Breaking Down Analytical and Computational Barriers in Energy Data Analytics

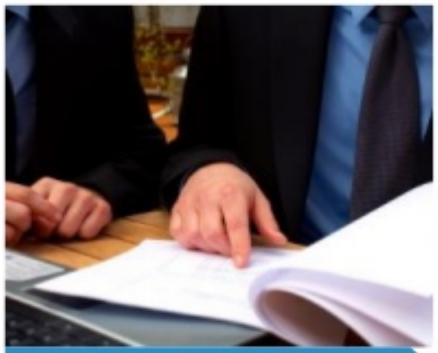
Jonathan Farland
DNV GL Energy

SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE
JUNE 6-8, 2016 SAN FRANCISCO

Agenda

Who is DNV GL?	• Introductions
Energy Analytics	• Overview • Data Science
Statistical Computing Pilot	• Demonstration
Concepts in Development	• Plans
Q&A	• Discussion

DNV GL



Policy

Production

Transmission & Distribution

Use

DNV·GL

Policy, Advisory and Research

Demand Side Management

Energy Analytics

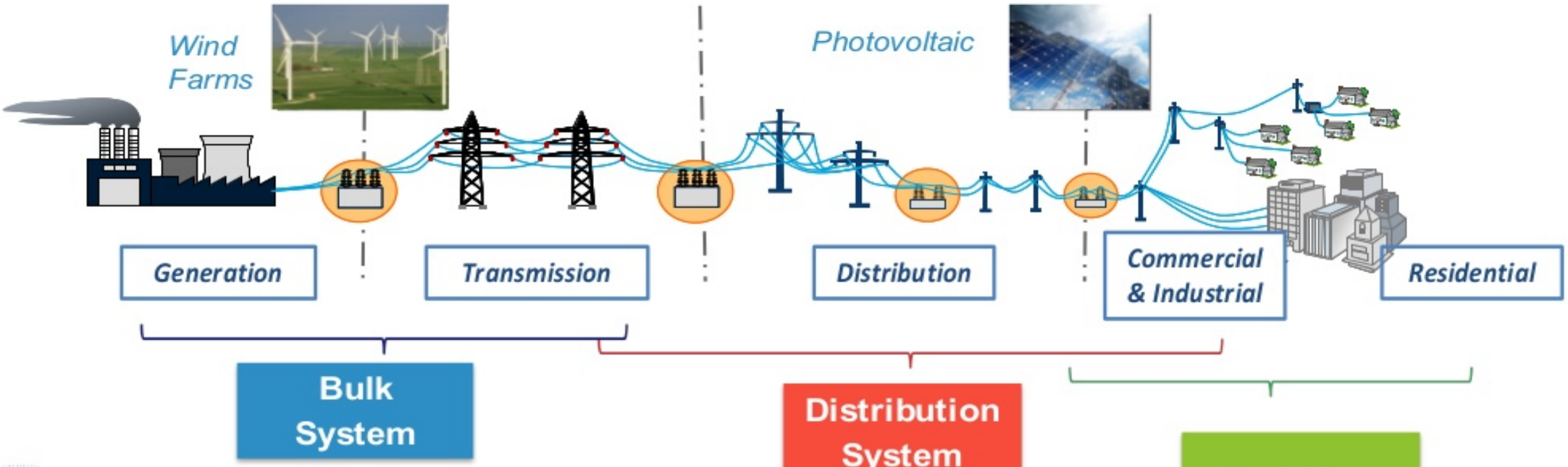
Load Research Services

Market Research and Program Evaluation

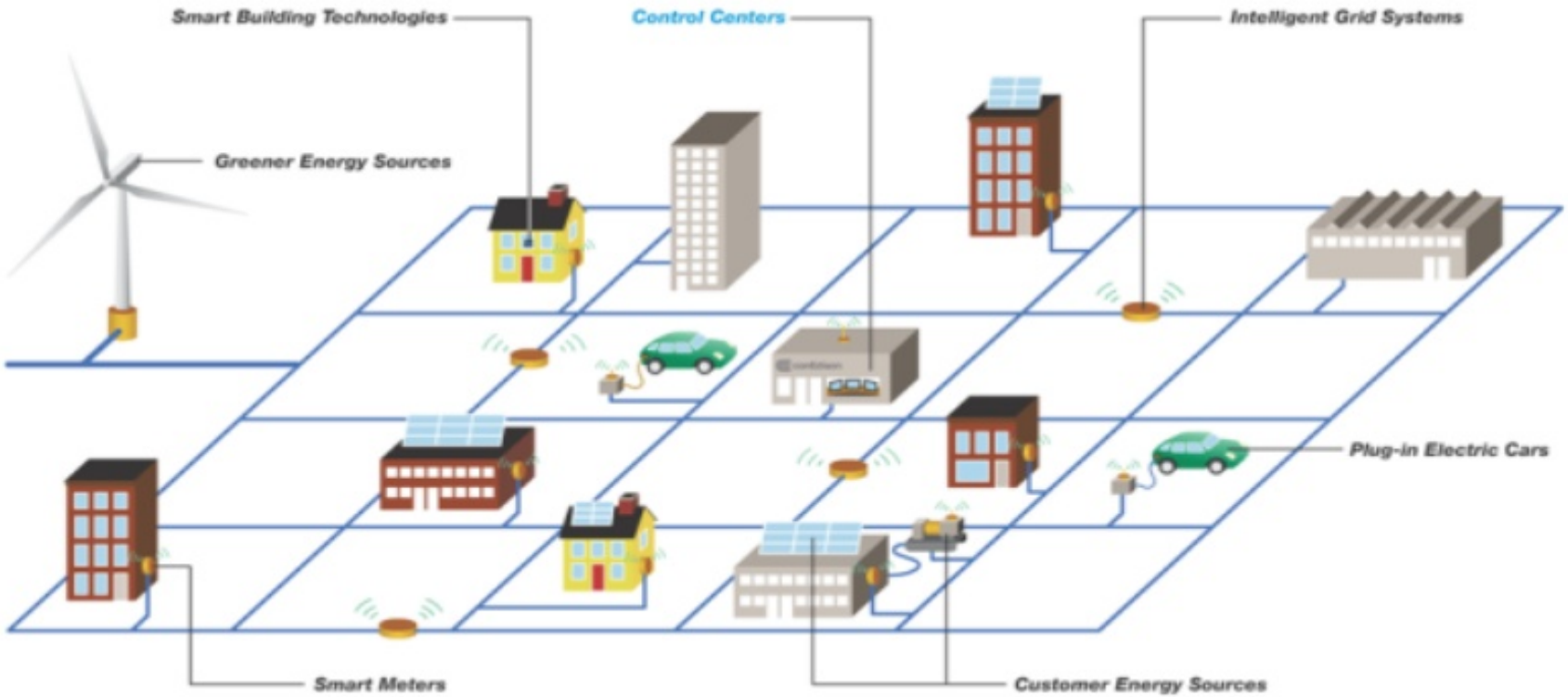
DNV·GL

Electricity Distribution Grid

Electricity Distribution Grid



The Rise of The Smart Grid



Energy Data Science

Jonathan Farland
DNV GL Energy

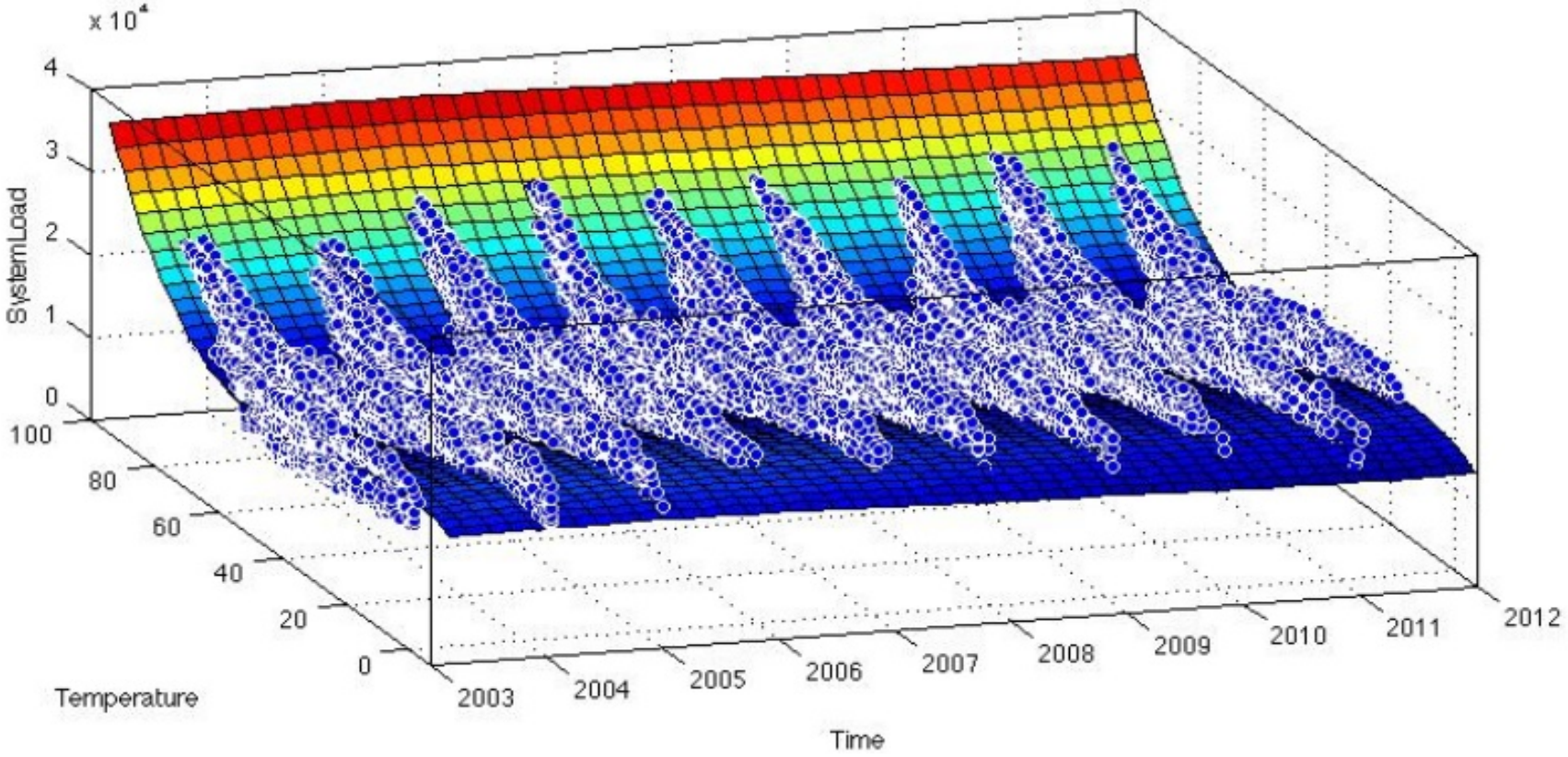
SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE
JUNE 6-8, 2016 SAN FRANCISCO

