Energy data analytics: how software is changing the energy industry

Karen Yu

Presented at Harvard Energy Journal Club

25 February 2016

Software can improve the efficiency of our energy system

Demand side
- Building efficiency
- Demand response

Supply side
- Electric load forecasting
- Wind power forecasting
- Solar power forecasting
- Predictive maintenance

"Smart Grid"

Business and arbitrage
- Price forecasting

Topics

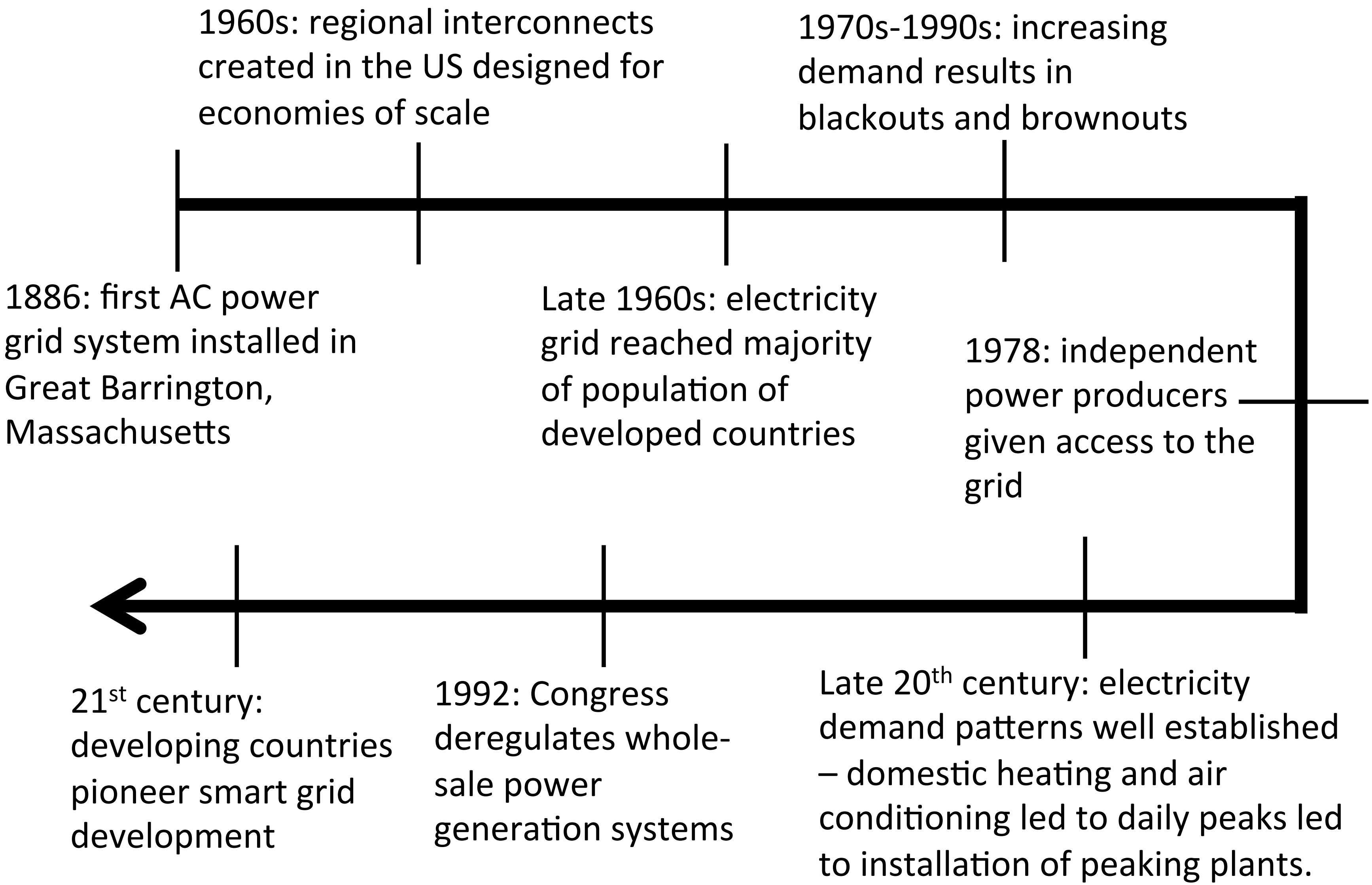
1) What is the smart grid and why do we need it?

2) What is machine learning and data analytics?

3) What are the computer science and software challenges for implementing the smart grid?

4) What have companies done so far to address this challenge?

