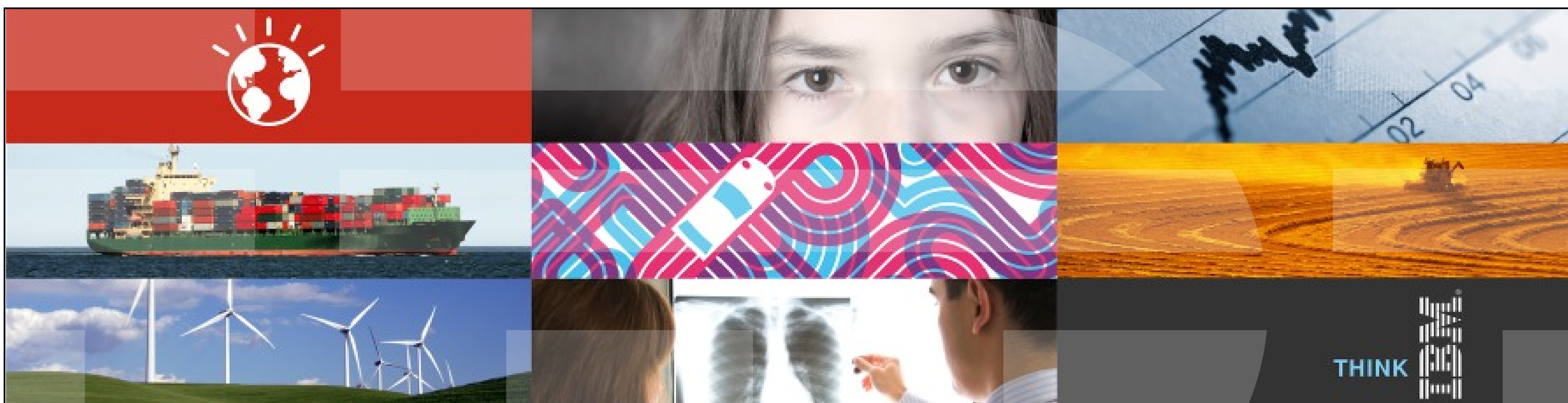
Energy Demand Forecasting
Industry Practices and Challenges

Mathieu Sinn (IBM Research)

12 June 2014
ACM e-Energy
Cambridge, UK



SmarterCities
Technology Centre

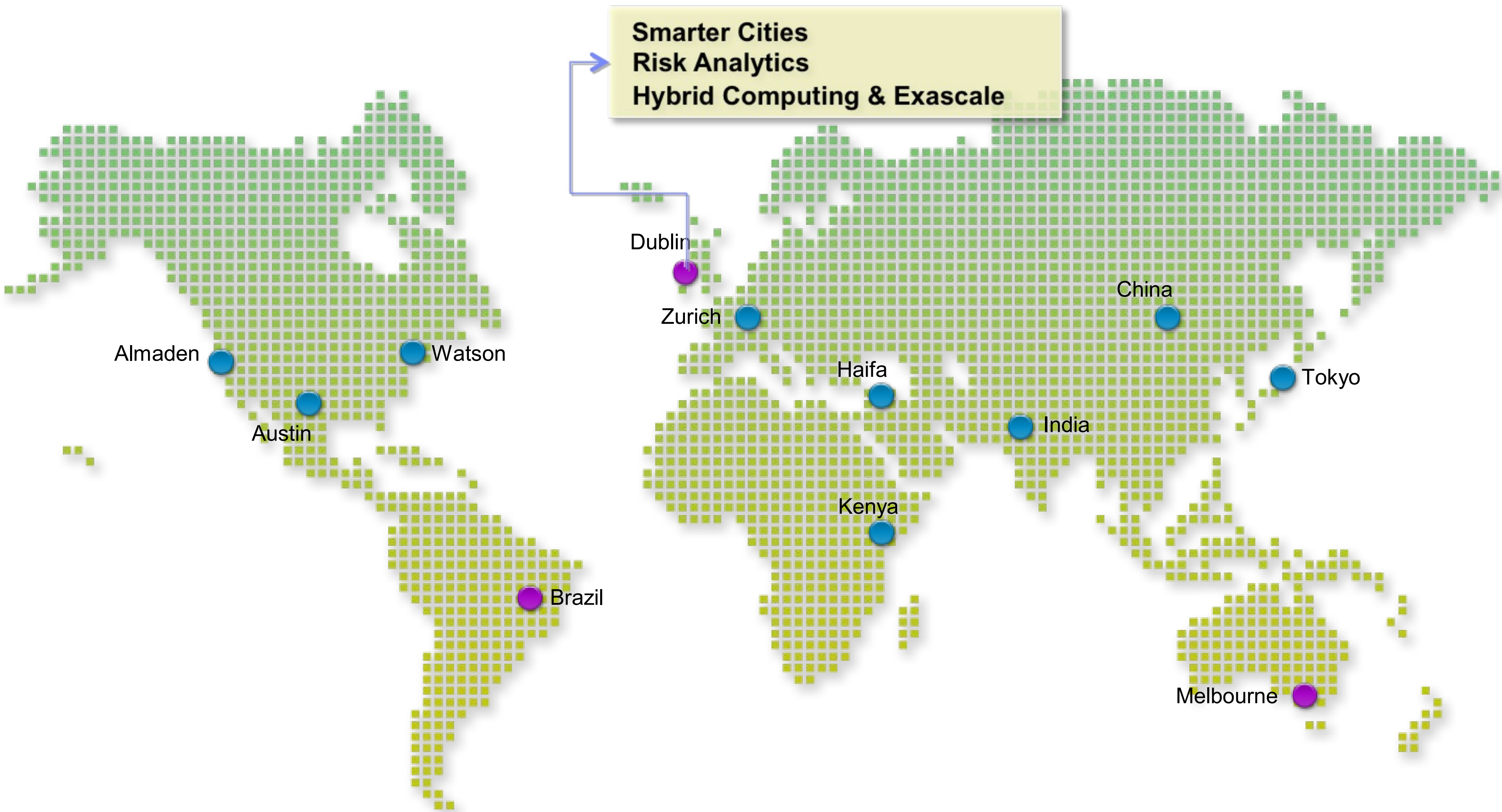
© 2014 IBM Corporation

Outline

- Overview: Smarter Energy Research at IBM
- Energy Demand Forecasting
  - Industry practices & state-of-the-art
  - Generalized Additive Models (GAMs)
  - Insights from two real-world projects
  - Ongoing work and future challenges
- Conclusions