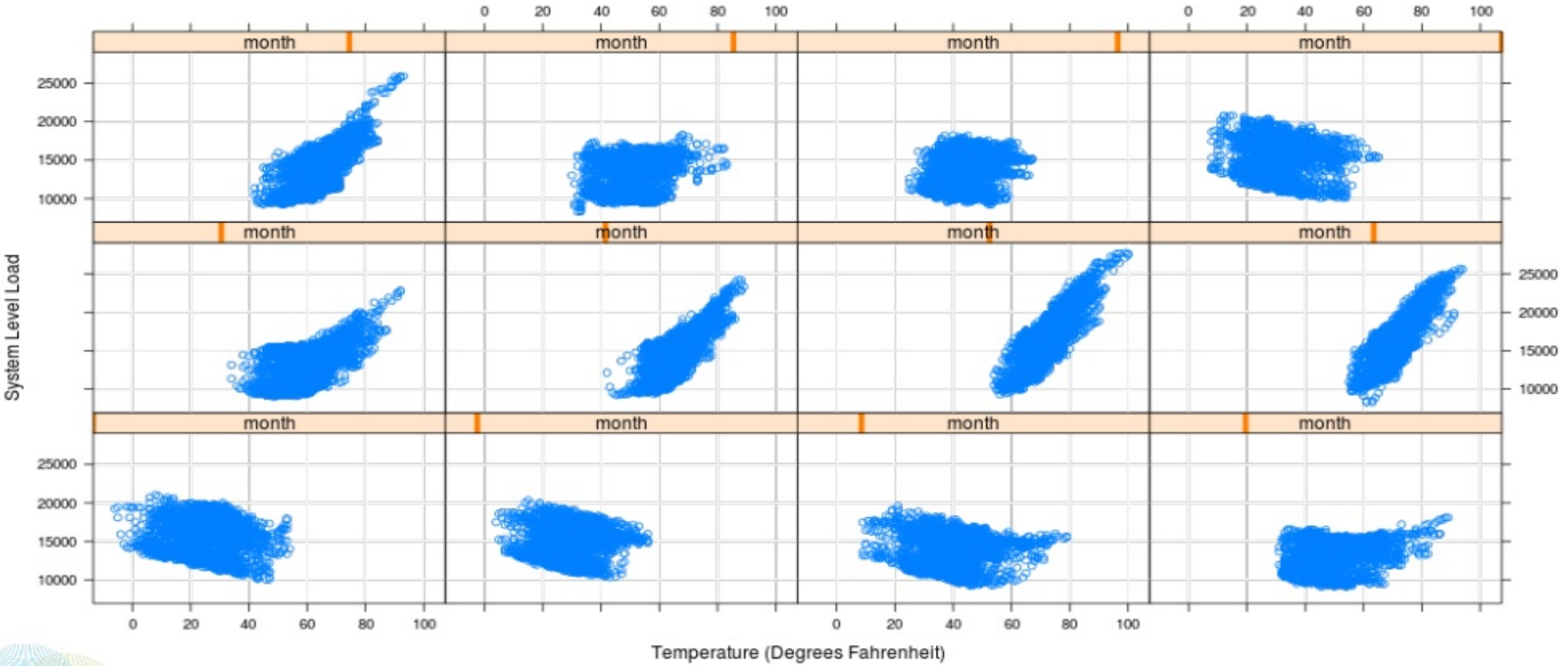
Terminology

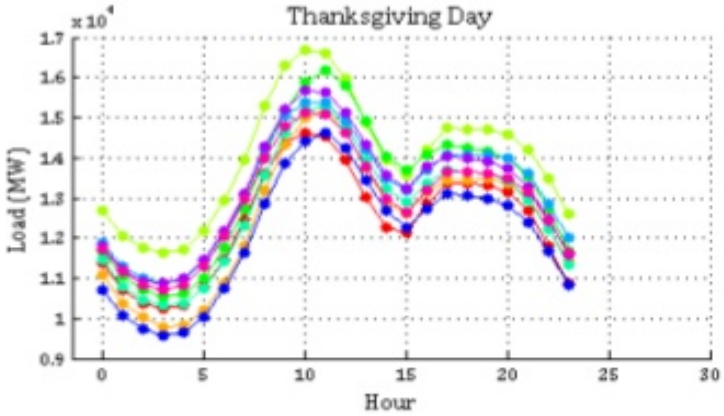
Energy (kWh, MWh, GWh):
Usage of energy over time

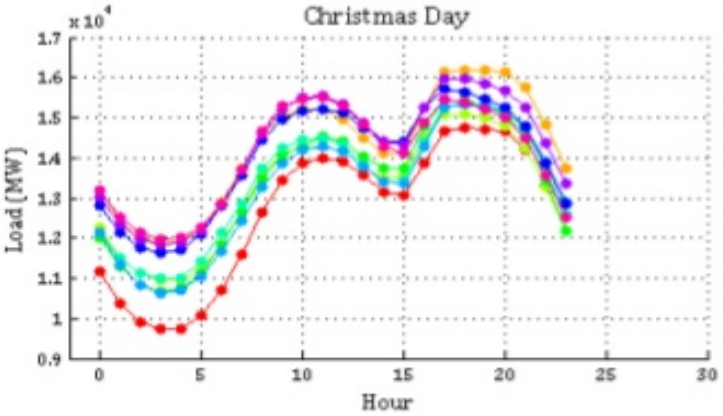
Demand (kW, MW, GW): Maximum power requirement of a system at a given time (e.g., an hour, a day, a month, a season of the year).

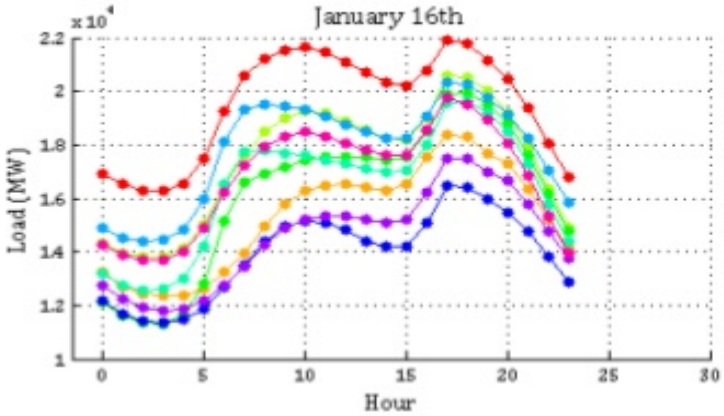
Load Forecast: The act of generating predictions for future demand or energy usage of an electrical grid.

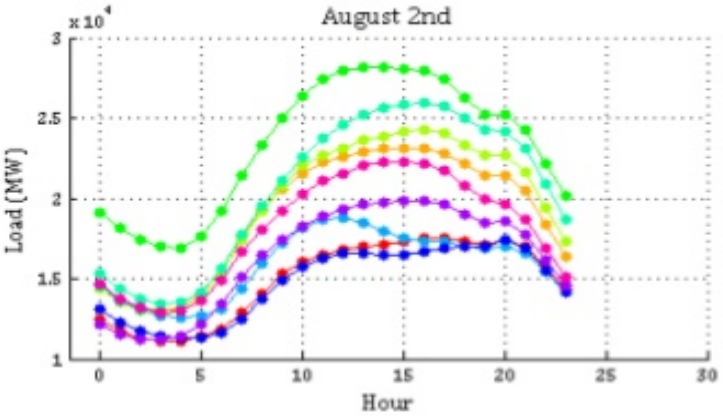

x 10^4
4
3
2
1
0
SystemLoad
100
80
60
40
20
0
Temperature
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Time