Brief history of the electric grid



The current electric grid has some severe limitations

- Low utilization of peaking generators leads to unnecessary redundancy
- Metering capabilities limit the degree to which price signals could be propagated through the system.
- Difficult to incorporate intermittency of renewable energy and ensure reliability
- Centralized power stations may increase vulnerability to terrorist attacks

Some precursors to the smart-grid attempted to address these issues:

- 1980s: automatic meter reading to monitor loads from large customers
    - Prevents physical trip to read meter
    - Allows for more accurate monitoring data
- 1990s: advanced metering infrastructure
    - Provide information about usage at different points in time
    - Allows for 2-way communication with meter so that information such as time-of-use pricing can be sent to the home

The smart grid doesn't refer to a specific technology, but rather a set of related technologies that include both physical infrastructure and digital applications

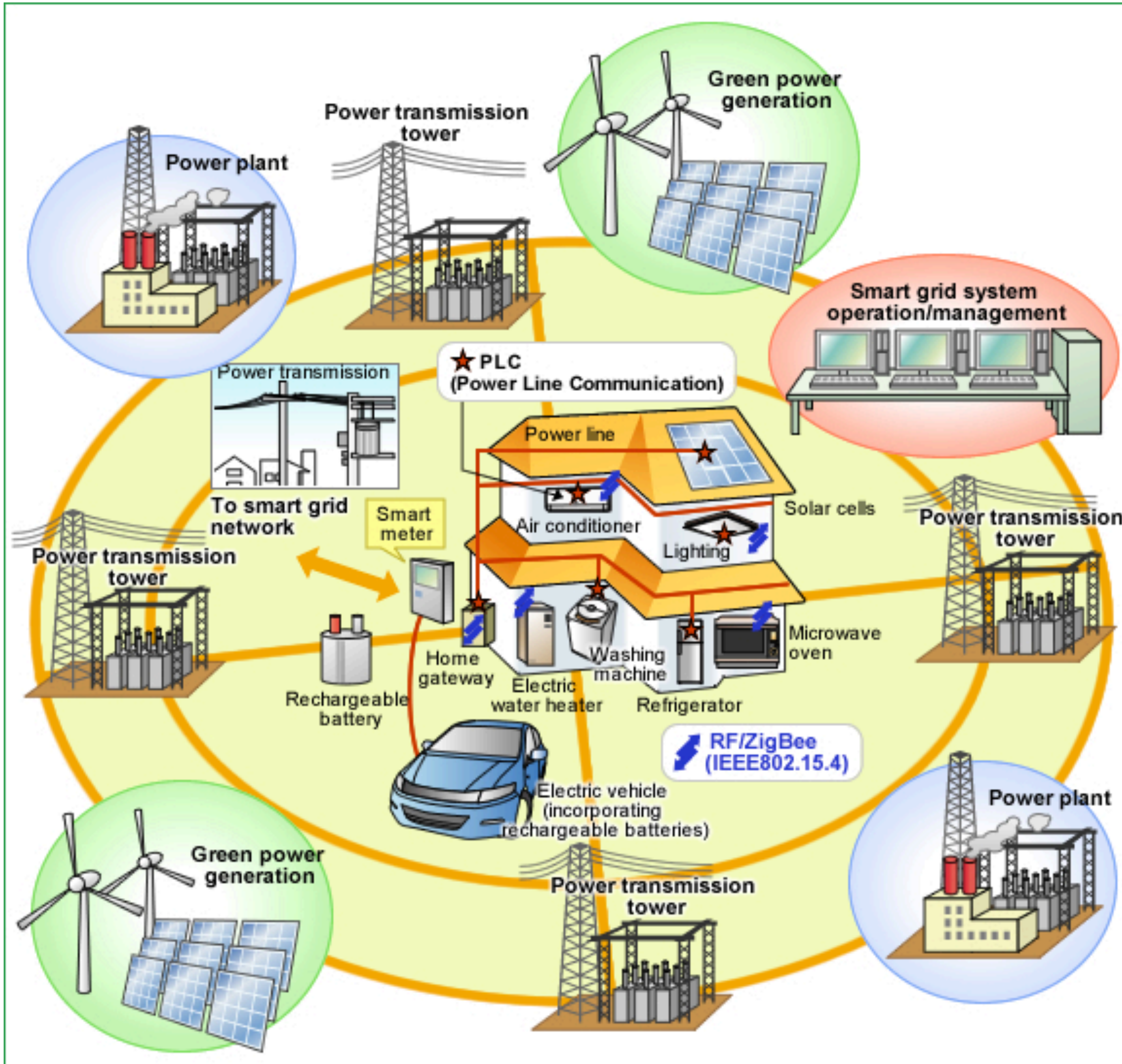


Image source: http://www.renesas.eu/ecology/eco_society/smart_grid/

Although there are numerous infrastructure and hardware challenges to implementing the smart grid, this talk emphasizes the role of software in the smart grid system

https://www.youtube.com/watch?v=t5OzUIpQWpM

Progress in infrastructure upgrades has been slow in part due to strong opposition:

- Concerns over privacy
- Concerns about fair availability of electricity
- Complex rate systems reduce clarity and accountability
- Remotely controlled kill-switch incorporated into smart meters

Software can be used to circumvent some “hardware” challenges

Problem:
Want to know how much energy each appliance is using

Hardware solution:
Attach monitors to each appliance

Software solution:
Measure electricity use at one point, then apply machine learning algorithm to disaggregate into components.

bidgely

- Recently raised $16 million
- Has ~ 20 large utility clients, including TXU, ComEd, London Hydro

https://www.youtube.com/watch?v=MSvzNlTIcfw

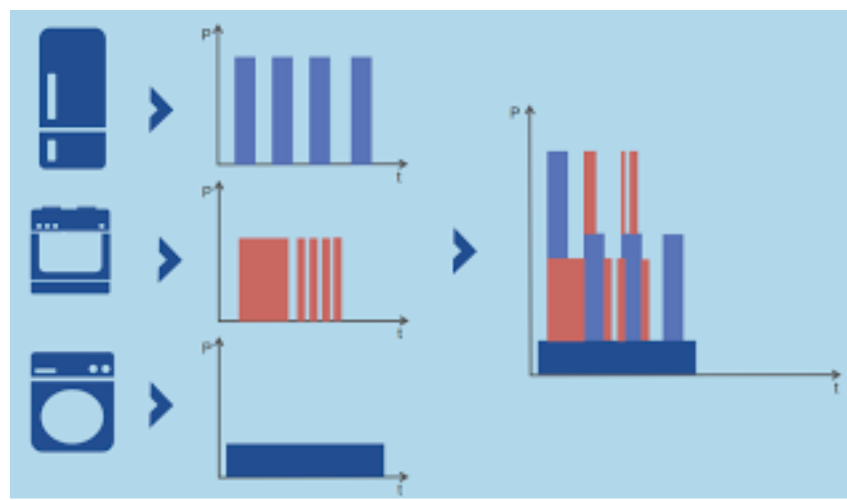
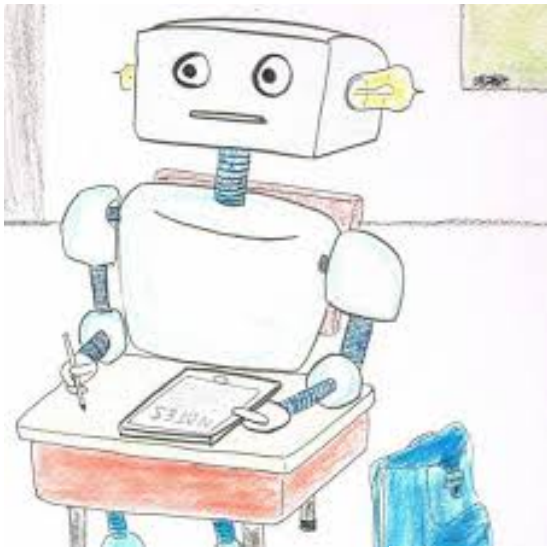


Image: http://smartmicrogrid.blogspot.com/2013_11_01_archive.html

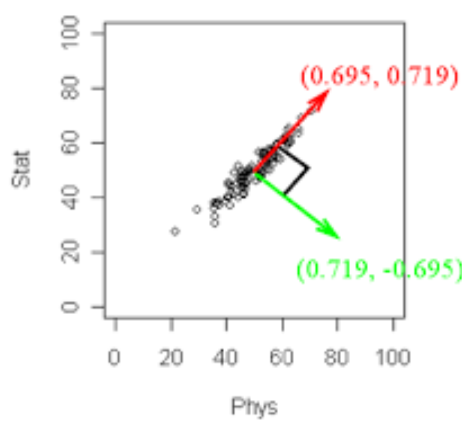
Machine Learning allows us to process large data sets with less human input



Introductory video:
https://www.youtube.com/watch?v=bHvf7Tagt18

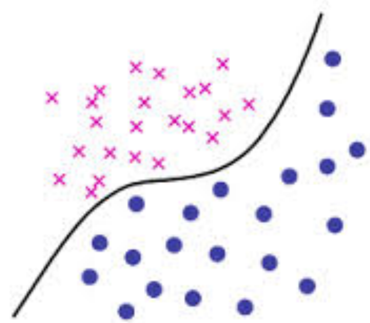
Categories of problems:

Dimensionality reduction



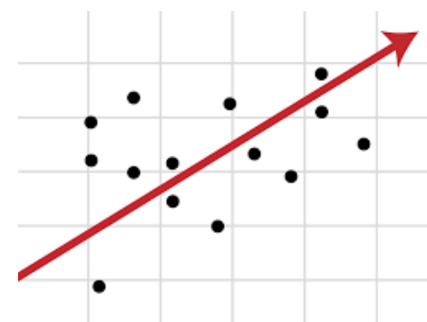
http://astrostatistics.psu.edu/su09/lecturenotes/pca.html

Clustering/ classification



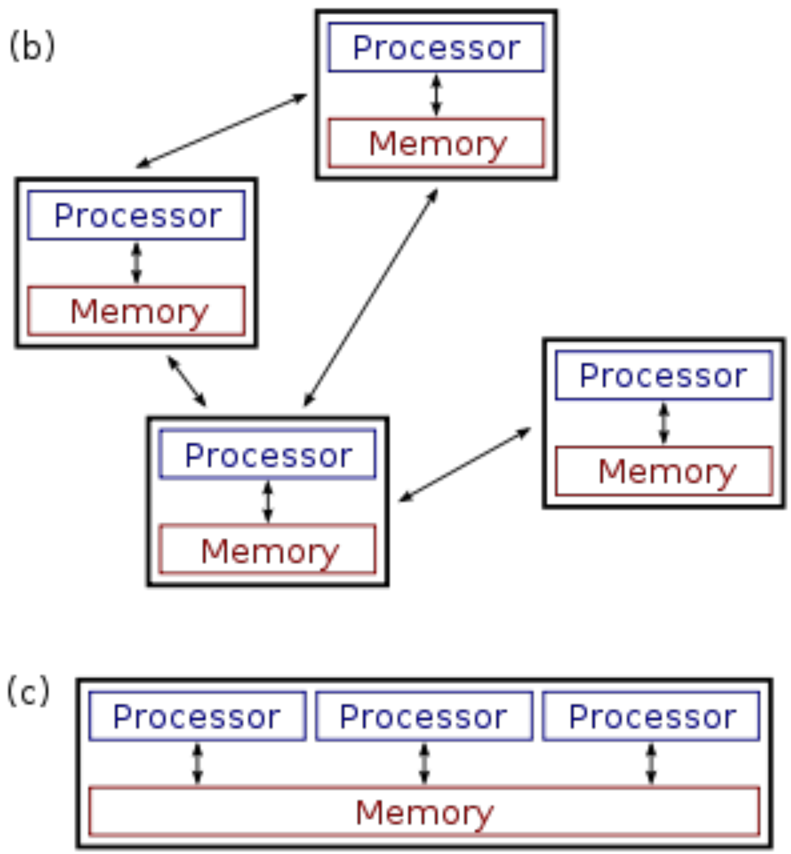
http://www.hact.org.uk/blog/2014/05/27/big-data-and-housing-part-3-machine-learning

Regression



http://news.mit.edu/2010/explained-reg-analysis-0316

Distributed computing makes handling these large data sets possible



Source:Wikipedia

Commercial services make distributed computing accessible

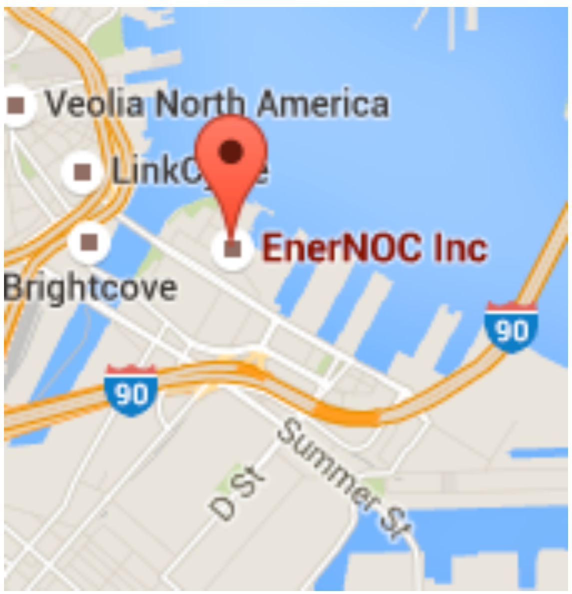
amazon web services™

Microsoft Azure

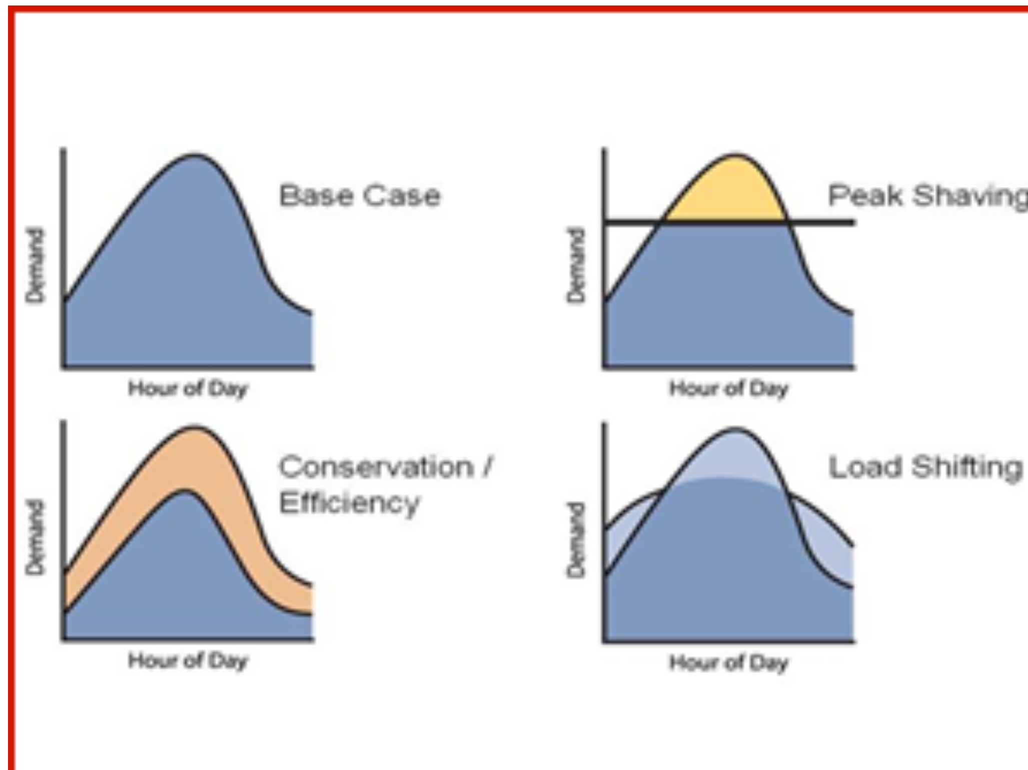
hadoop

Demand-side management

Traditional grid:
Supply is adjusted to match demand



Future grid:
Demand should be made more adaptive to supply conditions
Can be passive (automatic) or active (ask user to turn off device)



http://www.powerwise.gov.ae/en/research/programmes-projects/demand-side-management.html