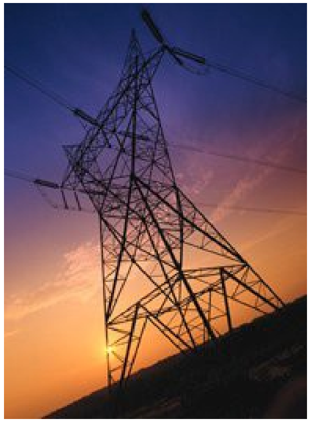
© 2014 IBM Corporation

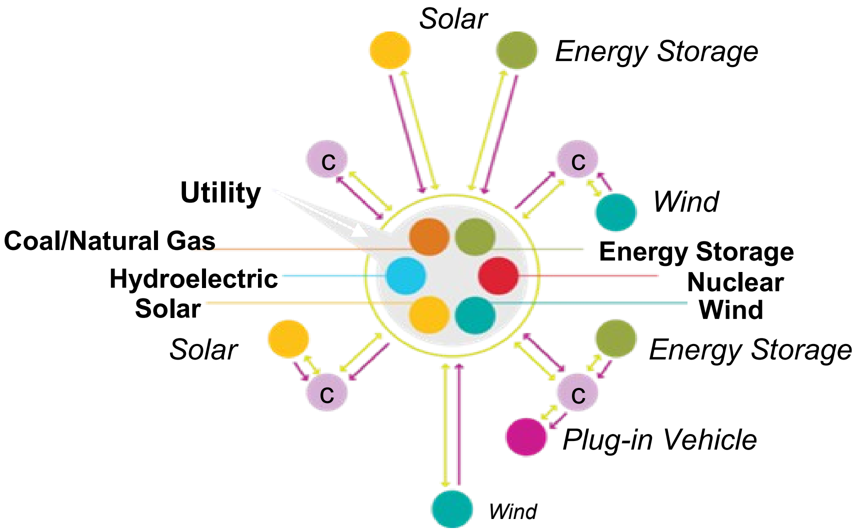
IBM Research Labs



© 2014 IBM Corporation

Smarter Energy Research at IBM
Overview





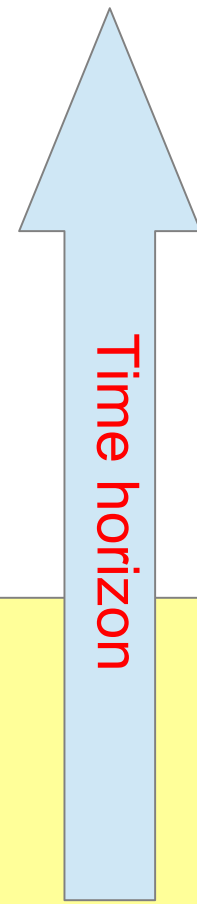
Non-exhaustive project list:

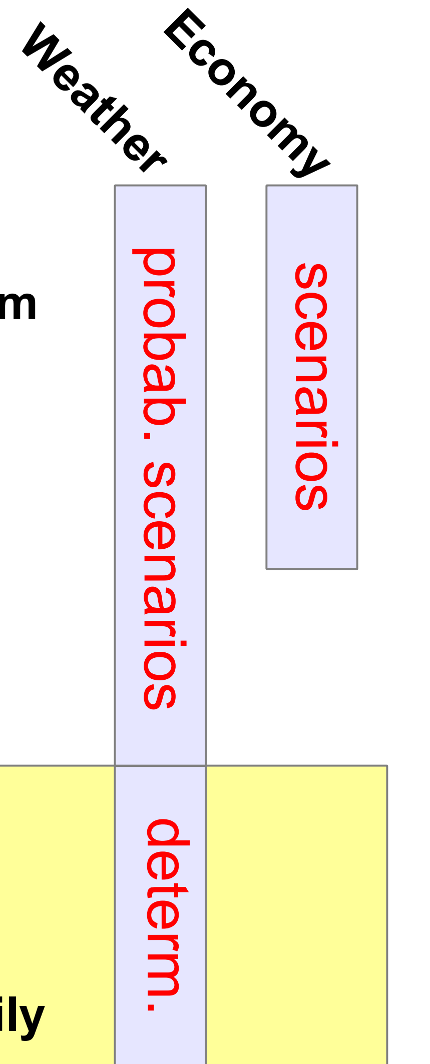
- Pacific Northwest Smart Grid (transactive control, internet-scale control systems)
- Renewable energy forecasts
  – Deep Thunder (weather), HyREF (wind power), Watt-Sun (PV)
- IBM Smarter Energy Research Institute (http://www.research.ibm.com/client-programs/seri/)
  – Outage Prediction and Response Optimization
  – Analytics and Optimization Management System (AOMS)

© 2014 IBM Corporation

Energy Demand Forecasting
Motivation

• Portfolio structuring	Long-term
• Power plants maintenance schedule • Future energy contracts	Yearly
• Energy storage management	Weekly
• Unit commitment, Economic dispatch • Day-ahead outage planning	Daily
• Market purchases/sales	Intra-daily

