

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



SPARK SUMMIT 2016

DNV GL



SPARK SUMMIT 2016



DNV·GL



SPARK SUMMIT 2016



Policy, Advisory and Research

Demand Side
Management

Energy
Analytics

Load Research
Services

Market Research
and Program
Evaluation



SPARK SUMMIT 2016

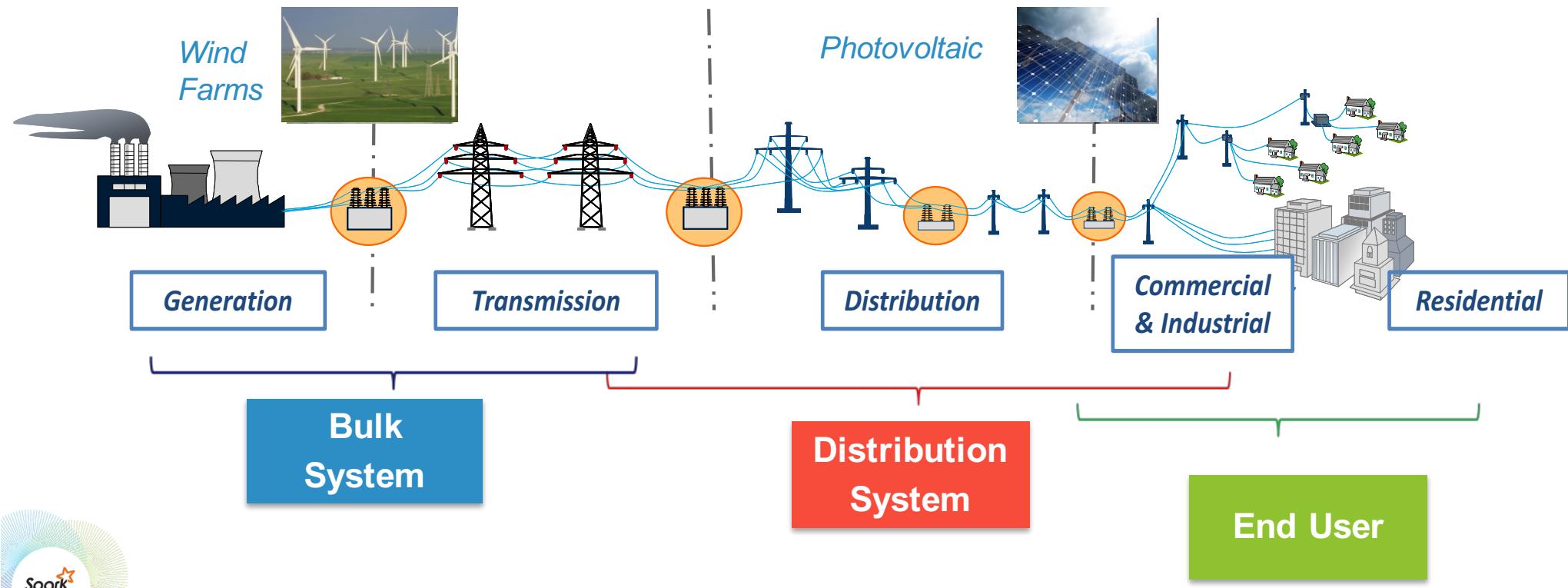


Electricity Distribution Grid



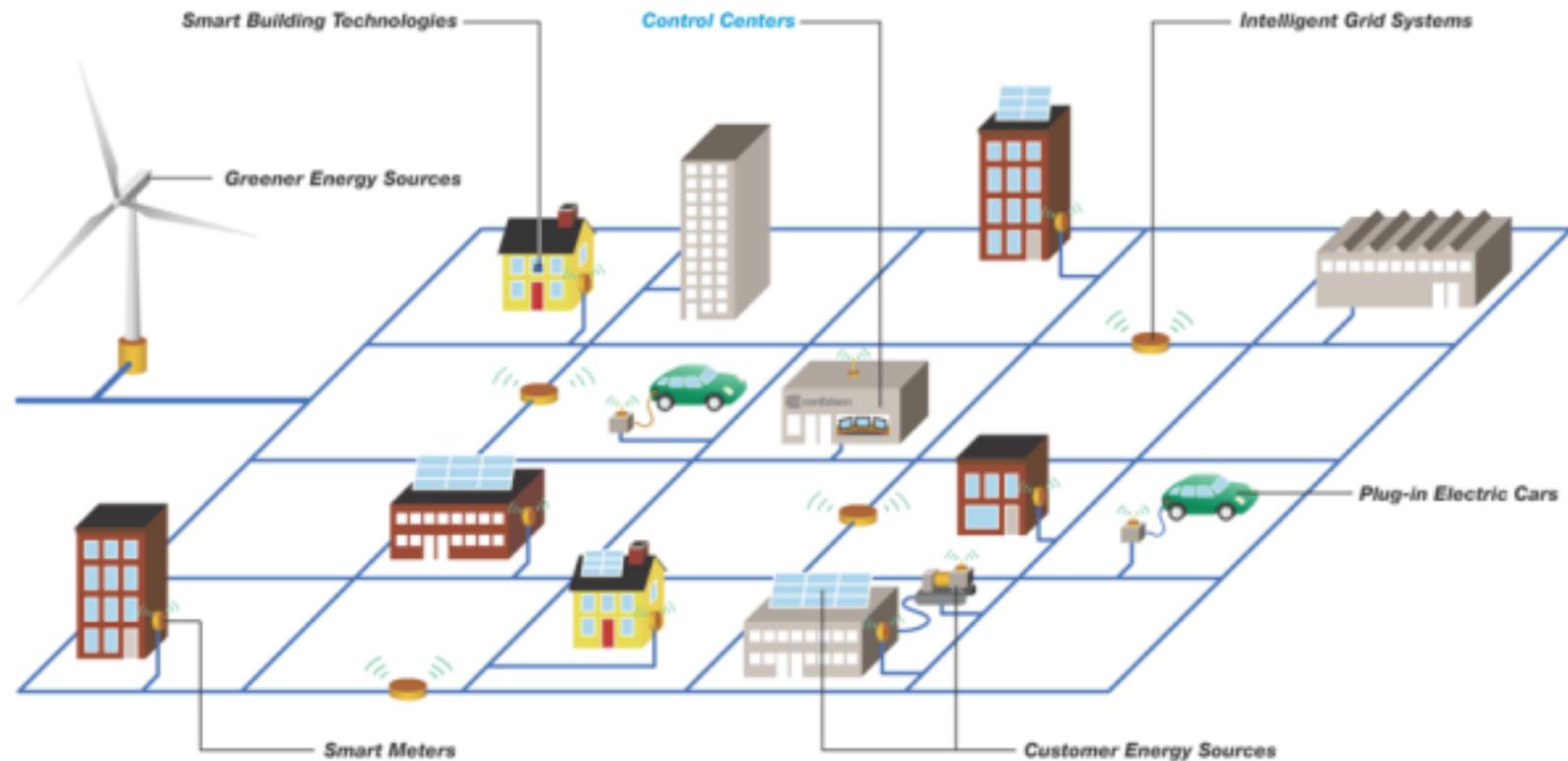
SPARK SUMMIT 2016

Electricity Distribution Grid



SPARK SUMMIT 2016

The Rise of The Smart Grid



SPARK SUMMIT 2016

Energy Data Science

Jonathan Farland
DNV GL Energy



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

Terminology



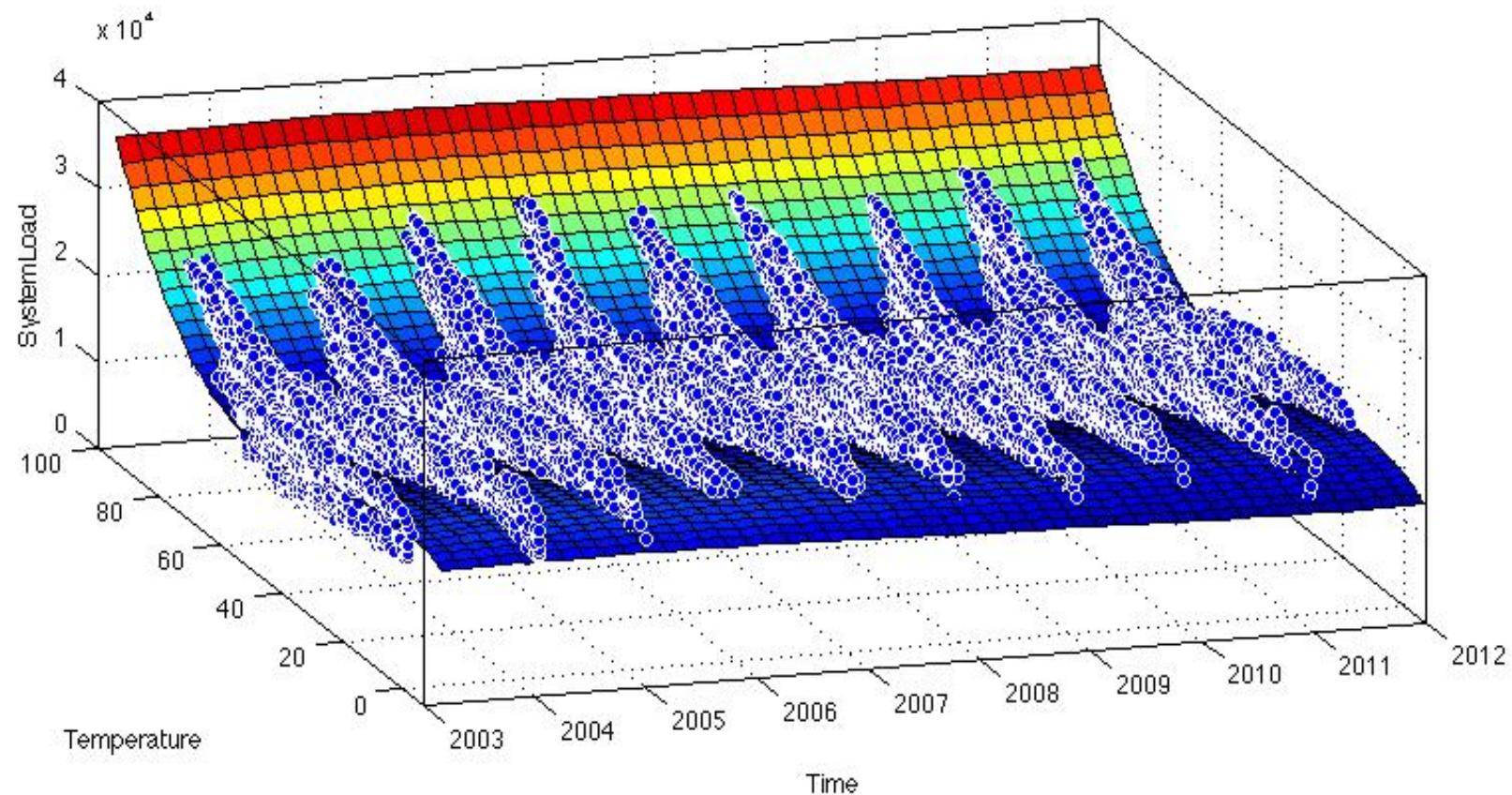
Energy (kWh, MWh, GWh):
Usage of energy over time

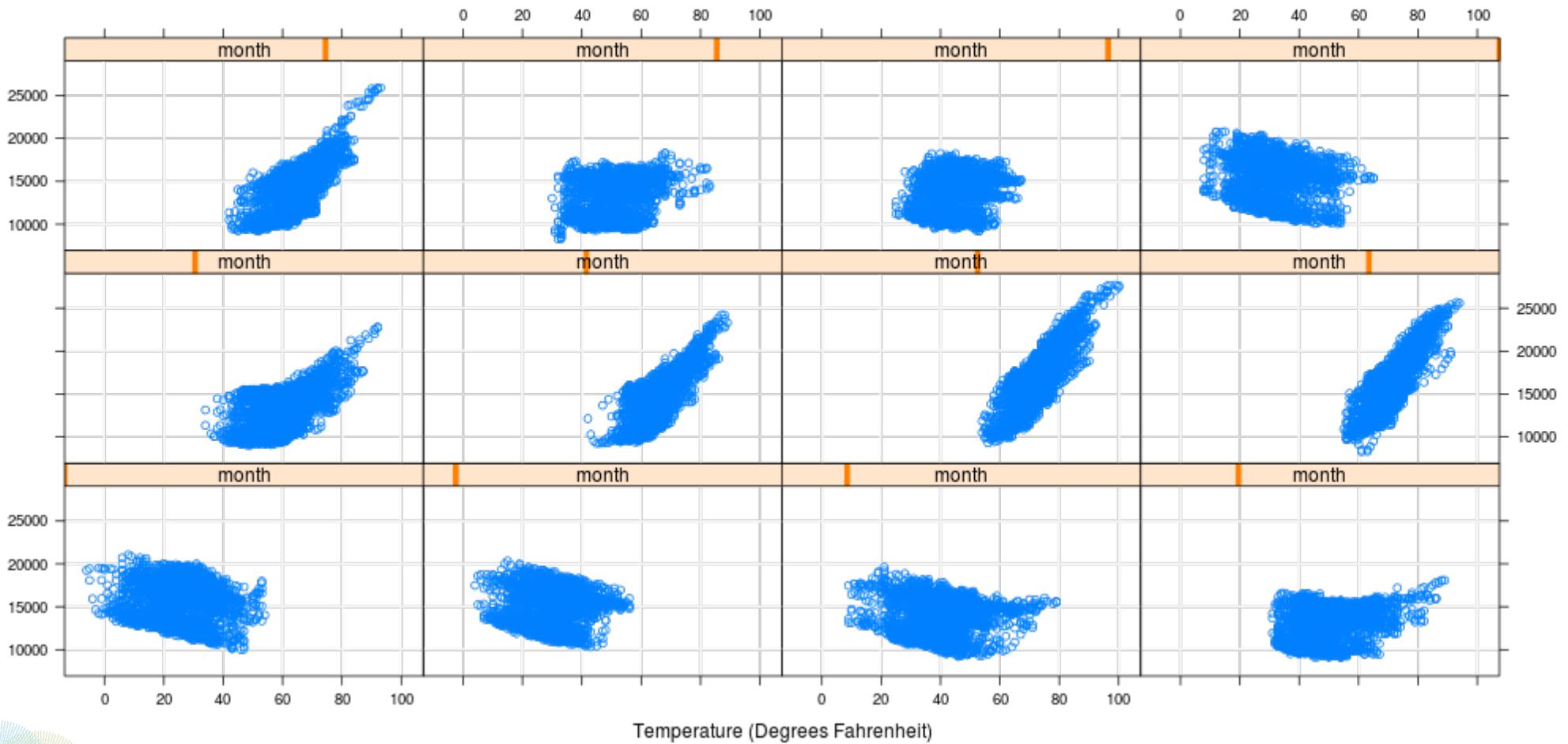


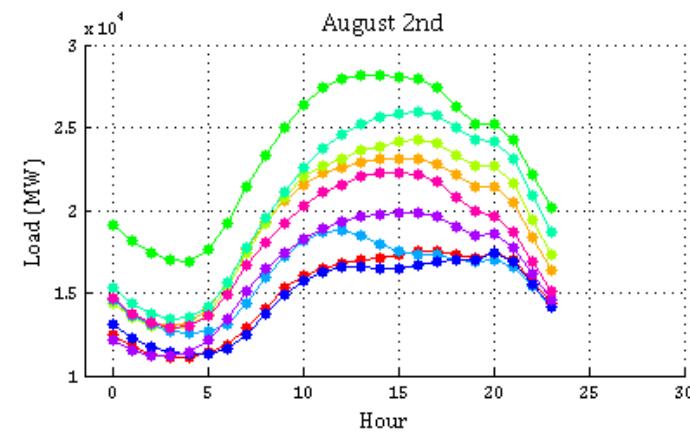
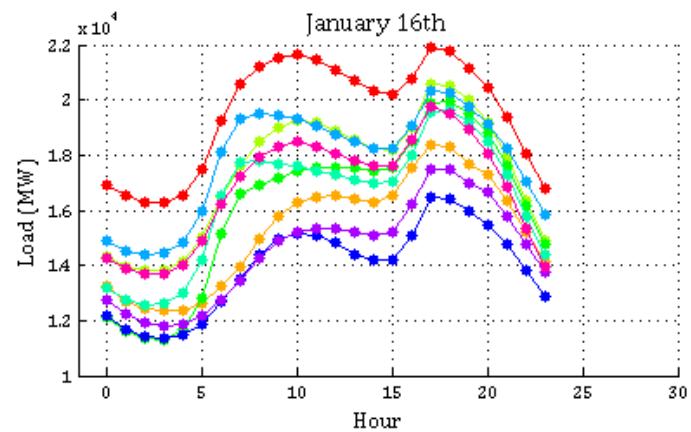
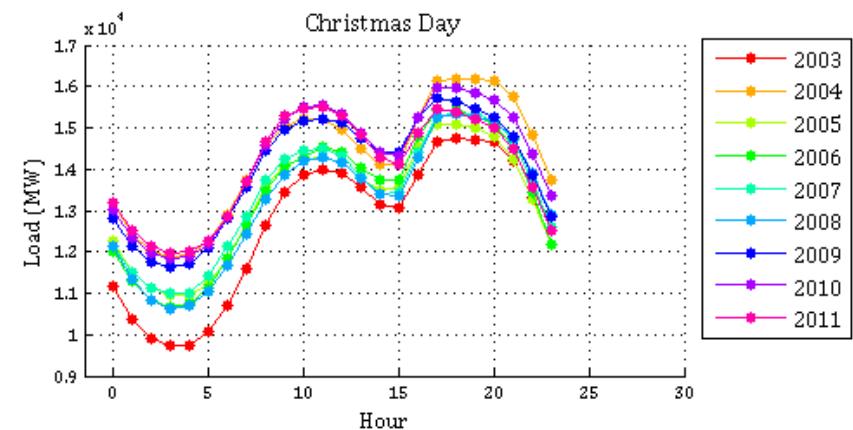
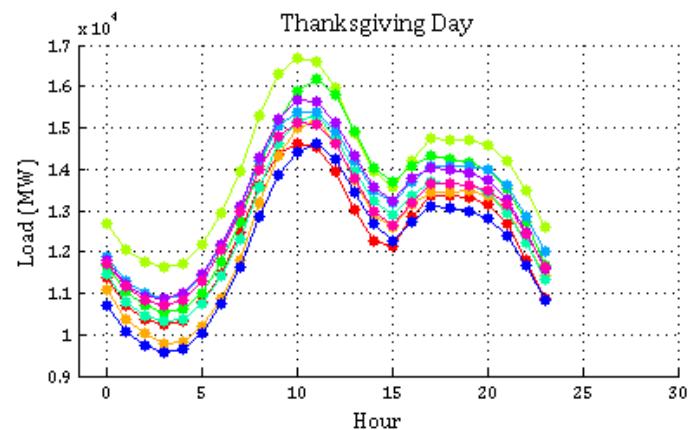
Demand (kW, MW, GW): Maximum power requirement of a system at a given time (e.g., an hour, a day, a month, a season of the year).

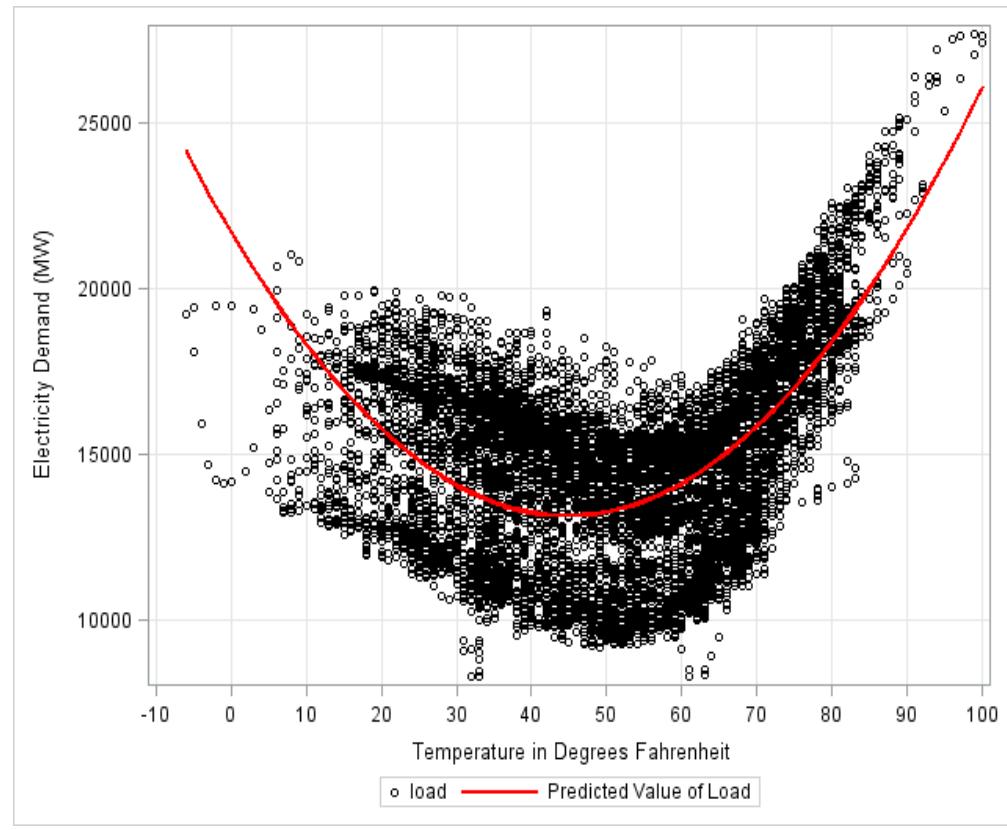
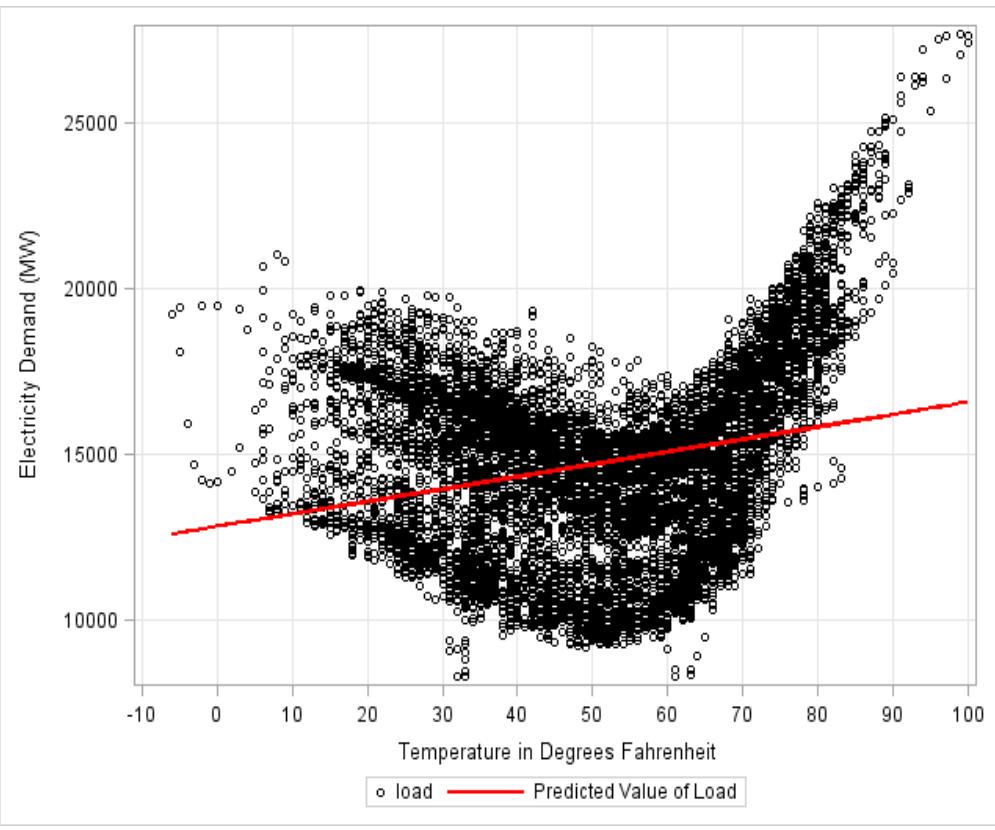


Load Forecast: The act of generating predictions for future demand or energy usage of an electrical grid.

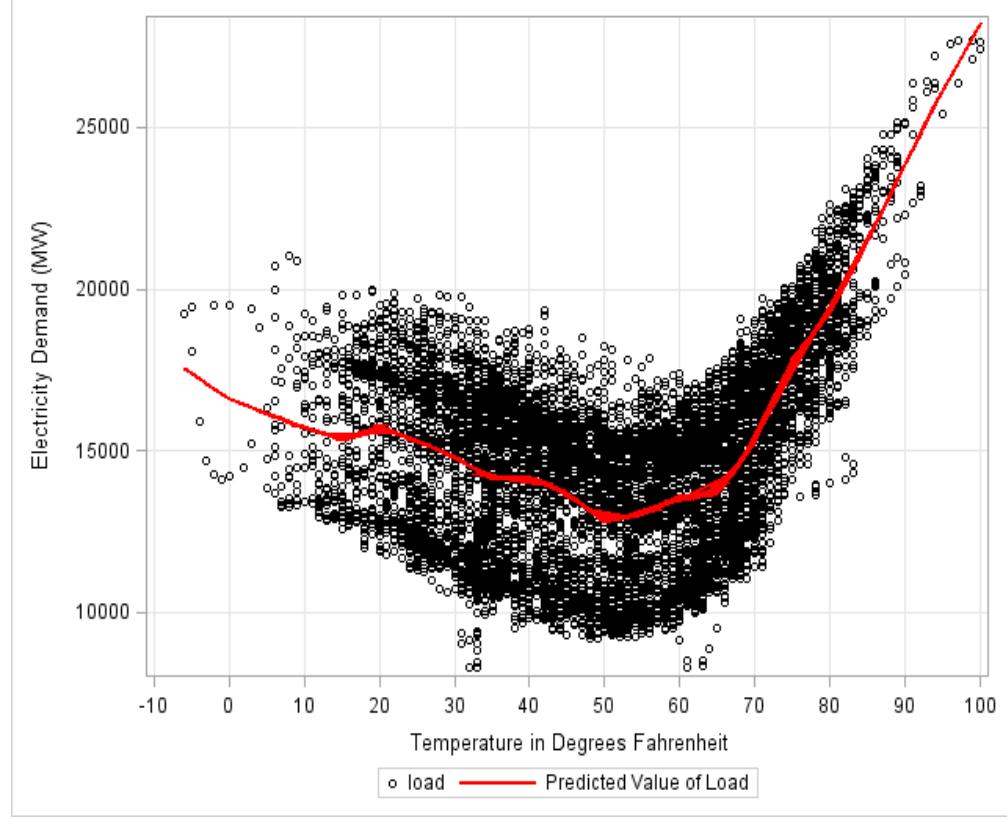
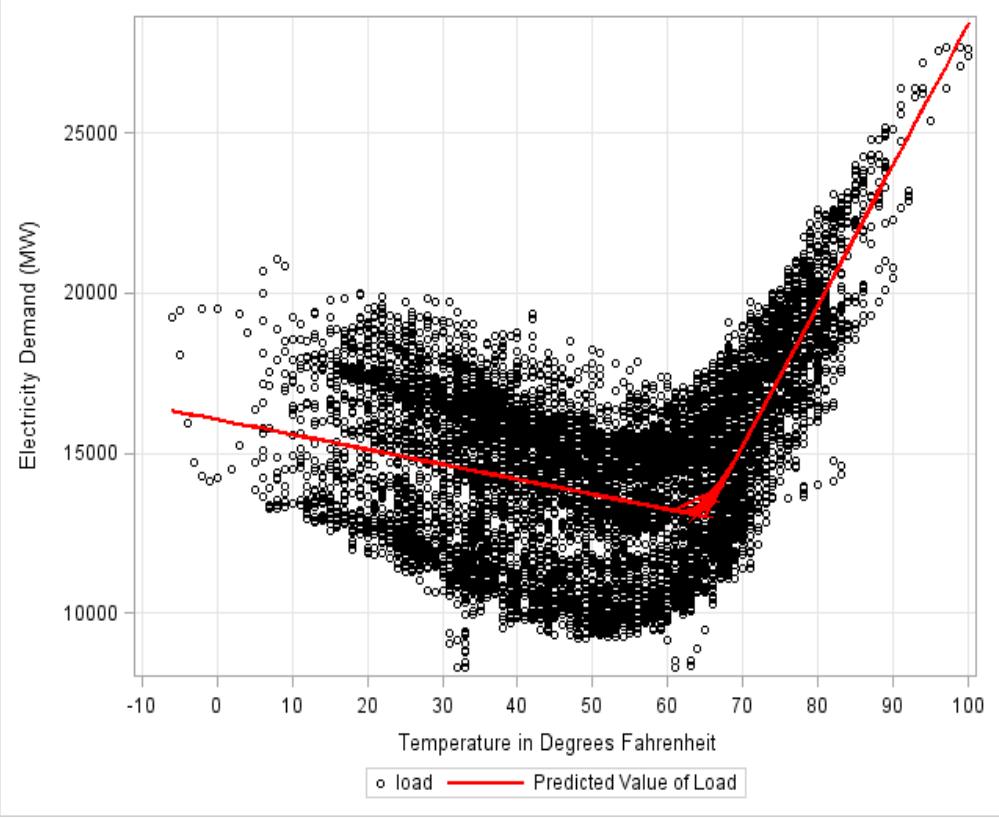




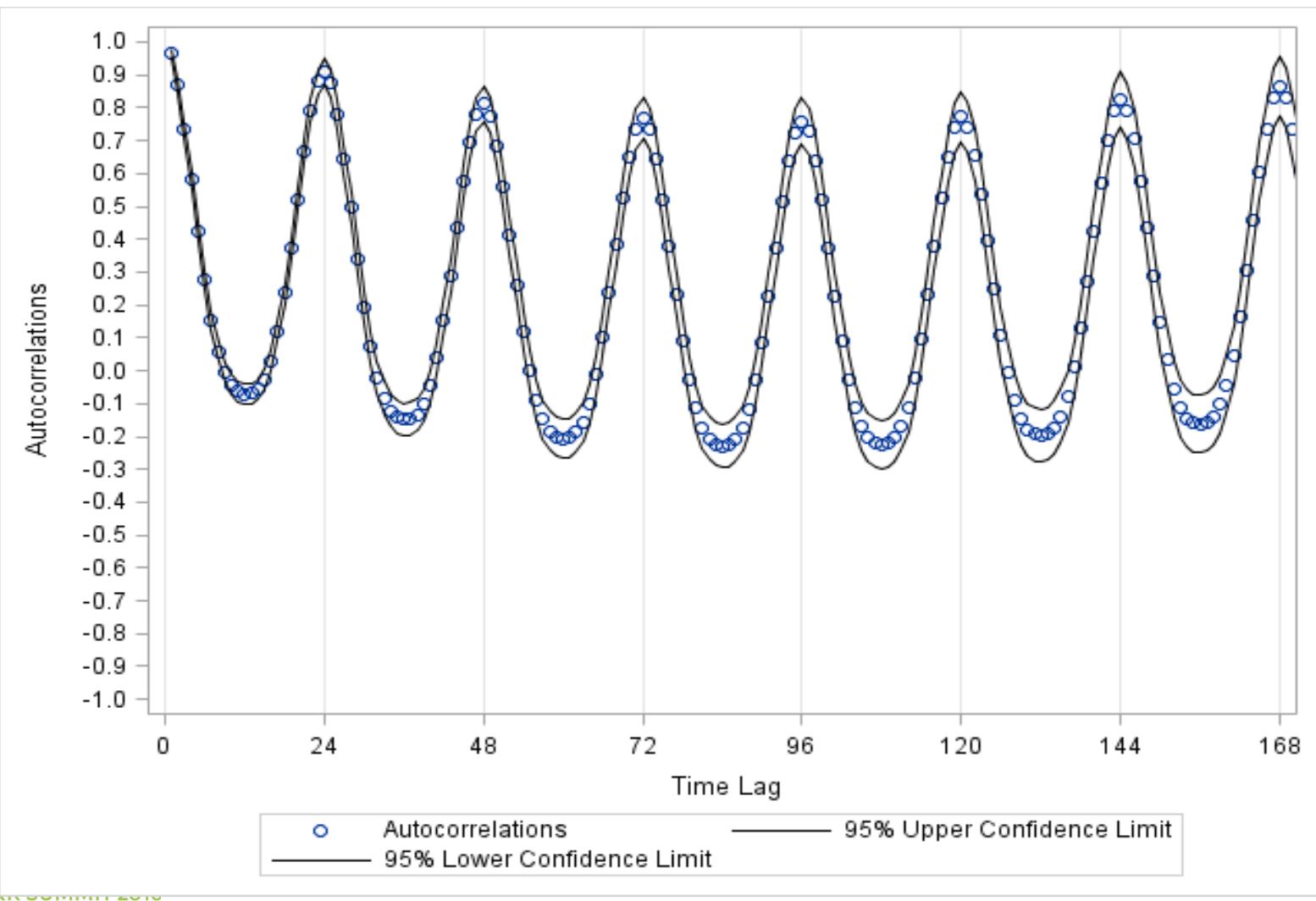




SPARK SUMMIT 2016



SPARK SUMMIT 2016



Forecasting Approaches

OLD SCHOOL

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

NEW HOTNESS

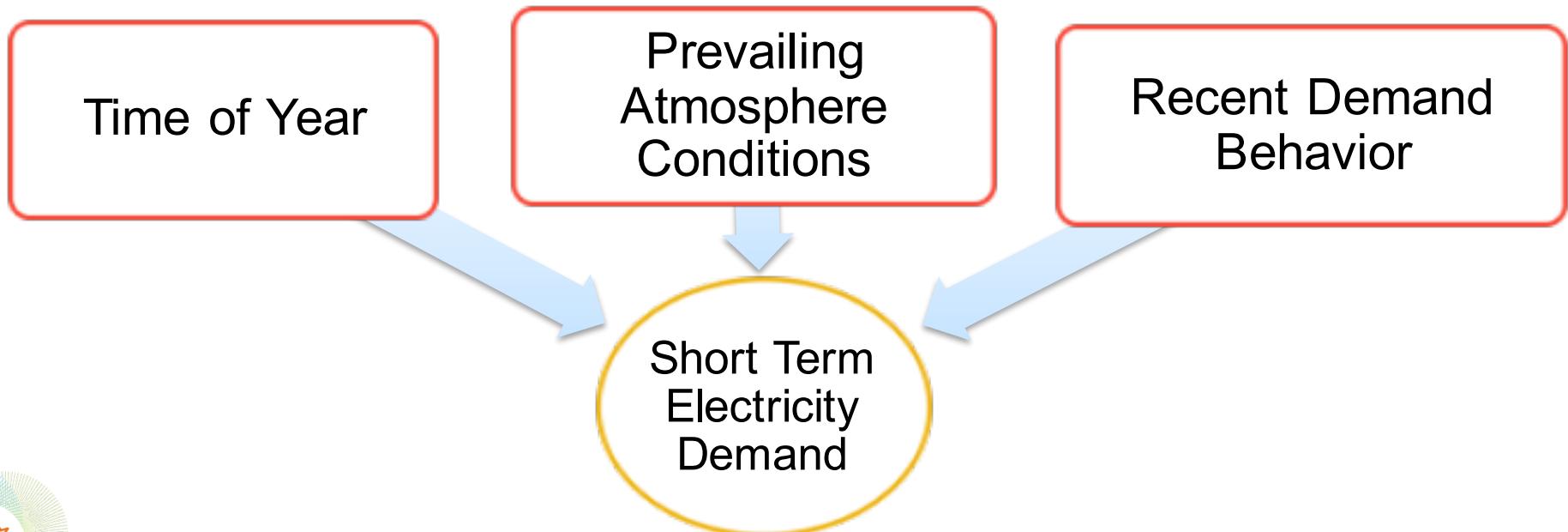
- Machine / Statistical Learning
- Semiparametric Regression
- Artificial Neural Networks
- Fuzzy Logic
- Support Vector Machines
- Gradient Boosting