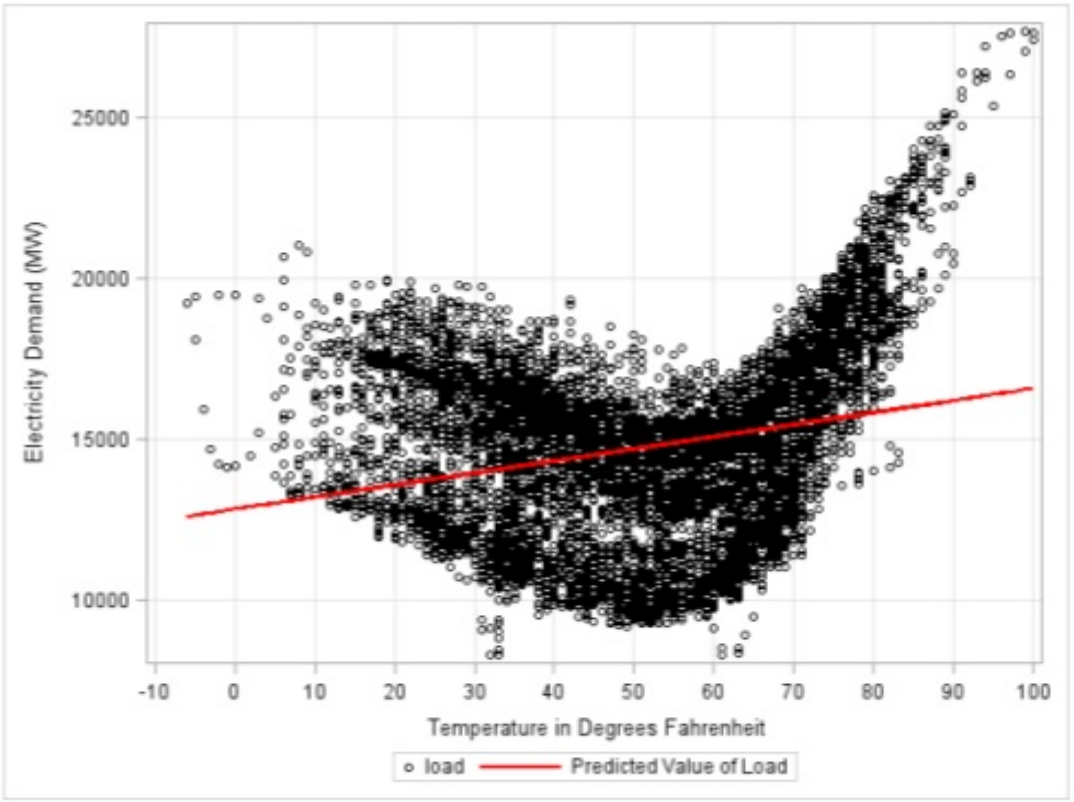

month
System Level Load
Temperature (Degrees Fahrenheit)
0
20
40
60
80
100
10000
15000
20000
25000