Demand-side management consists of elements with varying levels of technical complexity and user participation

Passive: Utility automatically cuts off certain loads when demand is high.

Active: Utility asks commercial and industrial customers to curtail usage during peak hour and offer compensation in return

Better software can make these processes more automatic and efficient.

Price-based: real-time pricing discourages customers from using electricity when prices are high

Electric vehicles can act as both a demand-side and supply-side resource

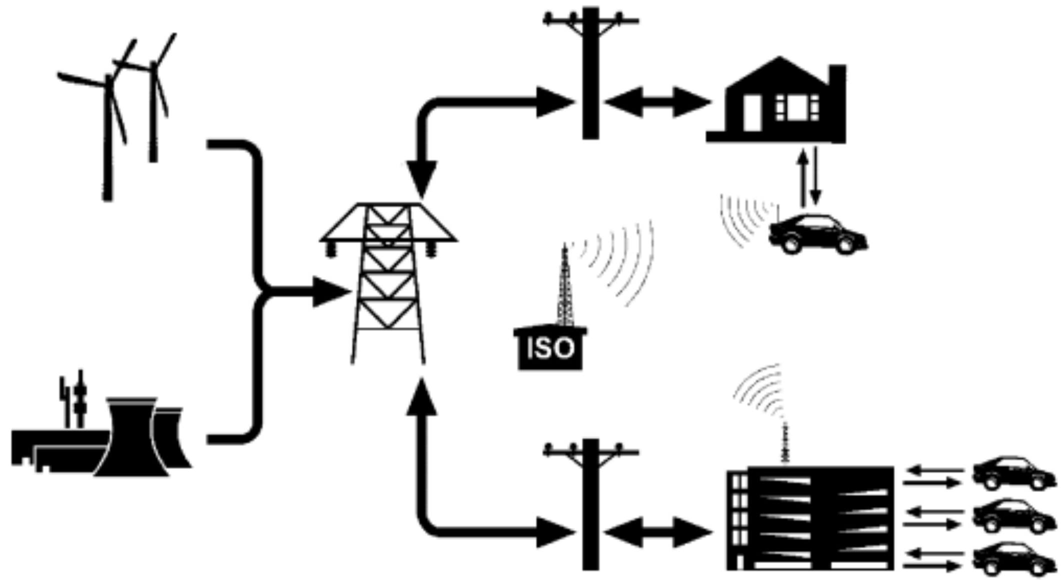


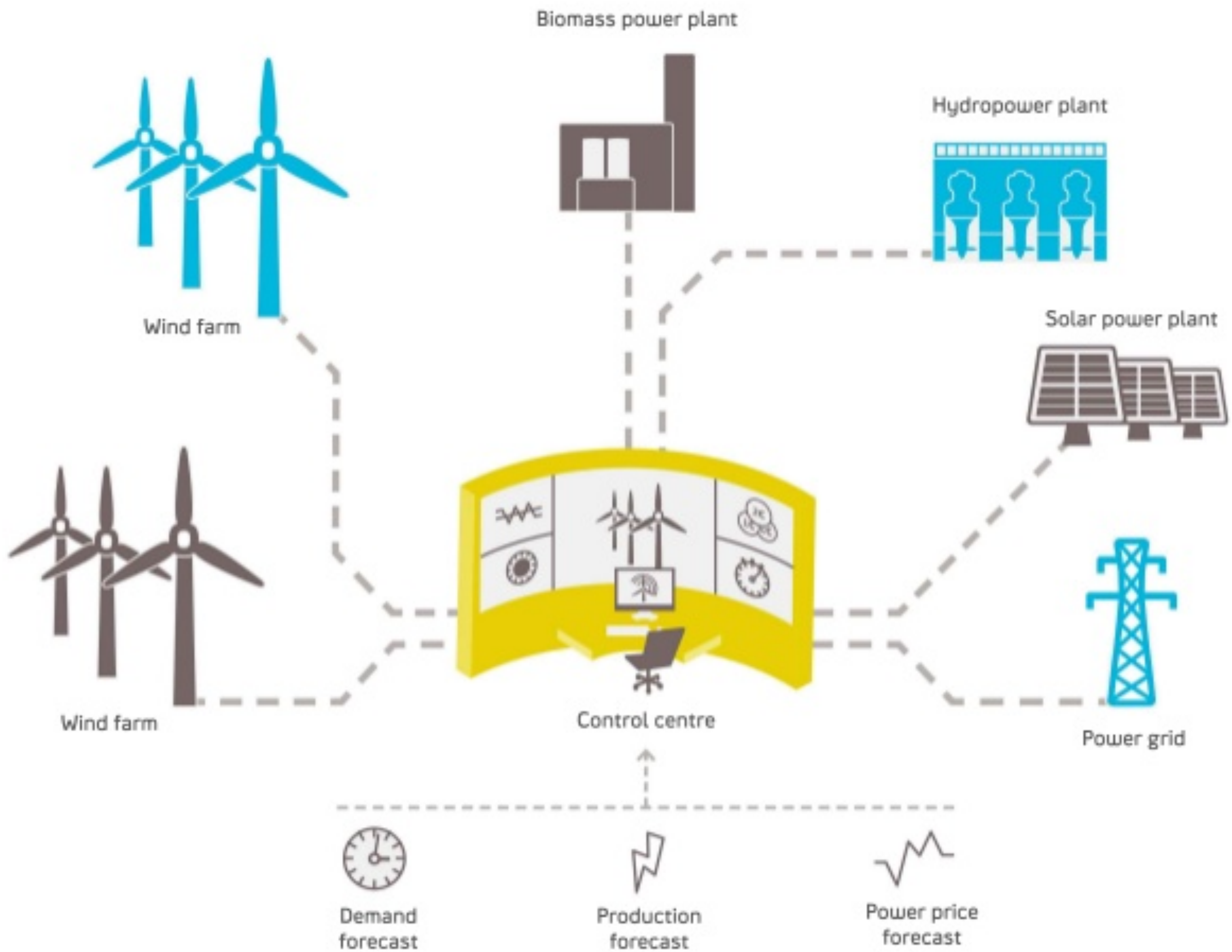
Fig. 1. Illustrative schematic of proposed power line and wireless control connections between vehicles and the electric power grid.

Kempton and Tomic, 2005

Software challenges:

- Predicting individual users' EV charging needs and designing algorithms to optimize charging cycles of EVs based on users' predicted needs
- Predicting aggregate EV charging demands at different points in the network
- Designing decentralized control mechanisms that coordinate the movement of EVs to different charge points

Virtual power plants allow coordination of numerous smaller grid components, enabling them to participate in the market as one power plant



http://www.slideshare.net/statkraft/statkrafts-virtual-power-plant-in-germany

This makes VPPs visible to grid operators, allowing them to displace conventional *capacity*

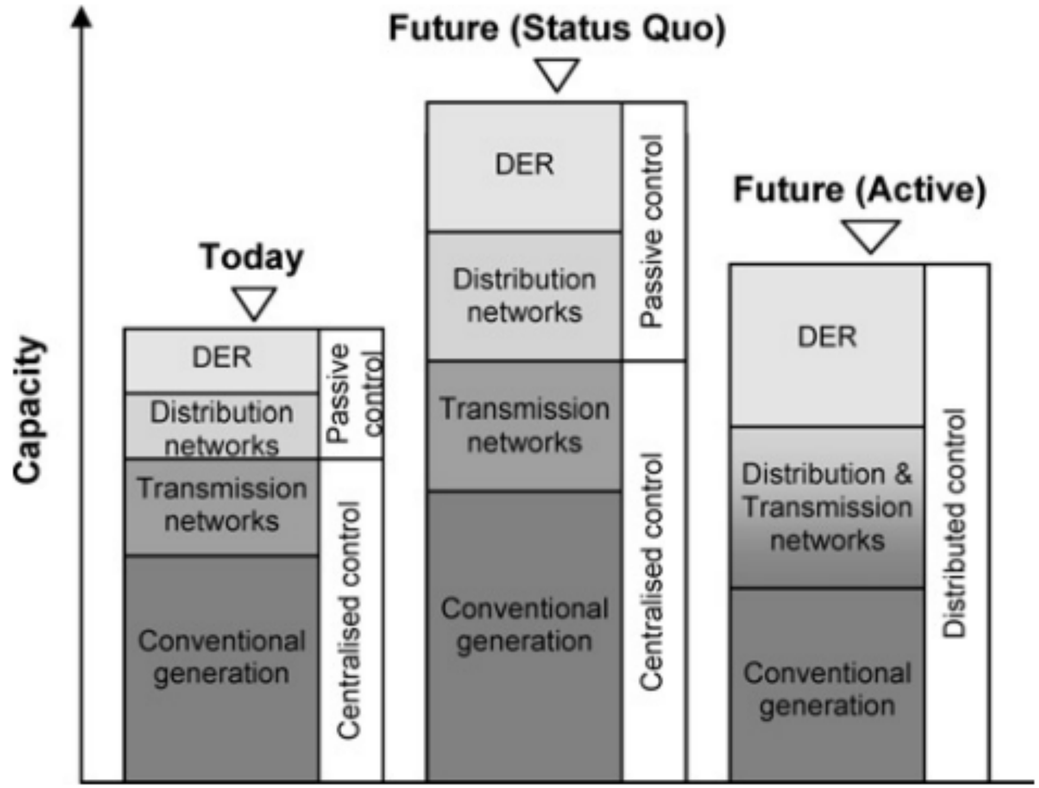


Fig. 1 *Relative levels of system capacity*

- Current distributed energy resources are not visible to system operators.
- Can displace energy produced by fossil fuels, but cannot displace capacity
- VPPs allow distributed generation to be represented on the system.
- Can trade in wholesale energy markets.
- Less capacity built → reduced costs!

Pudjianto et al., 2007

Software for VPPs is already in use, but several key challenges remain

https://www.youtube.com/watch?v=sNVrDwRQn3Q

Software challenges:

- Designing models of different VPP actors and processes in order to optimize technical arrangements needed to form and manage VPPs.
- Designing online mechanisms to form statistically correct trust measures for energy providers
- Designing search algorithms and negotiation mechanisms for individual actors to agree on which VPP to form at different points in time and how to share the profits

Data analytics can help people purchase/sell energy

“Decision Intelligence for the Wholesale Electricity Market”

https://www.youtube.com/watch?v=mH3qxQ8s4ZI

For buyers:

- Recommendation engine for energy customers who participate in deregulated electricity markets (PJM, NYISO, MISO, ERCOT)
- Build market price forecasts
- Load forecasting for industrial customers so they can reduce usage during peak hours

For sellers:

- Price forecasting for renewable generators to sell in day-ahead markets

Self-healing networks: envisioning an energy internet

Responding in real-time

Power and productivity for a better world™ ABB

ABB developed software that tracks grid flows and feeds information to control systems that can respond within a minute. They claim they can make outages 100 times less likely.

Predictive maintenance

C3 IoT

Real-time data and predictive algorithm allows grid operators to identify high-risk assets and perform maintenance before they fail. Helps operators prioritize maintenance.

Also see: Rudin et al., 2012: Machine Learning for the New York City Power Grid

Many startups offer customer engagement products

OPOWER

Customer engagement platform:

- Show customers personalized energy and billing insights
- Build satisfaction and loyalty
- Market products specific to the customer
- Drive actions with rewards

plotwatt®
Save your energy™

TENDRIL®

bidgely

EnergyHub®

ONZO®

Energy Elephant

FIRSTFUEL

News room > News releases >

Machine Learning Helps IBM Boost Accuracy of U.S. Department of Energy Solar Forecasts by up to 30 Percent

Makes Solar Forecasts Available to States to Advance Integration of Solar Power into the Nation's Energy Pipeline

https://www.youtube.com/watch?v=cj2RXjvRKOA

"Big data meets science"

Self-learning weather Model and renewable forecasting Technology (SMT)
Combine machine learning techniques with "domain models" ie. Meteorological models, atmospheric models

The smart grid in action: the Los Angeles Smart Grid Project

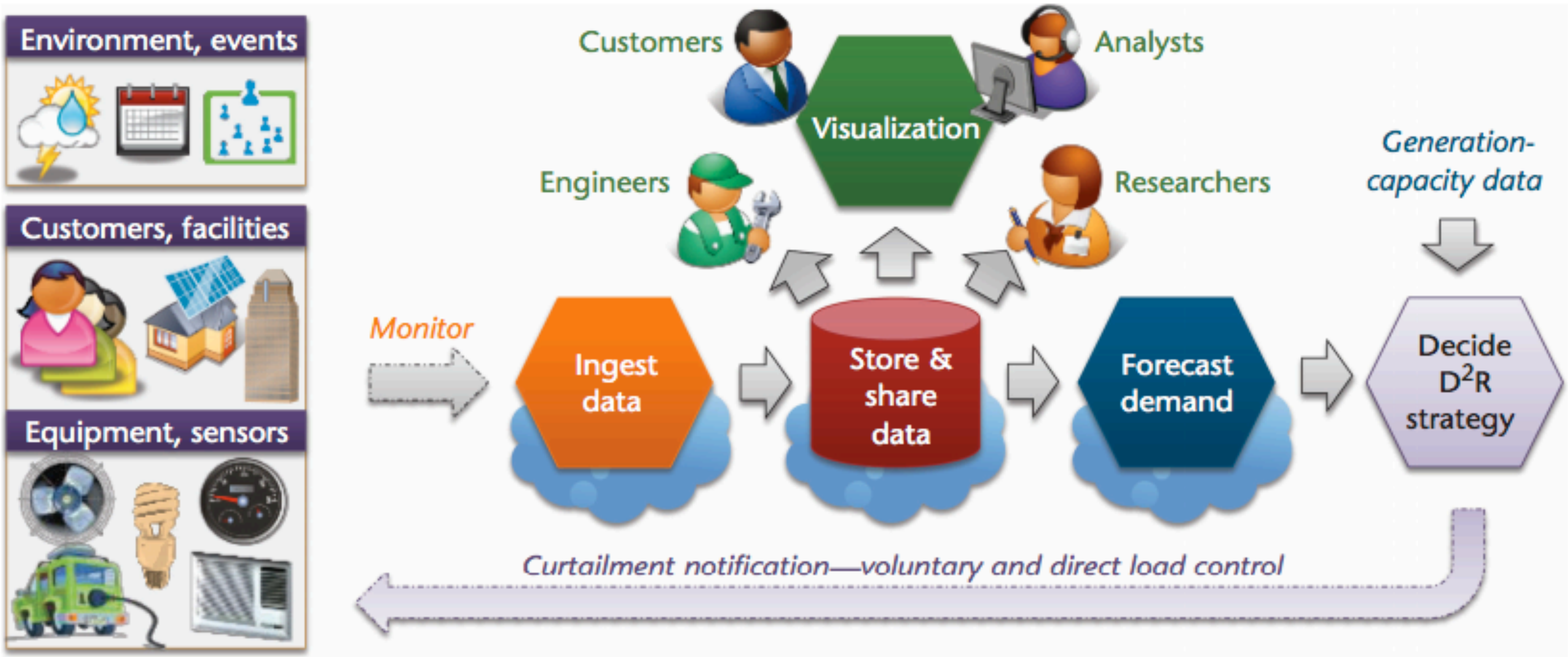


Figure 2. Dynamic Demand Response (D^2R) lifecycle in the University of Southern California (USC) campus microgrid using our cloud-based software platform. This lifecycle forms an observe, orient, decide, and act (OODA) loop.

Source: Simmhan et al., 2013

The smart grid in action: the Los Angeles Smart Grid Project

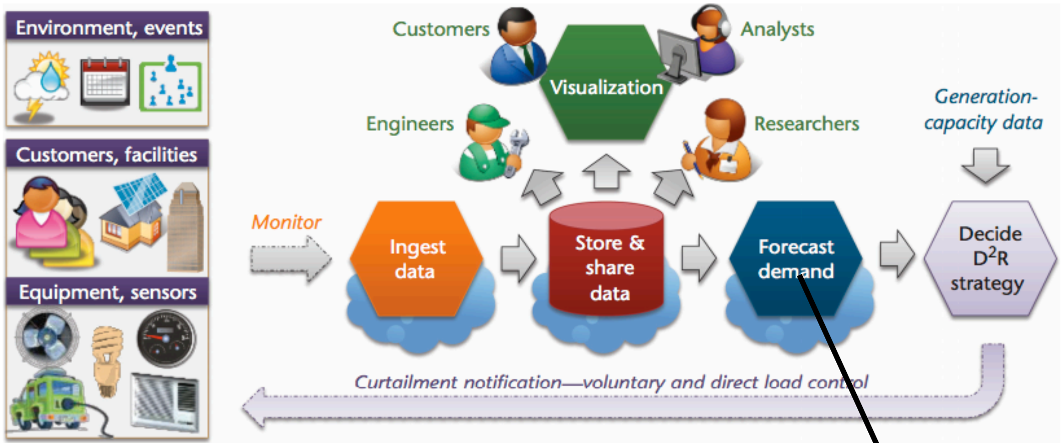
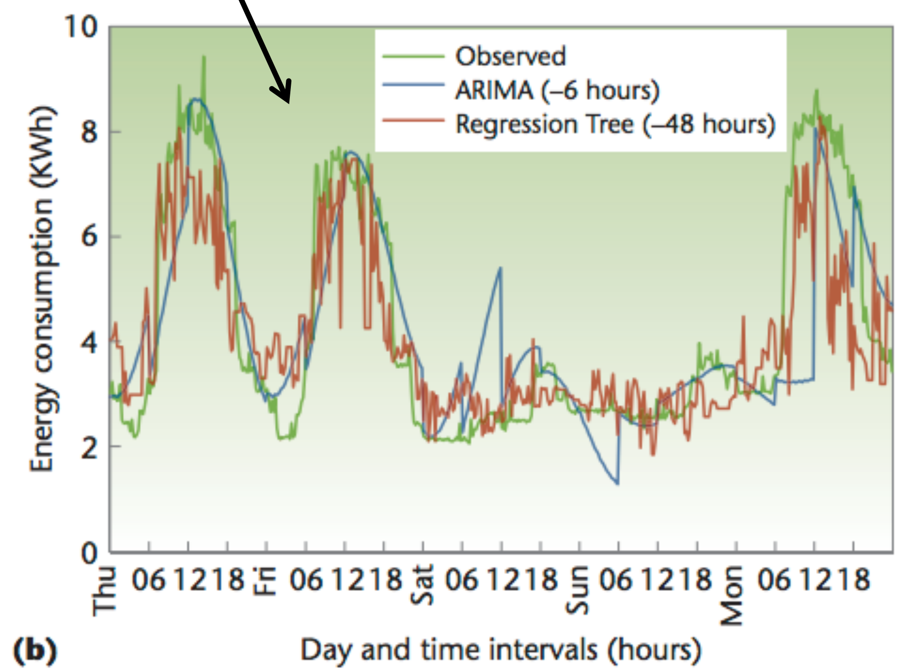


Figure 2. Dynamic Demand Response (D^{2}R) lifecycle in the Univer
microgrid using our cloud-based software platform. This lifecycle
(OODA) loop.

Simmhan et al., 2013



Computational requirements:

- 5 years of training data for 100 campus buildings has ~17 million feature vectors
- Training one model takes 7 hours on a computer

To scale up to an entire city is only possible with distributed computing!

The smart grid is not just an infrastructure challenge, but also a data analytics and software challenge