SPARK SUMMIT 2016

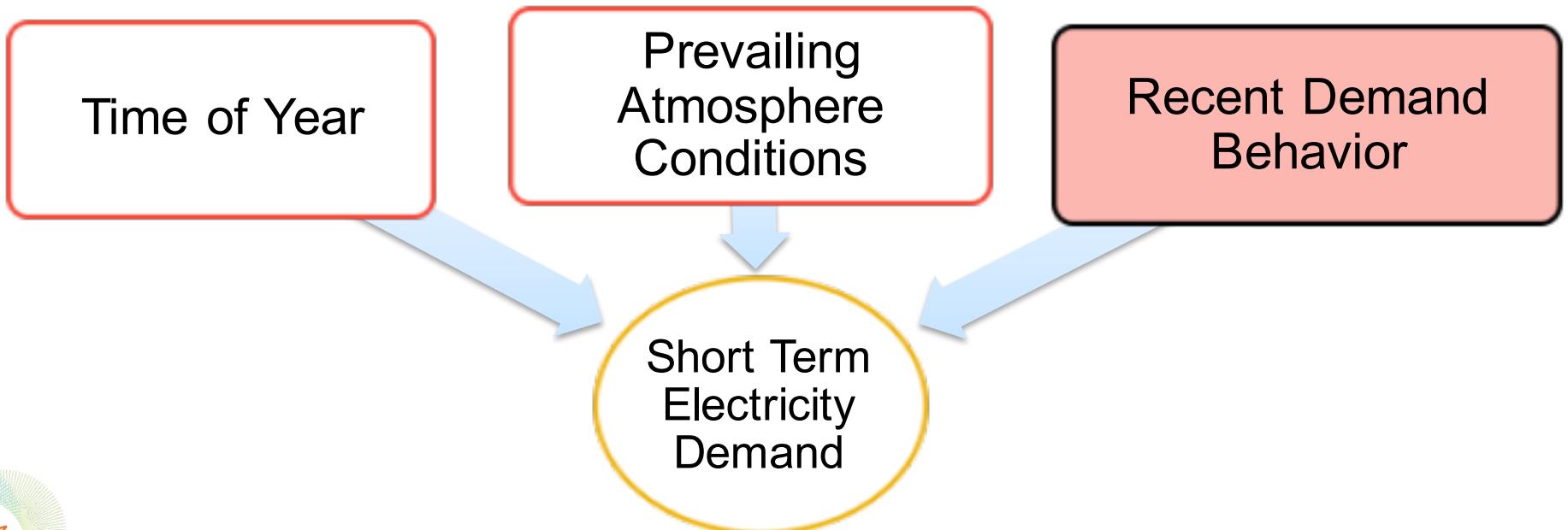
Additive Semiparametric Model

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



Additive Semiparametric Model

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



Emerging Technologies



Photovoltaic
Cells (e.g., Solar)



Electric Vehicles



Storage



Wind



Energy Efficiency

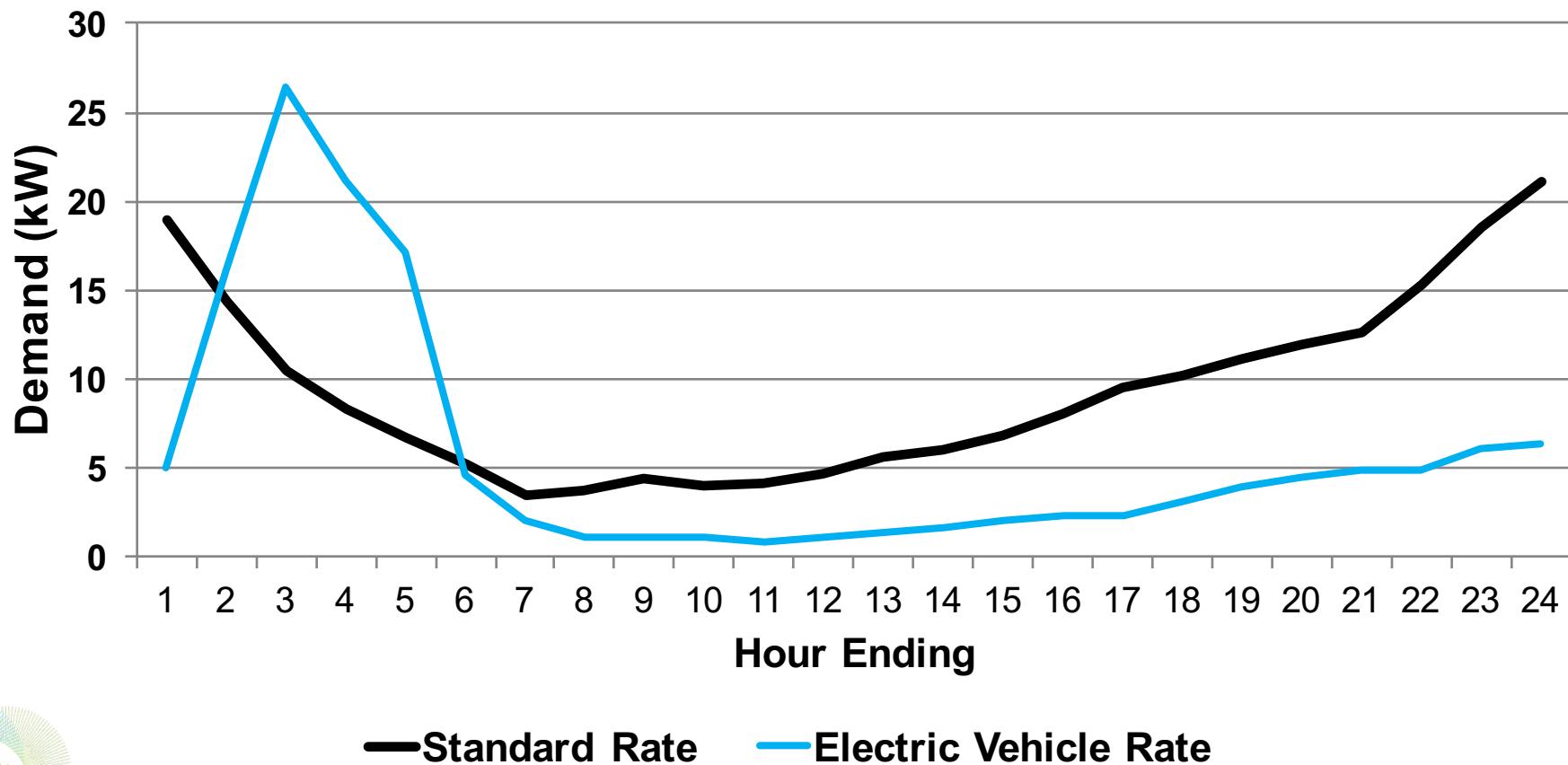


Demand
Response

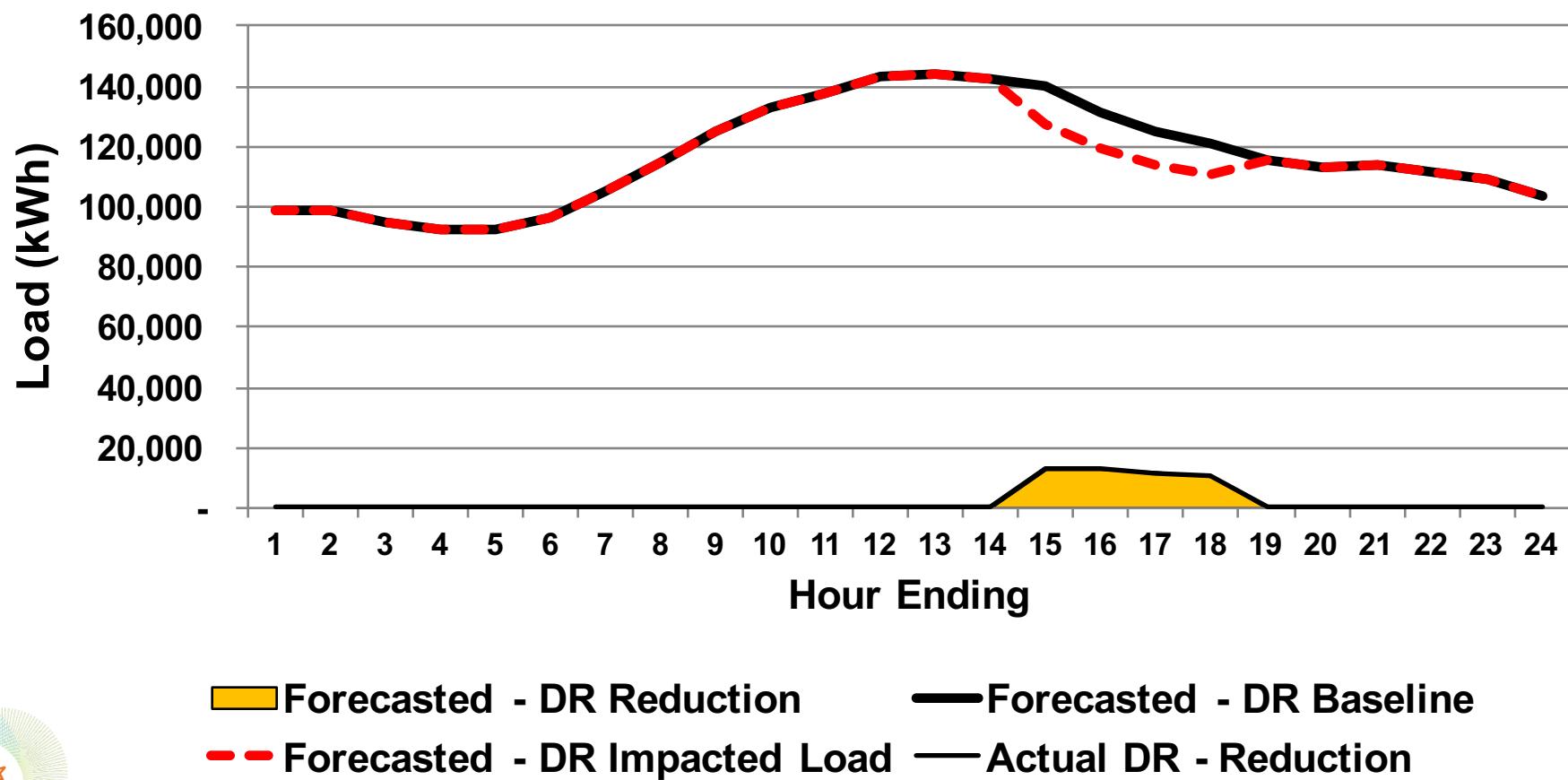


SPARK SUMMIT 2016

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



SPARK SUMMIT 2016

Statistical Computing Pilot

Pilot
Design

Data
Generating
Process

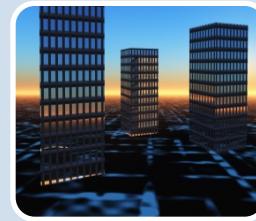
Analytics



SPARK SUMMIT 2016

Statistical Computing Pilot

Data Diversity



Energy Consumption

Climatic
- Temperature
- Humidity
- Wind Speed
- Solar

Demographic Firmographic

Economic Financial

Energy Efficiency Program Tracking

Grid Infrastructure

Key Focus Areas

Performance

Scalability

Granularity



SPARK SUMMIT 2016

Going Further



SPARK SUMMIT 2016

Use Cases



SPARK SUMMIT 2016

DEMONSTRATION



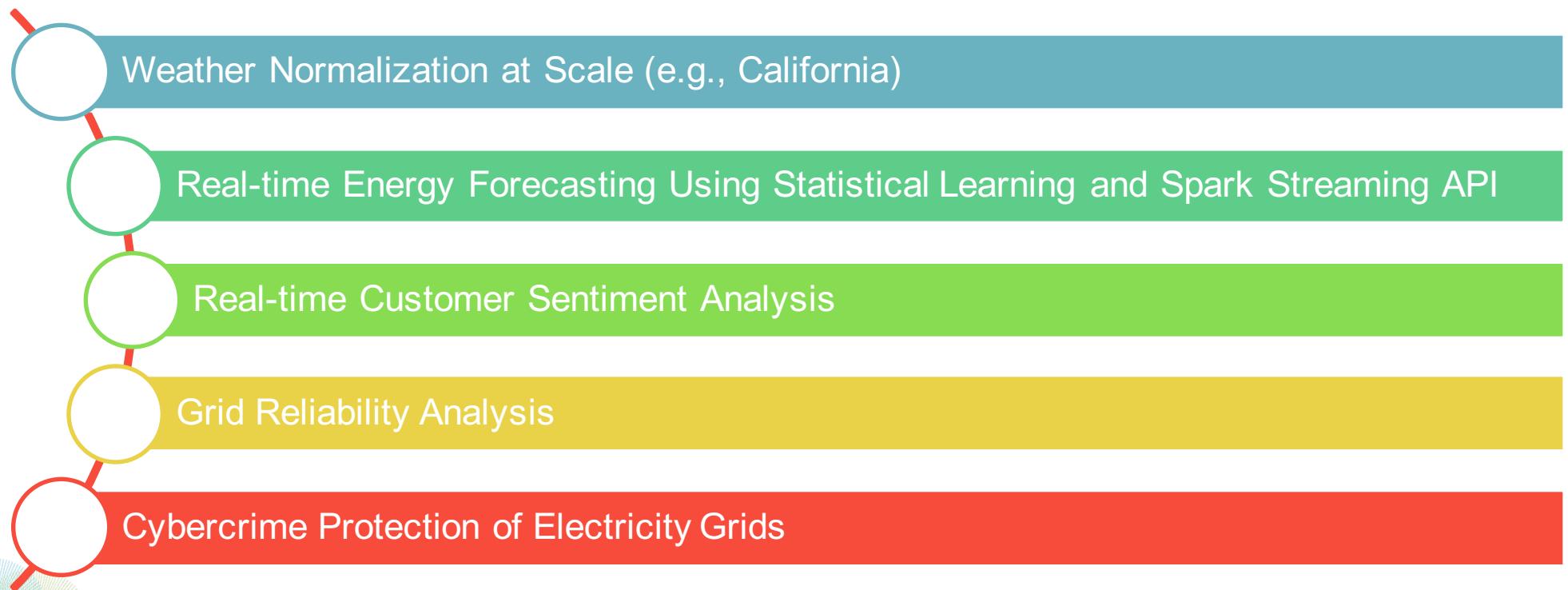
SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE
JUNE 6-8, 2016 SAN FRANCISCO

VISION OF THE FUTURE



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

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