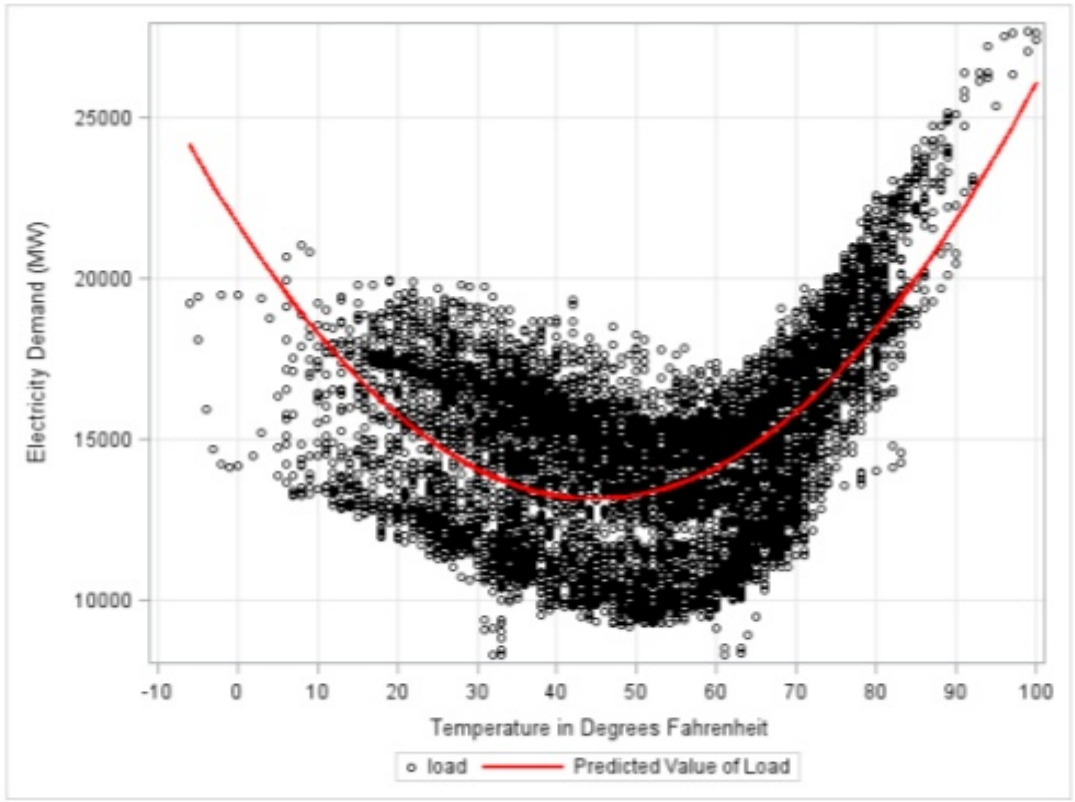

Thanksgiving Day
x 10^4
Load (MW)
Hour

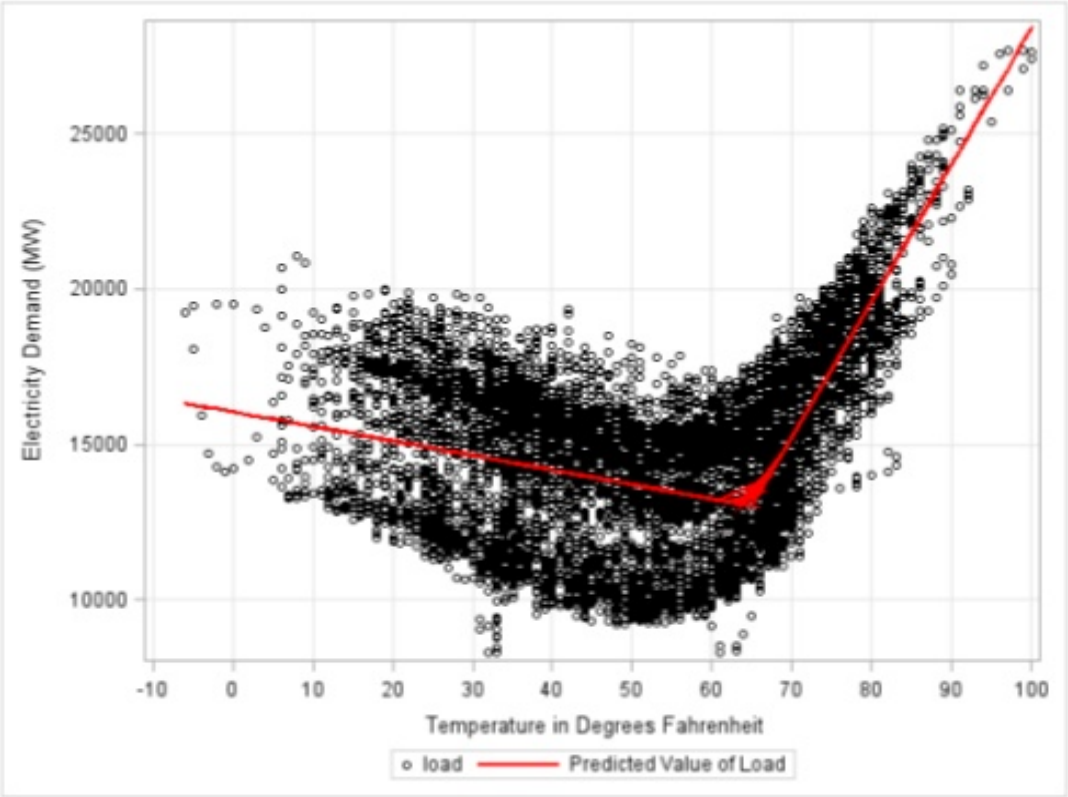

Christmas Day
x 10^4
Load (MW)
Hour
2003
2004
2005
2006
2007
2008
2009
2010
2011

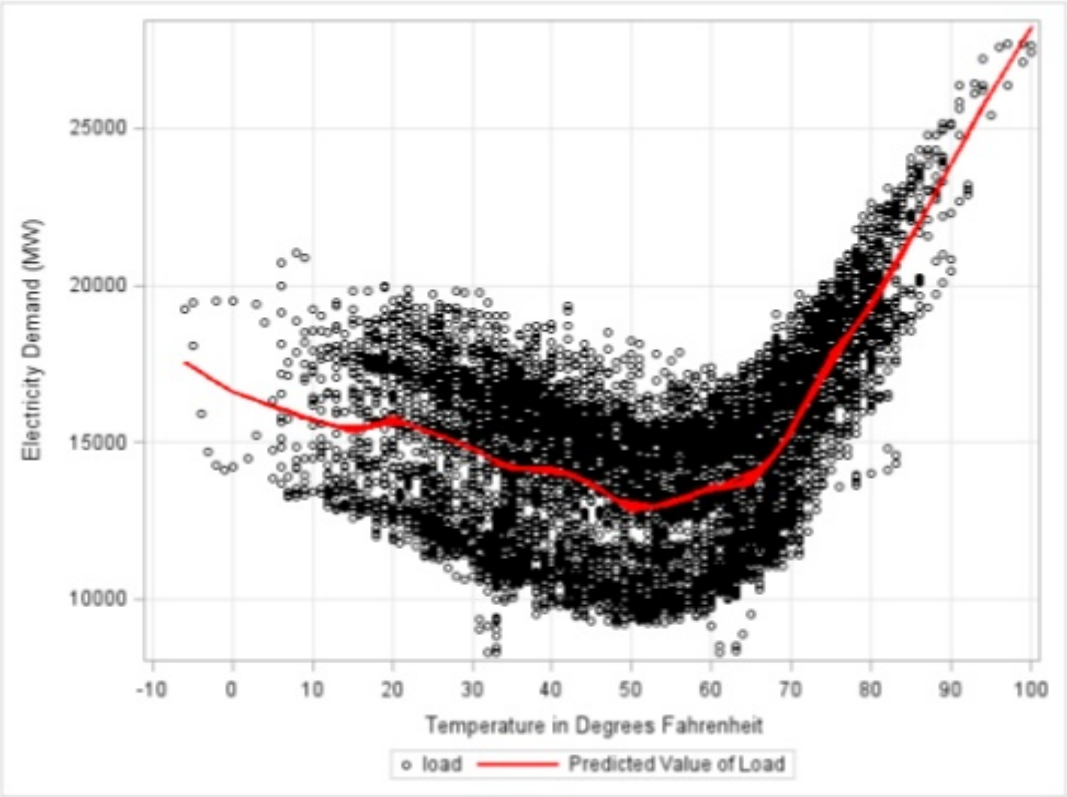

January 16th
x 10^4
Load (MW)
Hour

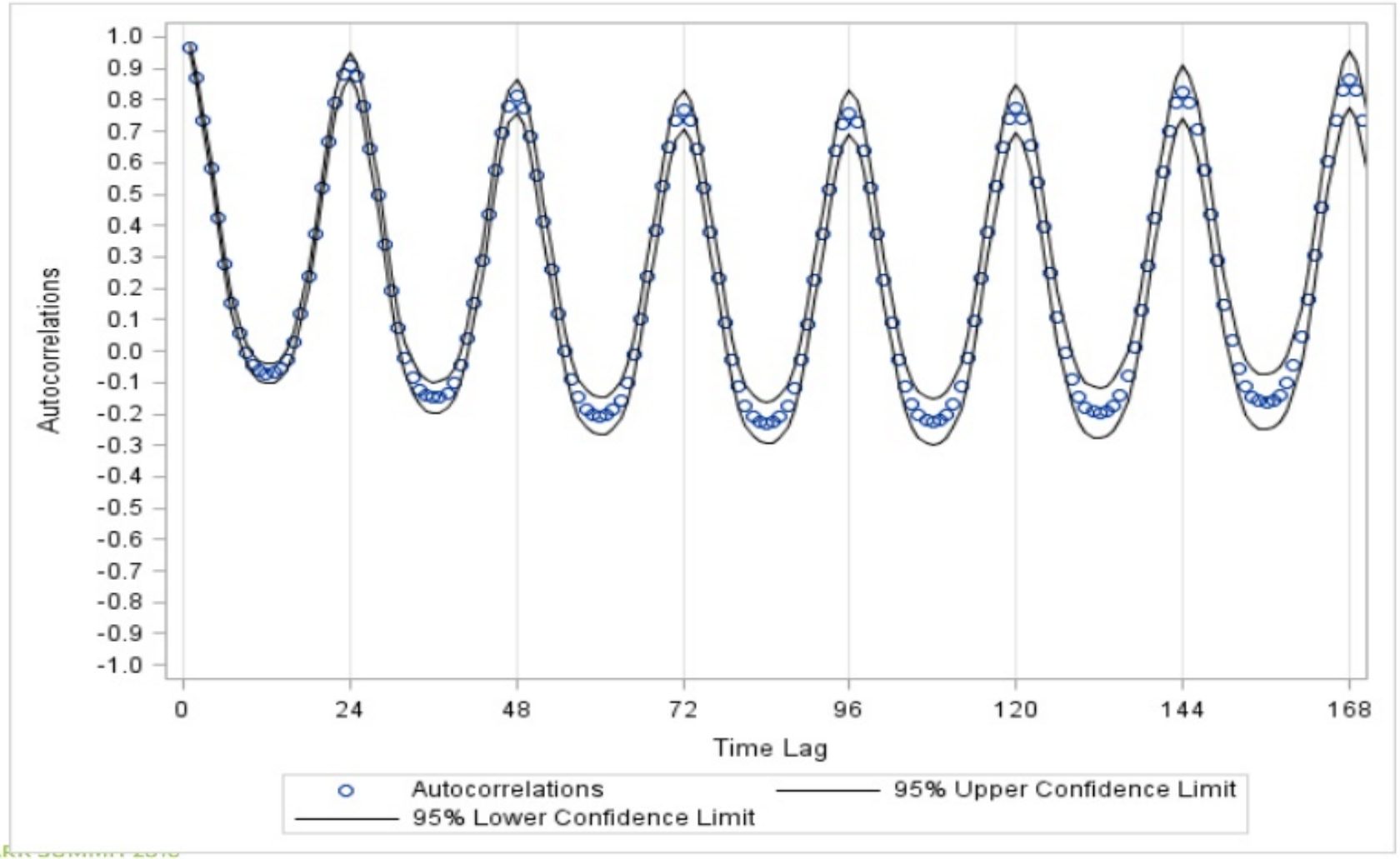

August 2nd
x 10^4
Load (MW)
Hour


Electricity Demand (MW)
25000
20000
15000
10000
-10
0
10
20
30
40
50
60
70
80
90
100
Temperature in Degrees Fahrenheit
load
Predicted Value of Load


Electricity Demand (MW)
25000
20000
15000
10000
-10
0
10
20
30
40
50
60
70
80
90
100
Temperature in Degrees Fahrenheit
load
Predicted Value of Load


Electricity Demand (MW)
25000
20000
15000
10000
-10
0
10
20
30
40
50
60
70
80
90
100