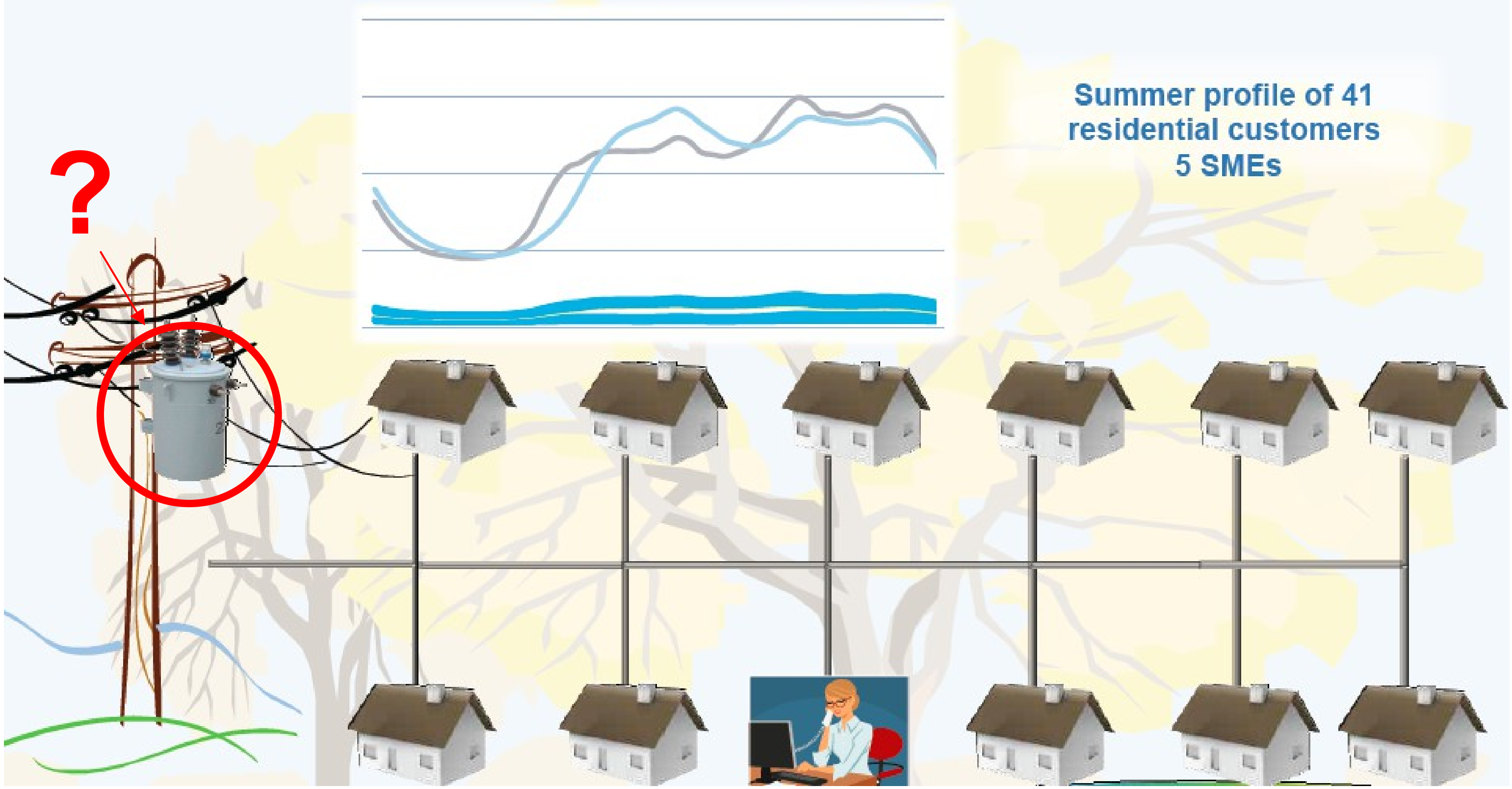




© 2014 IBM Corporation

Energy Demand Forecasting
Motivation

Beyond forecasting: Load modeling and prediction



Source E.Diskin: Can "big data" play a role in the new DSO definition?
European Utility Week, Amsterdam October 2013

© 2014 IBM Corporation

Energy Demand Forecasting

Current practice:

- Forecasting few, highly aggregated series
- Manual monitoring and fine-tuning

Challenges:

- Forecasts at lower aggregation levels → huge amounts of data
- Changes in customer behavior
- Distributed renewable energy sources

} **dynamic!**

Requirements:

Analytical models	Systems
- accurate, flexible, robust	- scalability, throughput
- automated (online learning)	- data-in-motion and -at-rest
- transparent, understandable	- external interface

© 2014 IBM Corporation

Energy Demand Forecasting

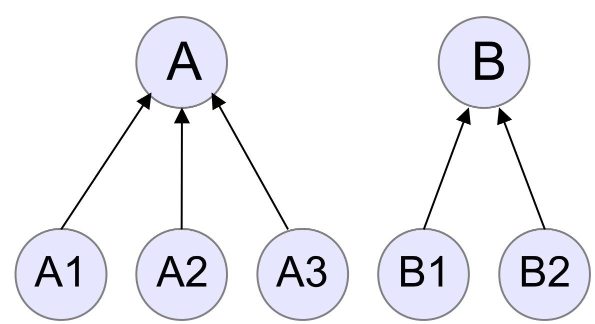
Objectives:

- Forecasting energy demand at various aggregation levels
  - Transmission and Subtransmission networks
  - Distribution substations and MV network
  - Breakdown by customer groups

top-down

Rationales:

- Disaggregate demand for higher forecasting accuracy
  - Local effects of weather, socio-economic variables etc.
- More visibility on loads in Subtransmission and Distribution networks
  - Understanding the effect of exogenous variables
  - Detecting trends, anomalies, etc.
  - Accounting for reconfiguration events

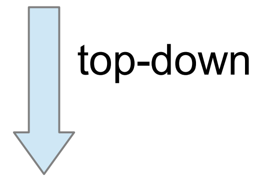


© 2014 IBM Corporation

Energy Demand Forecasting

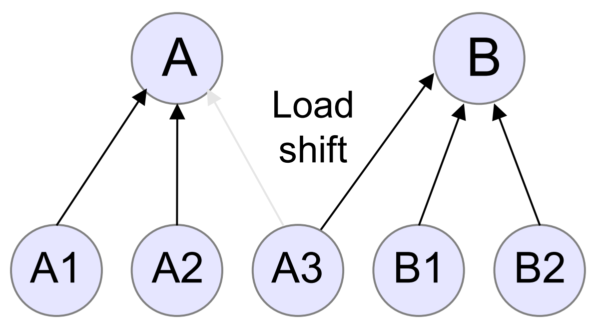
Objectives:

- Forecasting energy demand at various aggregation levels
  - Transmission and Subtransmission networks
  - Distribution substations and MV network
  - Breakdown by customer groups



Rationales:

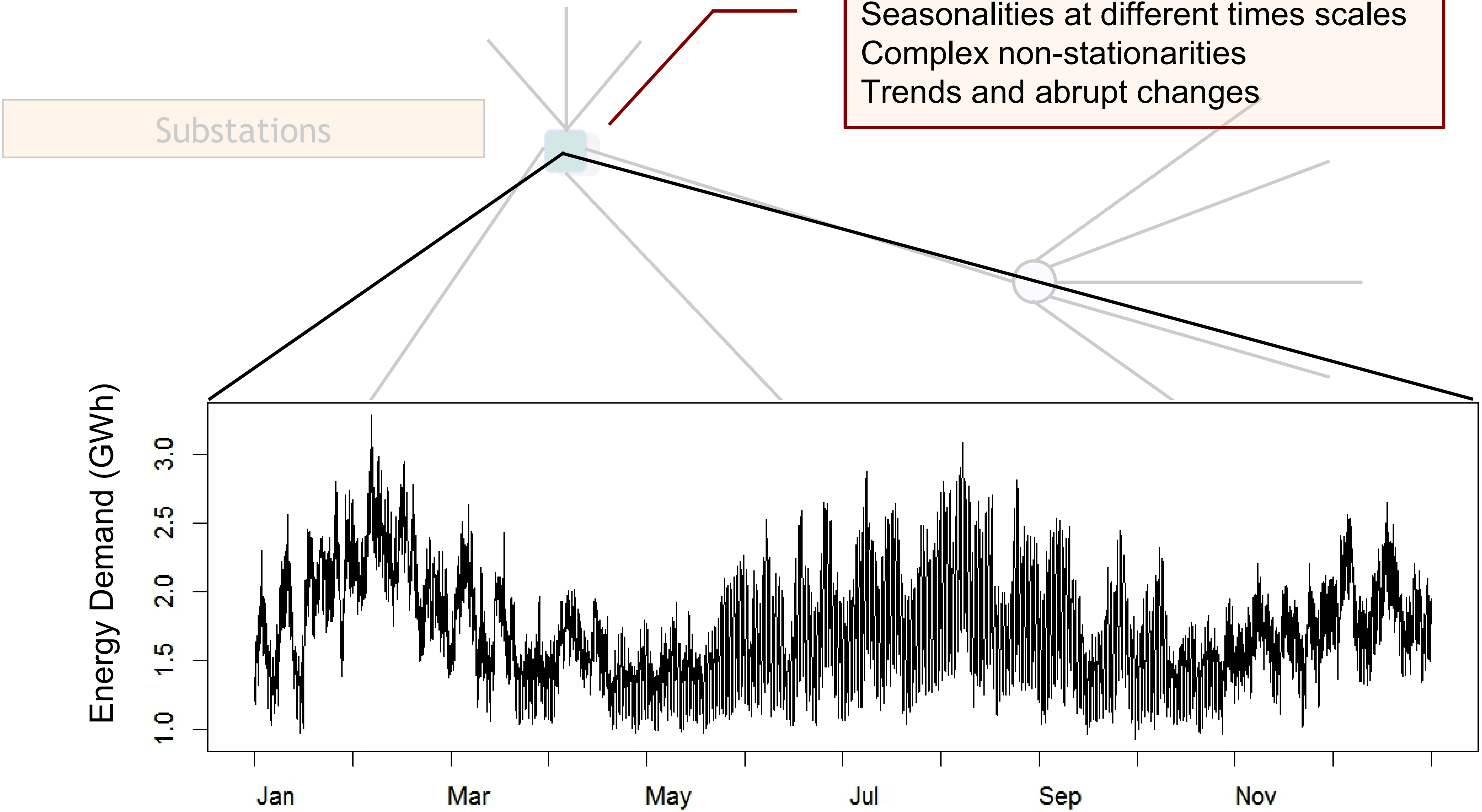
- Disaggregate demand for higher forecasting accuracy
  - Local effects of weather, socio-economic variables etc.
- More visibility on loads in Subtransmission and Distribution networks
  - Understanding the effect of exogenous variables
  - Detecting trends, anomalies, etc.
  - Accounting for reconfiguration events



© 2014 IBM Corporation

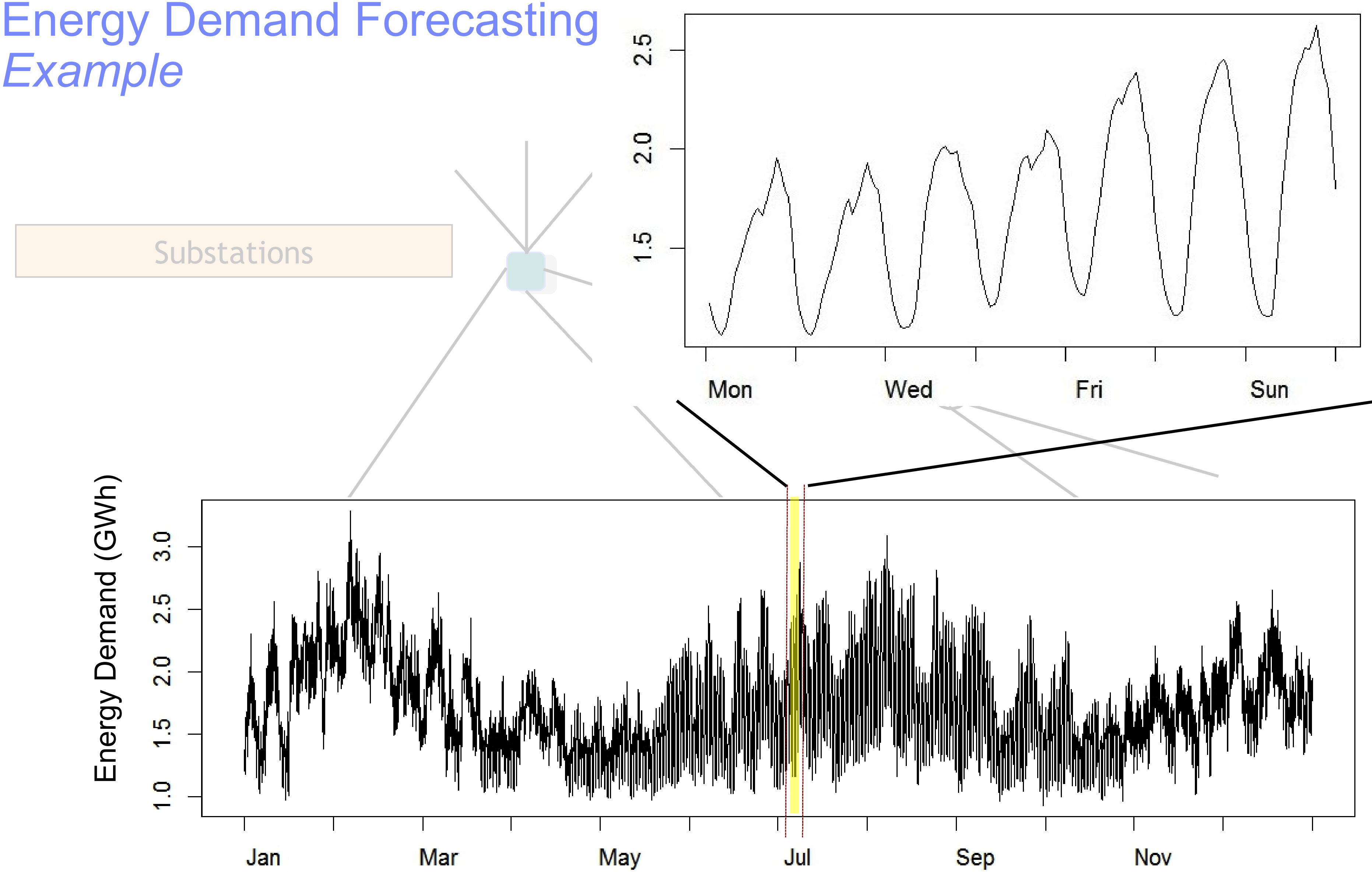
Energy Demand Forecasting
Example

Seasonalities at different times scales
Complex non-stationarities
Trends and abrupt changes



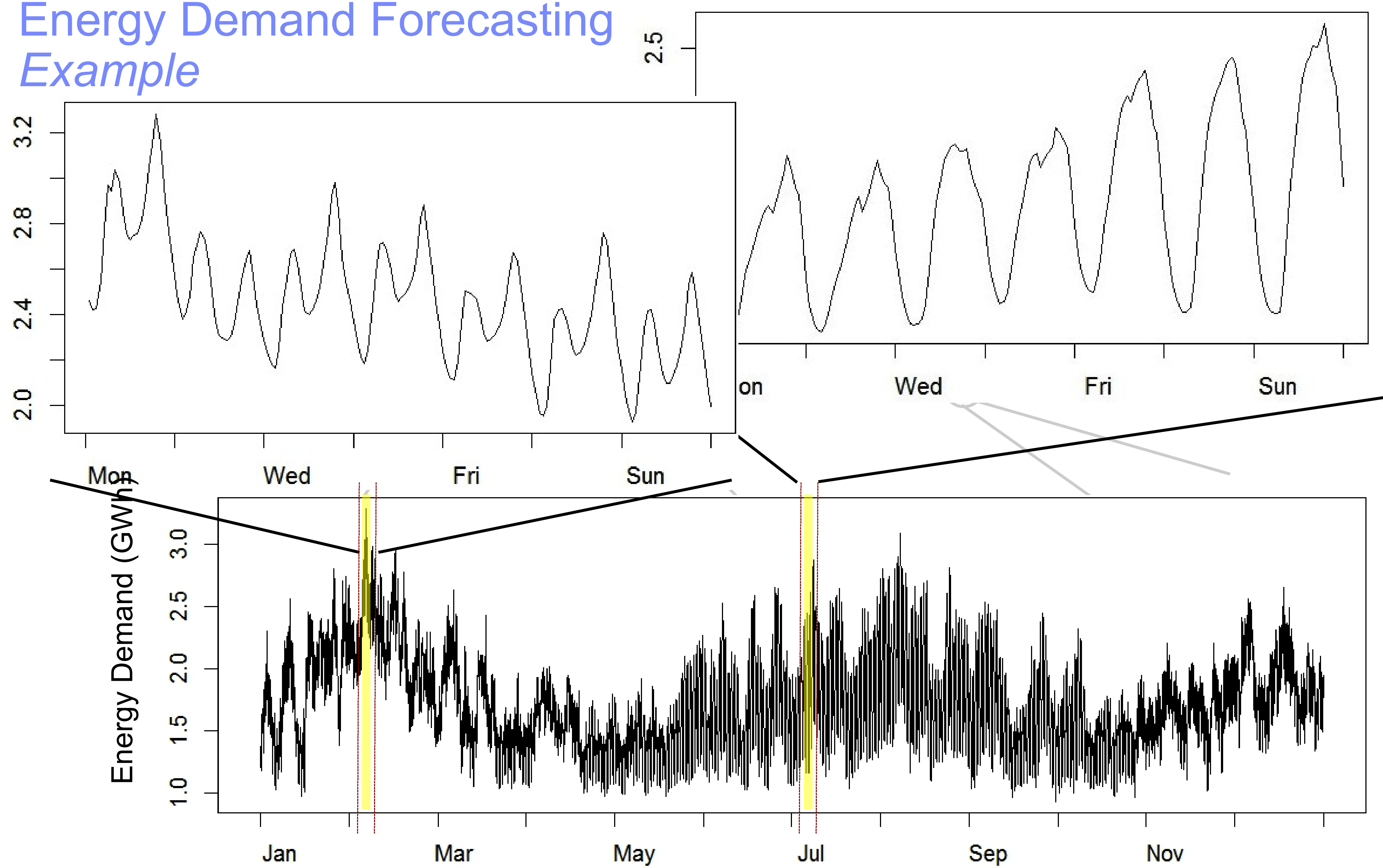
© 2014 IBM Corporation

Energy Demand Forecasting
Example



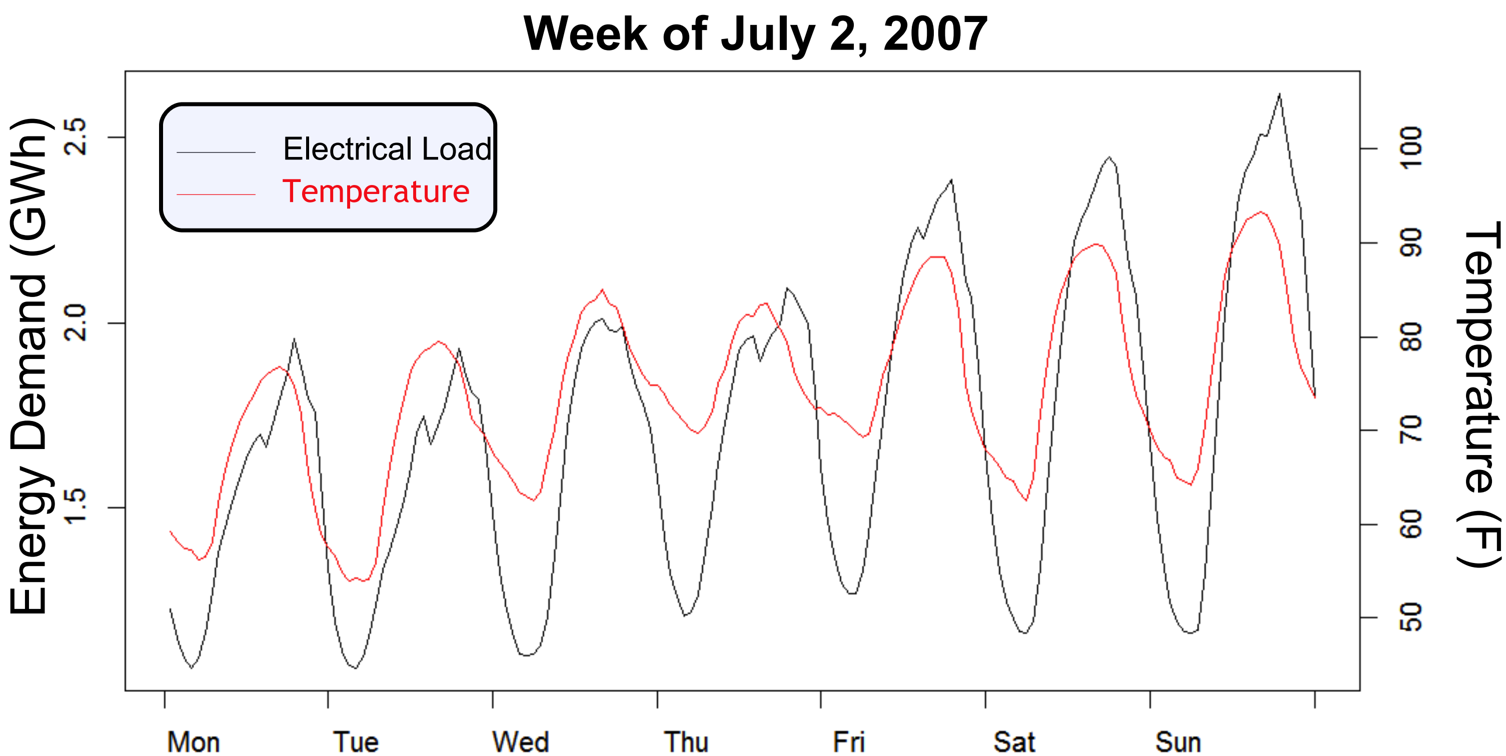
© 2014 IBM Corporation

Energy Demand Forecasting *Example*



© 2014 IBM Corporation

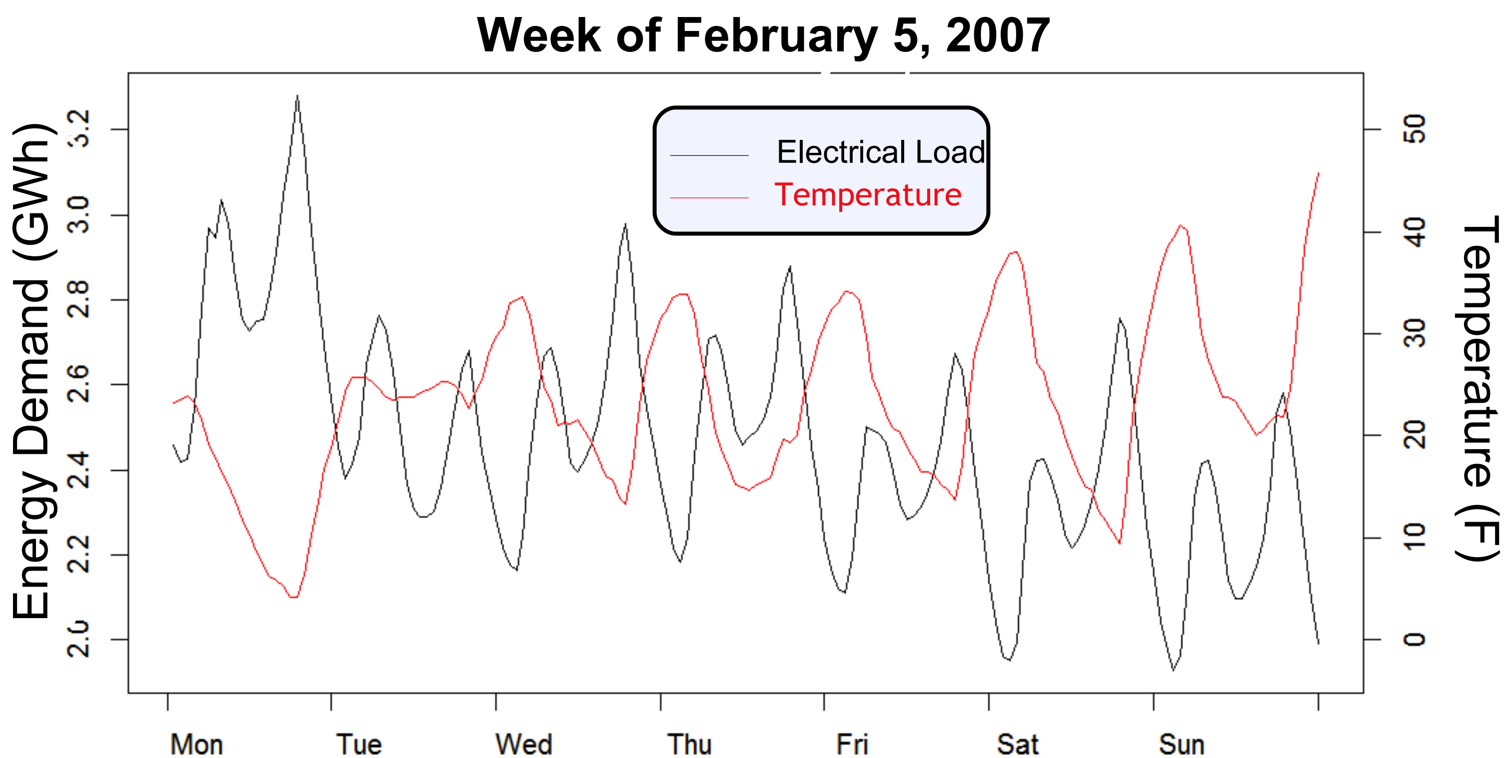
Energy Demand Forecasting
Example



SmarterCities
Technology Centre

© 2014 IBM Corporation

Energy Demand Forecasting
Example



© 2014 IBM Corporation

Energy Demand Forecasting
Additive Models

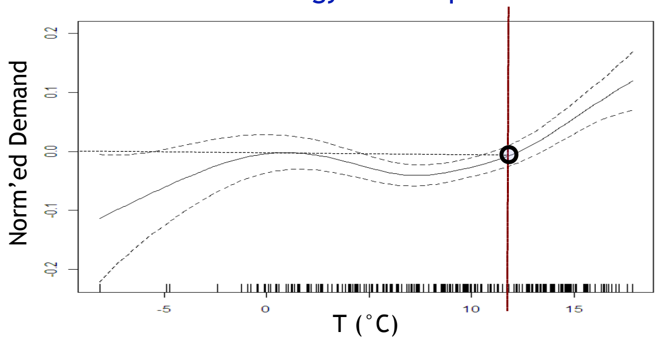
Assumption: effect of covariates is additive

Illustrative example:

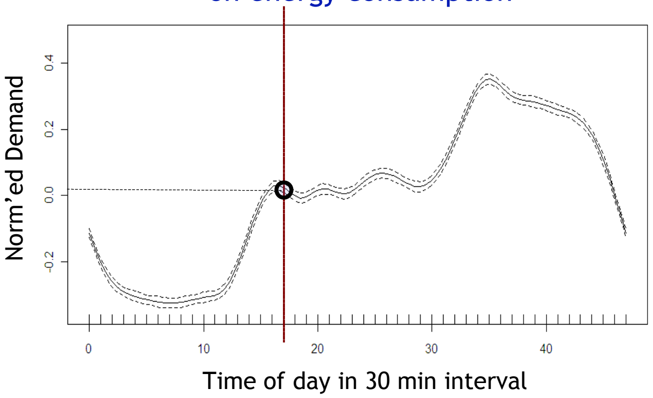
- $x_k = (x_k^{Temperature}, x_k^{TimeOfDay})$ *(covariates)*
- $y_k = f^{Temperature}(x_k) + f^{TimeOfDay}(x_k)$ *(transfer functions)*
- Say, x_k = (12°C, 08:30 AM)

$\rightarrow$ y_k = 0.0 GW + 0.02 GW

Contribution of temperature on energy consumption



Contribution of time of day on energy consumption



© 2014 IBM Corporation

Energy Demand Forecasting
Additive Models

Formulation:

Demand — Transfer functions — Noise

$$y_k = \sum_{i=1}^{I} f_i(x_k) + \epsilon_k$$

Transfer functions have the form:

Categorical condition — Basis functions

$$f_i(x_k) = \mathbf{1}_{A_i}(x_k)\beta_i^T b_i(x_k)$$

Weights

This includes:

- constant, indicator, linear functions
- cubic B-splines (1- or 2-dimensional)

Covariates:
- Calendar variables (time of day, weekday...)
- Weather variables (temperature, wind ...)
- Derived features (spatial or temporal functionals)
- ...

© 2014 IBM Corporation

Energy Demand Forecasting
Additive Models

Formulation:

$$y_k = \beta^T b(x_k) + \epsilon_k$$

Linear in basis functions

Training:

1) Select covariates, design features
2) Select basis functions (= knot points)
3) Solve Penalized Least Squares problem

Penalizer

$$\hat{\beta}_K = \arg\min_{\beta} \left\{ \|\boldsymbol{y}_K - \boldsymbol{B}_K \beta\|^2 + \beta^T \boldsymbol{S}(\lambda_K)\beta \right\}$$

where λ_K is determined using Generalized Cross Validation

S. Wood (2006): *Generalized Additive Models. An Introduction with R.*

© 2014 IBM Corporation

Energy Demand Forecasting
EDF-IBM NIPS model

Model for 5 years of French national demand (Feb 2006 – April 2011)[1]

Covariates:

$$x_k = (x_k^{\text{DayType}}, x_k^{\text{TimeOfDay}}, x_k^{\text{TimeOfYear}}, x_k^{\text{Temperature}}, x_k^{\text{CloudCover}}, x_k^{\text{LoadDecrease}})$$

- DayType: *1=Sun, 2=Mon, 3=Tue-Wed-Thu, 4=Fri, 5=Sat, 6=Bank holidays*
- TimeOfDay: *0, 1, ..., 47 (half-hourly)*
- TimeOfYear: *0=Jan 1st, ..., 1=Dec 31st*
- Temperature: *spatial average of 63 weather stations*
- CloudCover: *0=clear, ..., 8=overcast*
- LoadDecrease: *activation of load shedding contracts*

[1] A.Ba, M.Sinn, P.Pompey, Y.Goude: Adaptive learning of smoothing splines. Application to electricity load forecasting. Proc. Advances in Neural Information Processing Systems (**NIPS**), 2012

© 2014 IBM Corporation

Energy Demand Forecasting
EDF-IBM NIPS model

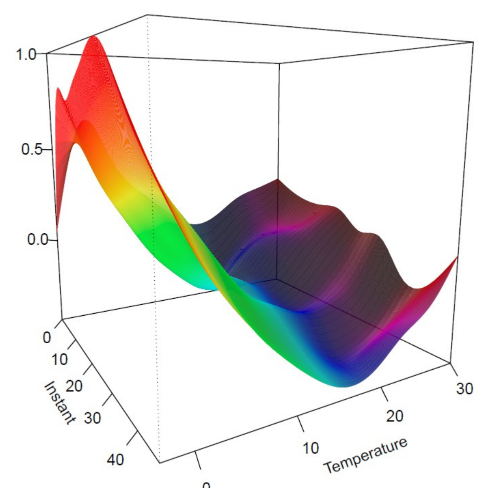
Model:

Trend | Lag load | Day-type specific daily pattern

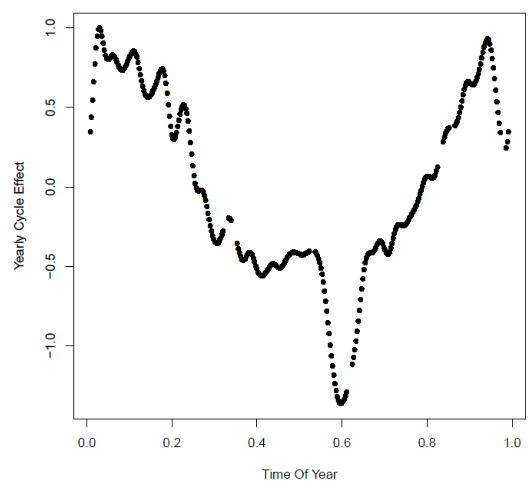
$$y_k = \beta^{\text{Intercept}} + f^{\text{Trend}}(k) + f^{\text{LagLoad}}(y_{k-48}) + \sum_{l=1}^{6} \mathbf{1}(x_k^{\text{DayType}} = l)(\beta_l^{\text{DayType}} + f_l^{\text{TimeOfDay}}(x_k))$$
$$+ f^{\text{CloudCover}}(x_k) + f^{\text{Temperature/TimeOfDay}}(x_k) + f^{\text{LagTemperature}}(x_{k-48})$$
$$+ f^{\text{TimeOfYear}}(x_k) + x_k^{\text{LoadDecrease}} f^{\text{LoadDecrease}}(x_k) + \epsilon_k.$$

Lag temperature (accounting for thermal inertia)

Transfer functions:



TimeOfDay / Temperature



TimeOfYear

Results:

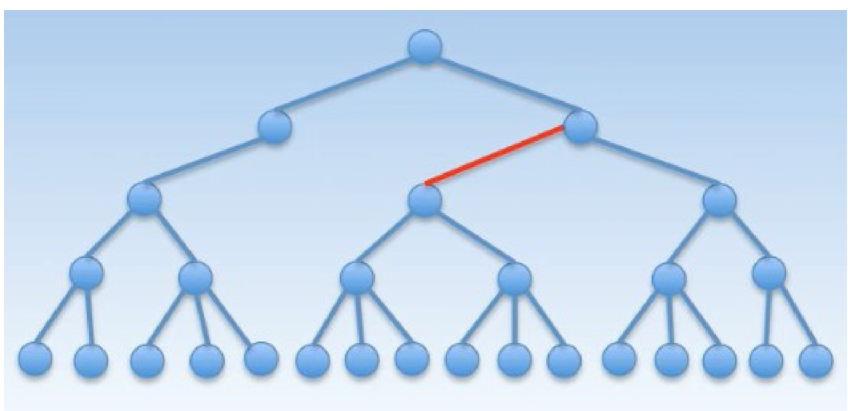
<u>1.63%</u> MAPE
<u>20%</u> improvement by online learning
<u>Explanation:</u> macroeconomic trend effect

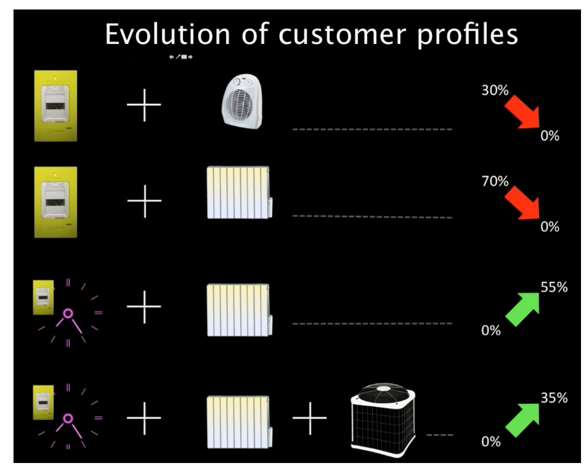
© 2014 IBM Corporation

Energy Demand
EDF-IBM Simulation platform

Simulation:

- Massive-scale simulation platform for emulating demand in the future electrical grid[2]
  - 1 year half-hourly data, 35M smart meters
  - Aggregation by network topology (with dynamic configurations)
  - Changes in customer portfolio
  - Distributed renewables (wind, PV)
  - Electric vehicle charging
- Built on IBM InfoSphere Streams





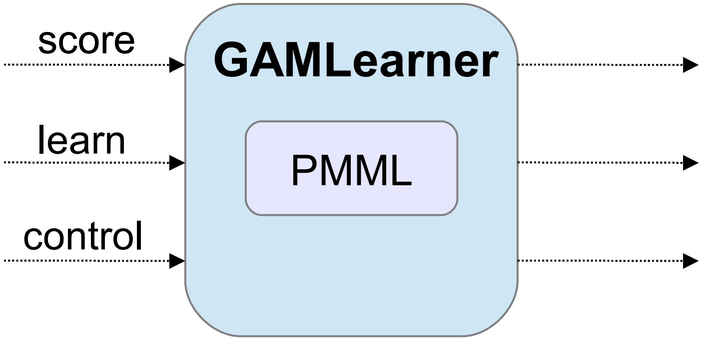
[2] P.Pompey, A.Bondu, Y.Goude, M.Sinn: Massive-Scale Simulation of Electrial Load in Smart Grids using Generalized Additive Models. *Springer Lecture Notes in Statistics (to appear)*, 2014.

© 2014 IBM Corporation

Energy Demand Forecasting
Online learning

Forecasting:

- Statistical approach: Generalized Additive Models (GAMs)
  - Accuracy, flexibility, robustness, understandability ...
- Developing GAM operators for IBM InfoSphere Streams



PMML = Predictive Models Markup Language

- Online learning:
  - Tracking of trends (e.g., in customer portfolio)
  - Reducing human intervention
  - Incorporating new information

© 2014 IBM Corporation

Energy Demand Forecasting
Online learning

Formulation of GAM learning as **Recursive Least Squares:**

"Kalman gain"

Model parameter

Forecasting error

$$\hat{\beta}_{K+1} = \hat{\beta}_K + \frac{\boldsymbol{P}_K b_{K+1}}{\omega + b_{K+1}^T \boldsymbol{P}_K b_{K+1}} (y_{K+1} - b_{K+1}^T \hat{\beta}_K)$$

Precision matrix

$$\boldsymbol{P}_{K+1} = \omega^{-1} \left(\boldsymbol{P}_K - \frac{\boldsymbol{P}_K b_{K+1} b_{K+1}^T \boldsymbol{P}_K}{\omega + b_{K+1}^T \boldsymbol{P}_K b_{K+1}} \right)$$

- Adapt model once actual demand becomes available ($\hat{\beta}_K \rightarrow \hat{\beta}_{K+1}$)
- Implementation:
  - Forgetting factor $\omega \in (0, 1]$ (discounting past observations)
  - Complexity: $O(p^2)$ (p = number of spline basis functions)
  - Sparse matrix algebra → 1000 tuples per second
  - Adaptive regularization

© 2014 IBM Corporation

Energy Demand Forecasting
Online learning

Stability:

- