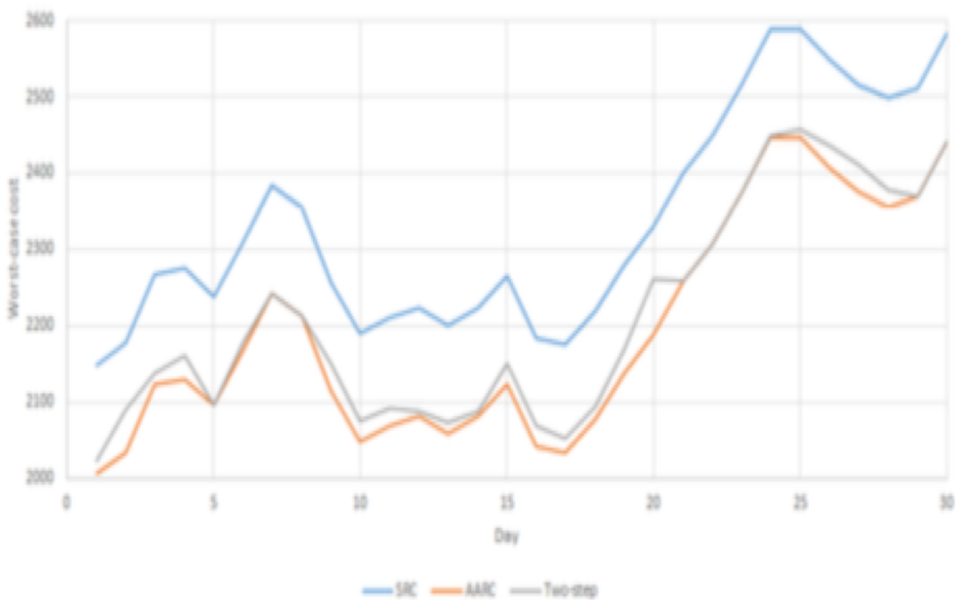
Model can be fitted with supervised learning methods.

IEEE SMARTGRID

IEEE Advancing Technology for Humanity

Results

In terms of worst-case cost



Measuring performance via "worst-case" cost (over allowable uncertainty):

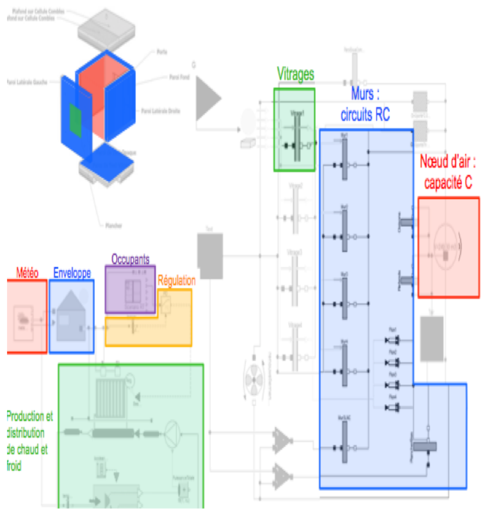
- Original LP model can lead to unsatisfied constraints.
- Robust LP and more sophisticated variants bring down worst-case cost greatly.

IEEE SMARTGRID

IEEE Advancing Technology for Humanity

Design of a complex energy system

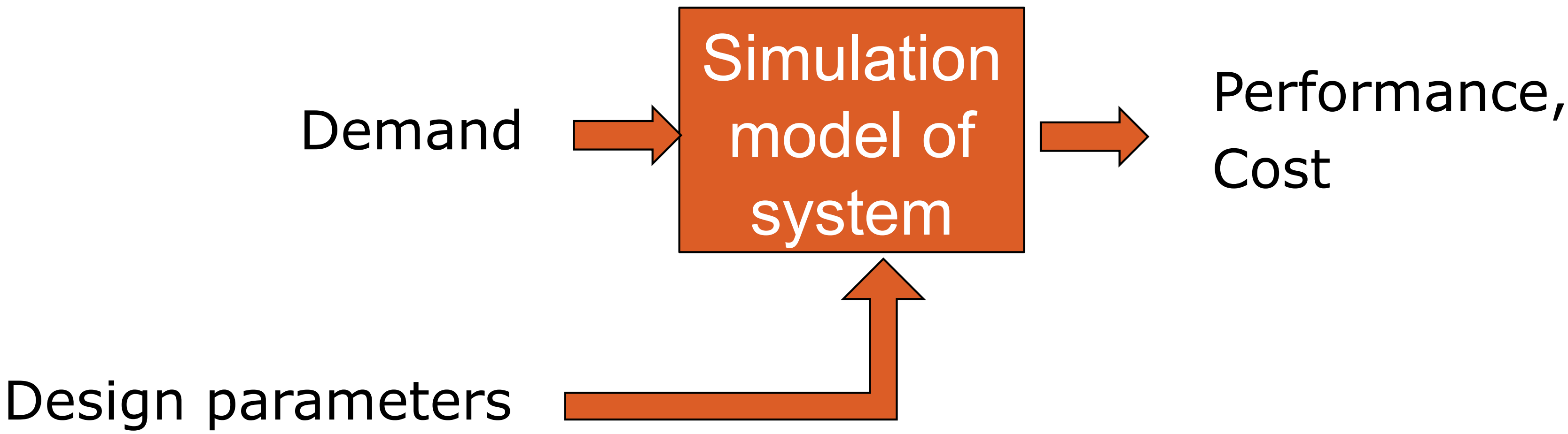
Renovation of a set of 22 buildings



Goal: For a set of 22 buildings, optimize renovation parameters (wall thicknesses, isolation, etc), so as to minimize average energy consumption, taking into account

- Constraints on parameters.
- Temperature (comfort) limits.
- Uncertainty on future demand.

Classical approach



Classical approach:

- Setup a (complex) **simulation** model as accurate as possible.
- Optimize the simulation model by exhaustive or heuristic search in parameter space.

Challenges in classical approach:

- Parameter search is complicated and time-consuming.
- Might produce spurious optima, or not converge.

Optimization-friendly approach

Basic idea

Model system in a way that guarantees that the parameter search is easy.

- So-called "posynomial models" are examples of models that can effectively optimized.
- They generalize "power laws" that govern (or, accurately model) many physical systems.

$$f(p) = x_1 p_1^{a_{11}} p_2^{a_{12}} + x_2 p_1^{a_{21}} p_2^{a_{22}}$$

Optimization-friendly approach

Procedure

- Use a (complex, "un-friendly") simulation model to obtain a set of input-output data.
- Model the input-output behavior by a (friendly) **proxy** model.
- Optimize the proxy model.

The approach involves:

- Sampling the parameter space;
- For each set of parameter values, get an input-output pair.
- From those pairs, learning the model.

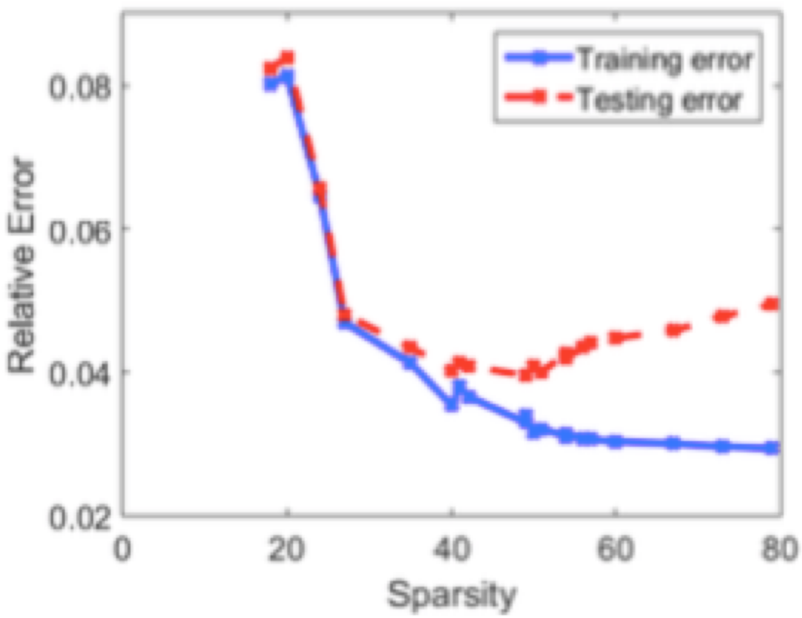
Getting simulation data

Using Dymola simulation model

- n = 3, 174 simulations.
- n sample points split into training and validation set (50-50).
- Simulated 2,4192,000s (28 days) with Δt = 3,600s.
- Latin hypercube sampling in [0, 1] for parameters (5 of them)

Model Fitting Results

With a posynomial model



Using sparse machine learning we identified a model with **44 terms** that is ~4% accurate, for both energy and temperature.

This is the function below, with **specific** values of *x*, *a*, and n=44:

$$f(p) = x_1 p_1^{a_{11}} p_2^{a_{12}} \ldots p_5^{a_{15}} + \ldots + x_n p_1^{a_{n1}} p_2^{a_{n2}} \ldots p_5^{a_{n5}}$$

Parameter optimization

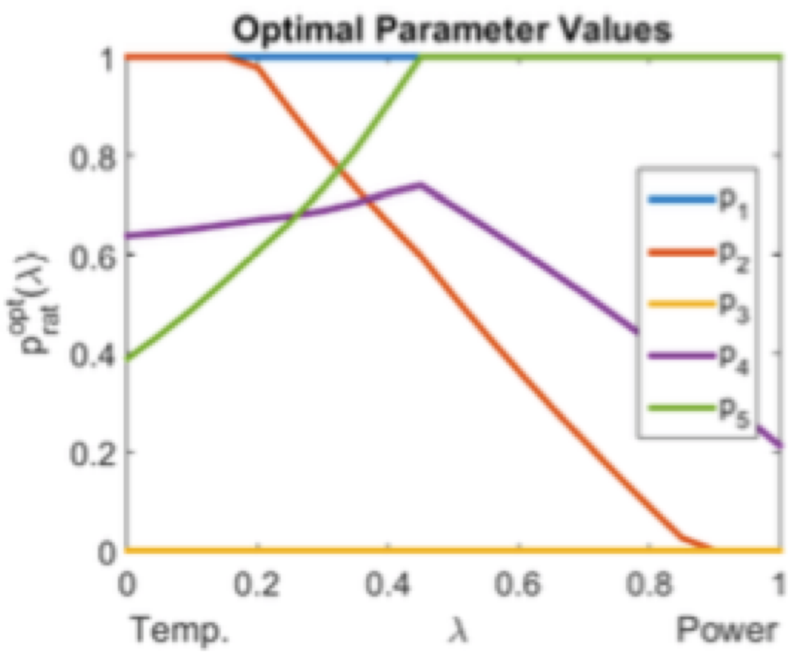
With a posynomial model

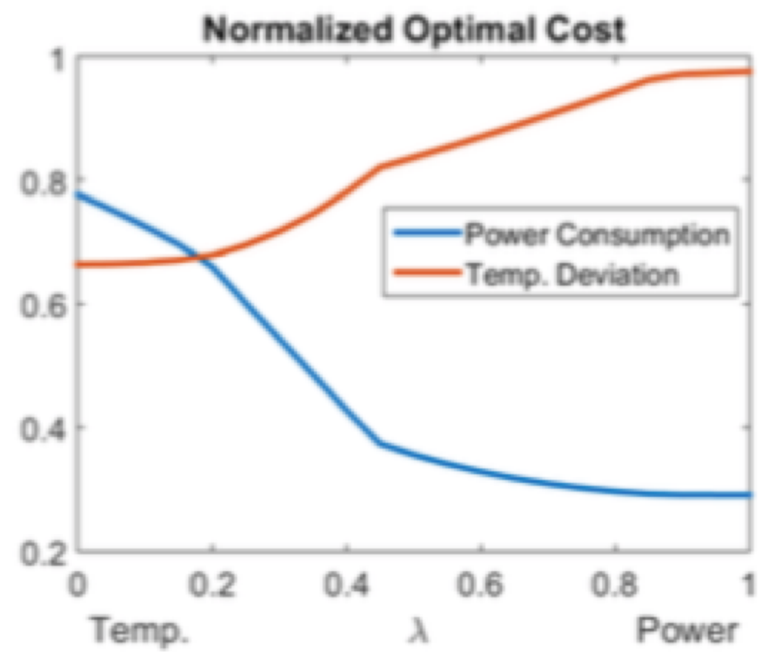
Optimization problem:

$$\begin{array}{ll} \underset{p \in \mathbf{R}^5}{\text{minimize}} & \lambda \frac{\hat{E}(p)}{\bar{E}} + (1-\lambda)\frac{\hat{M}(p)}{\bar{M}} \\ \text{subject to} & p \in \mathcal{H} \end{array}$$

- p: vector of parameters
- Set **H** describes constraints on p
- *Ê, M* are proxy energy and temperature functions

Parameter optimization results





- Clear trade-off between power consumption and temperature deviation
- Power-temperature trade-off depends on p_2, p_4, and p_5
- Trade-off does not depend on p_1 and p_3

Parameter optimization results

With a posynomial model

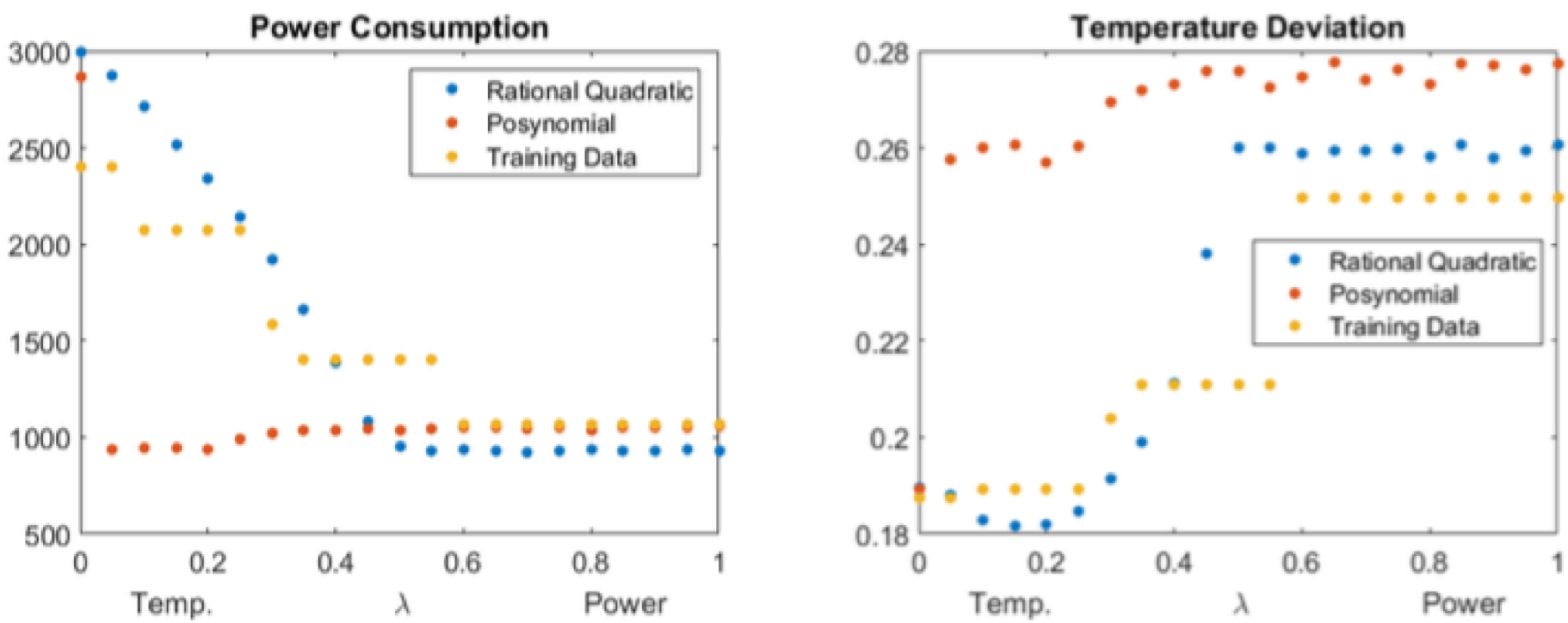


Figure 14: Dymola simulation of optimal parameters versus sampled data.

IEEE SMARTGRID

IEEE Advancing Technology for Humanity

Unsupervised learning

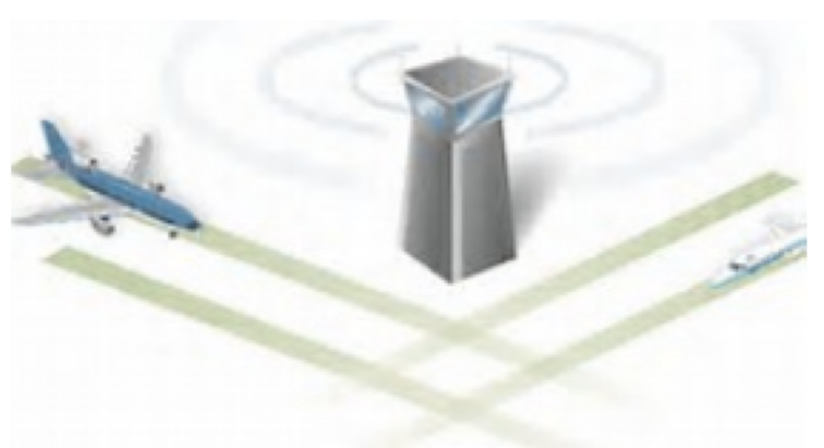
Text data

ASRS data:

A collection of ~25K reports on flight safety written by commercial pilots in the US, maintained by NASA.

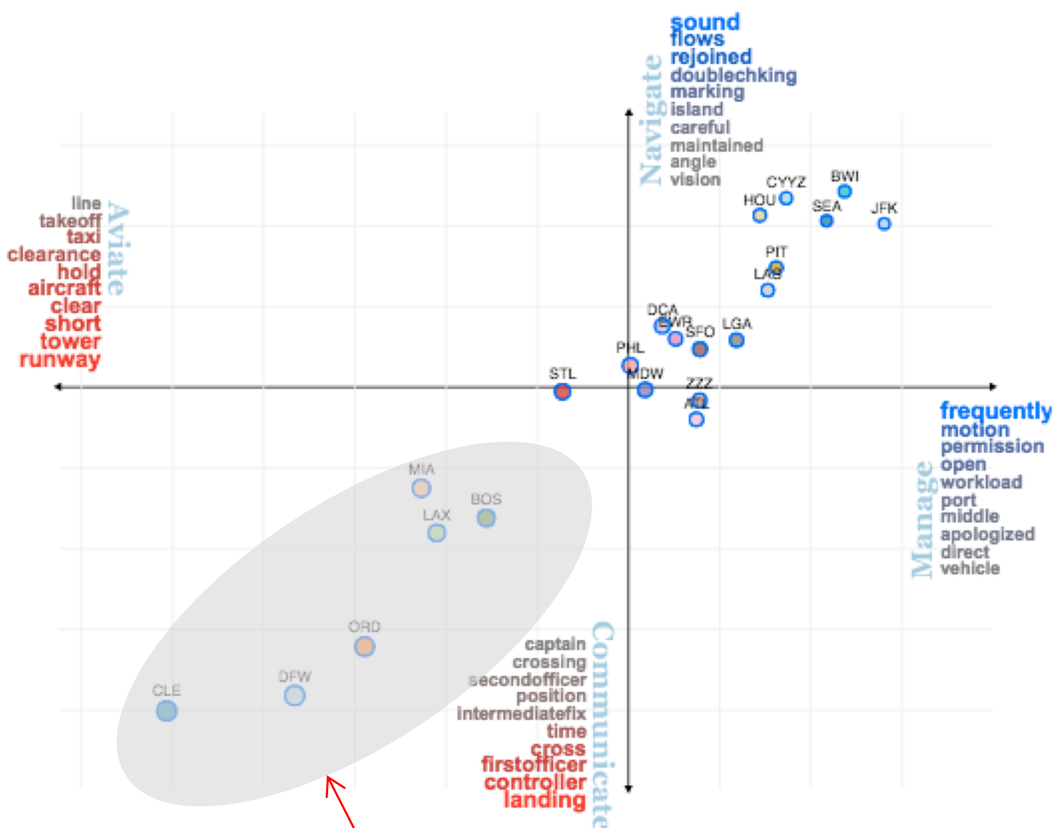
Goals:

- Understand and diagnose issues.
- If possible, predict incidents.



Unsupervised learning

PCA of ASRS data



Highest variance directions correspond to four main pilot tasks:

- Navigate (fly)
- Aviate (on runway)
- Communicate (with tower)
- Manage

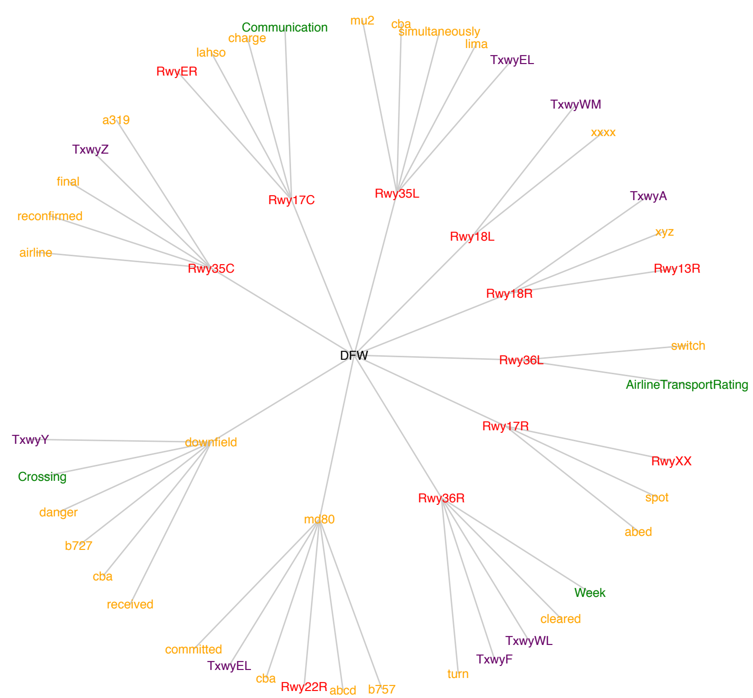
Communications / runway issues predominant in big airports

IEEE SMARTGRID

IEEE Advancing Technology for Humanity

Supervised learning

Sparsity: learning the relevant features



Goal: Analyze the relevant features in classifying reports from one airport against all others

- At DFW we find the terms “Rwy36R” and “TxwF”.
- This corresponds to an intersection with lots of near-miss collisions, due to lack of visibility from Tower.

Conclusion and perspectives

- Methodology
  - Estimating demand uncertainty based on statistical models
  - Optimization-friendly models: Kriging (sampling) methods

- Applications, looking forward
  - Text mining approach to support predictive maintenance of power assets.
  - Applications of complex modeling / control to Electric Grid

References

- Chris Meissen, Vu Pham, Laurent El Ghaoui, Andy Packard. “Posynomial Modeling for Power Consumption Optimization”, EECS Technical report, UC Berkeley, 2016.
- Jean Pauphilet, Diego Kiner, Damien Faille, Laurent El Ghaoui. “A Tractable Numerical Strategy for Robust MILP and Application to Energy Management”, to appear in IEEE Conf. Decision & Control, Dec. 2016.
- El Ghaoui, Laurent, et al. "Understanding large text corpora via sparse machine learning." Statistical Analysis and Data Mining 6.3 (2013): 221-242.
- D. Faille, M. Fouquet, B. Péchiné, S. Ahallal, L. El Ghaoui, J. Pauphilet. Robust planning for combined heat and power production. IFAC 2015.
- Laurent El Ghaoui. “Optimization models”, Livebook available upon free registration athttp://livebooklabs.com/keeppies/c5a5868ce26b8125/6267fbb41c3f1521
- Siebel Energy Institute : related article on the Energy Institute website: http://www.siebelenergyinstitute.org/machine-learning-tool-manages-power-generation-efficiently-human-operators/

THANK YOU!
QUESTIONS?

Today's presentation will be made available on the IEEE Smart Grid Portal

Smartgrid.ieee.org

IEEE SMARTGRID

IEEE Advancing Technology for Humanity