Temperature in Degrees Fahrenheit
load
Predicted Value of Load


Electricity Demand (MW)
25000
20000
15000
10000
-10
0
10
20
30
40
50
60
70
80
90
100
Temperature in Degrees Fahrenheit
load
Predicted Value of Load


Autocorrelations
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1.0
0
24
48
72
96
120
144
168
Time Lag
Autocorrelations
95% Upper Confidence Limit
95% Lower Confidence Limit

Forecasting Approaches

OLD SCHOOL

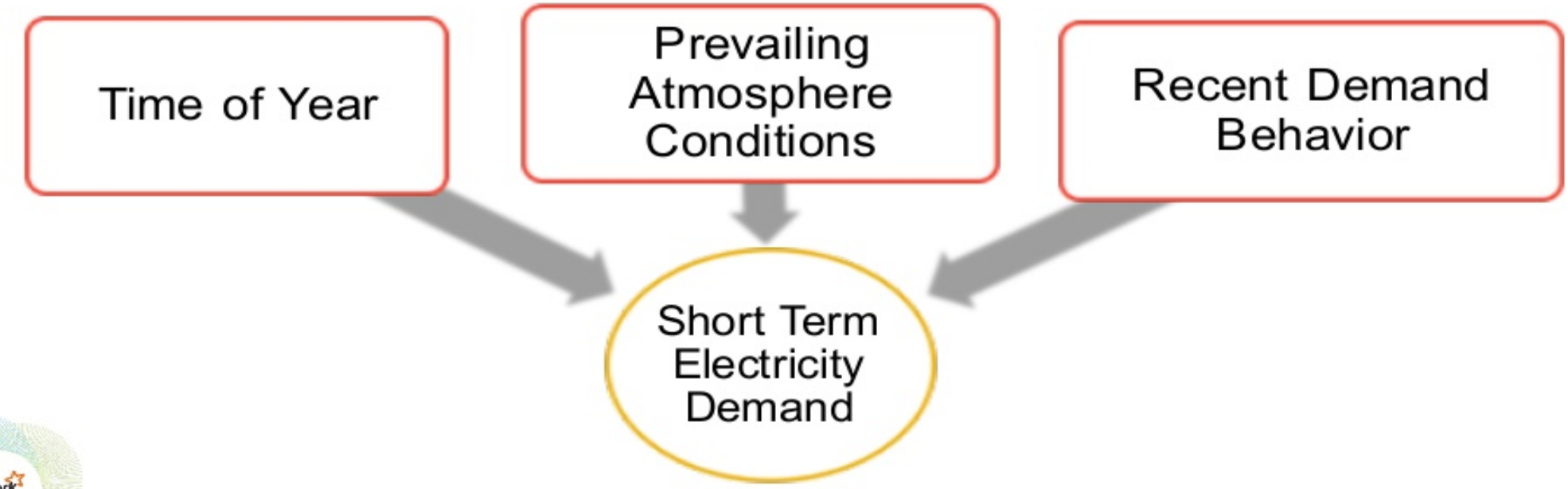
- Similar Day Matching
- Statistically Adjusted Engineering (SAE)
- Univariate Time Series (ARIMA)
- Multiple Linear Regression
- Econometric

NEW HOTNESS

- Machine / Statistical Learning
- Semiparametric Regression
- Artificial Neural Networks
- Fuzzy Logic
- Support Vector Machines
- Gradient Boosting

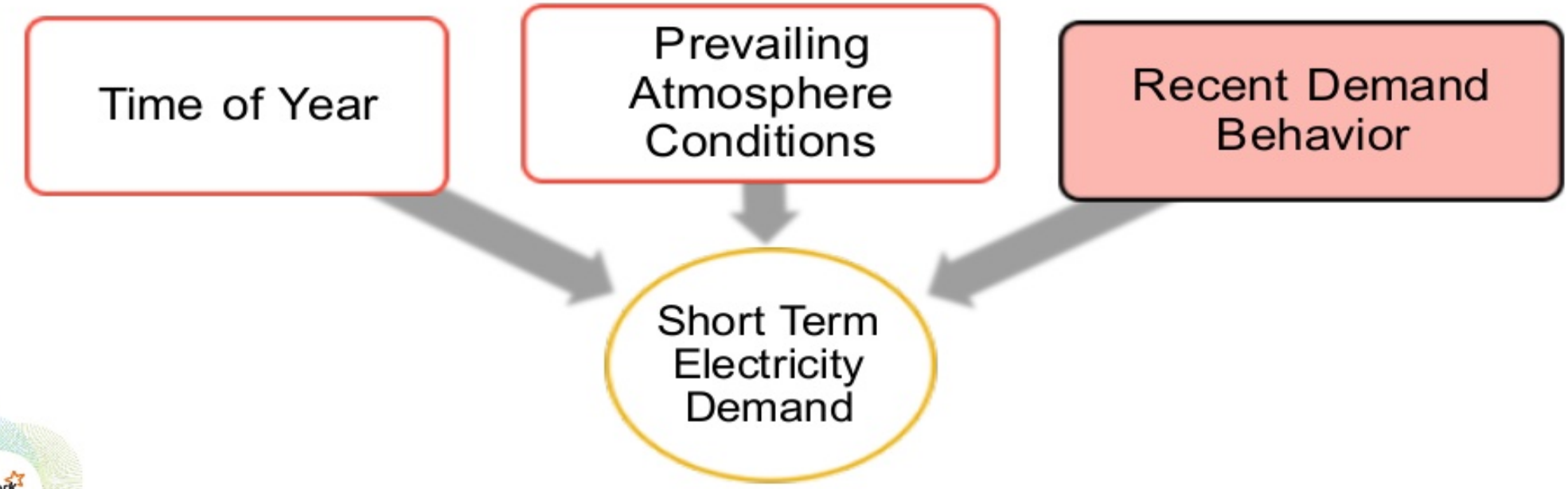
Additive Semiparametric Model

$$y_t = h(time) + f(weather) + \alpha(behavior) + \varepsilon_t$$



Additive Semiparametric Model

$$y_t = h(time) + f(weather) + \alpha(behavior) + \varepsilon_t$$



Emerging Technologies

Photovoltaic Cells (e.g., Solar)

Electric Vehicles

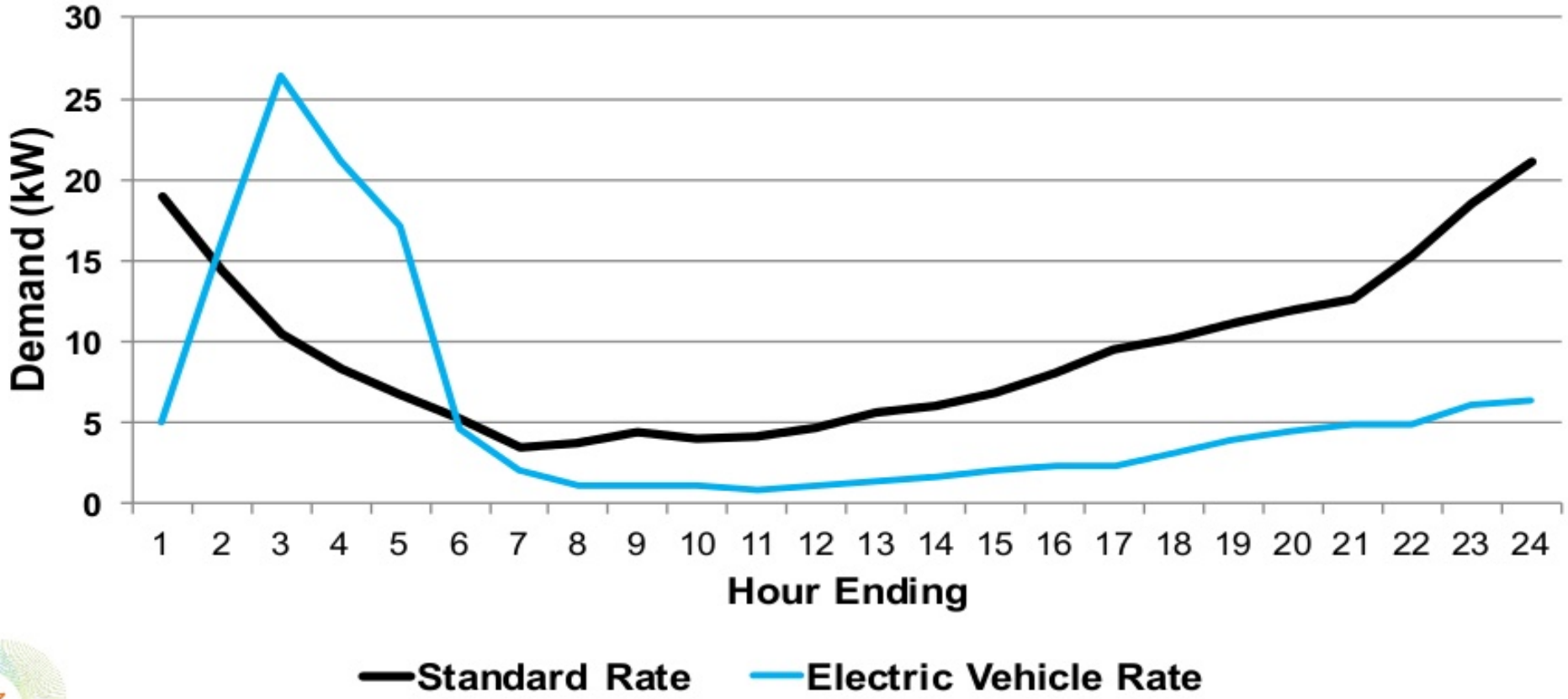
Storage

Wind

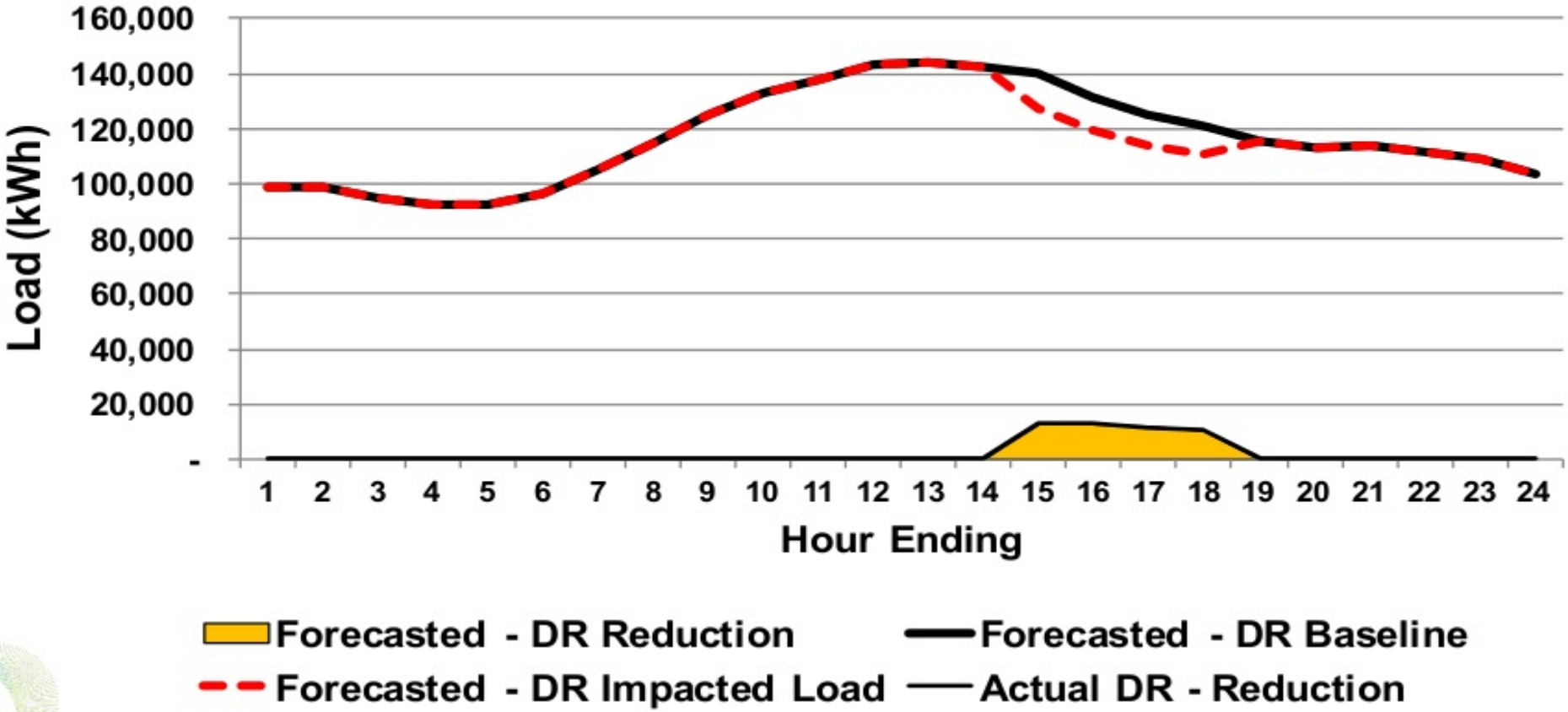
Energy Efficiency

Demand Response

Load Shifting: Electric Vehicles



Load Reduction: Demand Response



Databricks + Spark Pilot

SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE
JUNE 6-8, 2016 SAN FRANCISCO

Statistical Computing Pilot

Benefits of Big Data from Advanced Metering Infrastructure

- ✓ A deeper understanding of demand and therefore human behavior (think energy efficiency)
- ✓ Cost effective operating costs
- ✓ Real-time notification of power outages
- ✓ Improved System Planning and Reliability
- ✓ Allows for integration of disruptive technologies like Electric Vehicles

Statistical Computing Pilot

Statistical Computing Pilot

Data Diversity






Energy Consumption	Climatic	Demographic Firmographic	Economic Financial	Energy Efficiency Program Tracking	Grid Infrastructure
	- Temperature - Humidity - Wind Speed - Solar				

Key Focus Areas

Performance

Scalability

Granularity

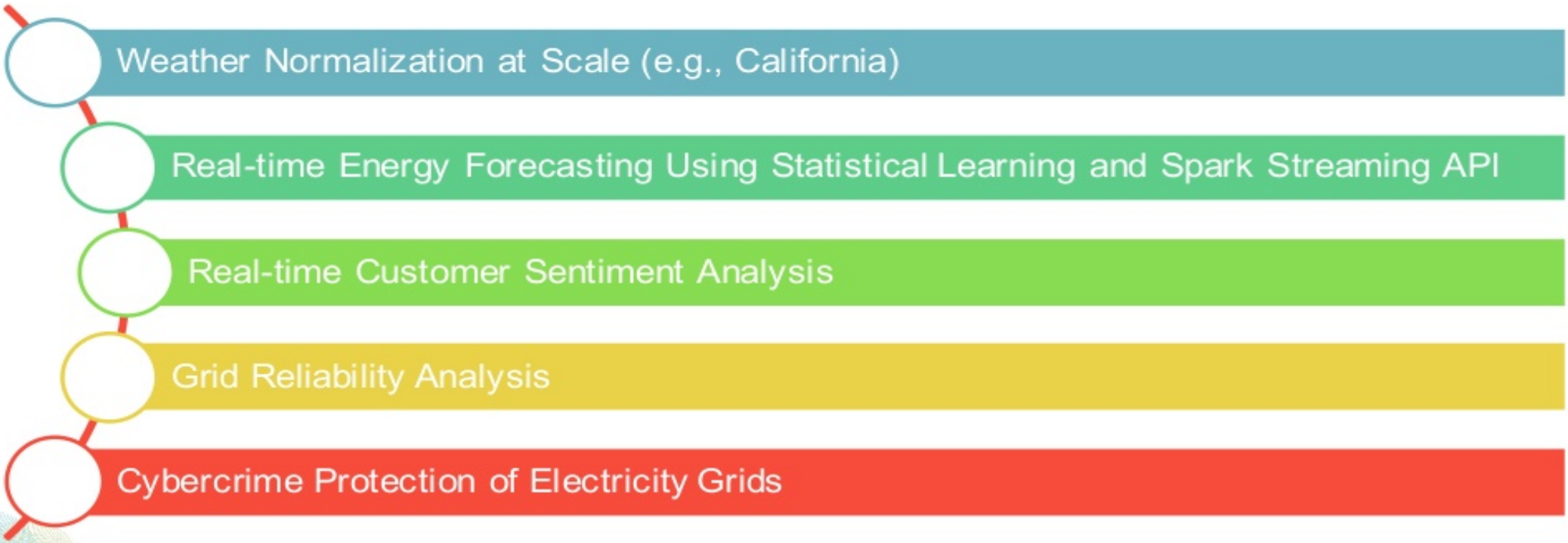
Going Further

Use Cases

DEMONSTRATION

VISION OF THE FUTURE

Current Concepts in Development



THANK YOU.

Jonathan Farland – DNV GL
jon.farland@dnvgl.com

Andrew Stryker – DNV GL
Andrew.stryker@dnvgl.com

SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE
JUNE 6-8, 2016 SAN FRANCISCO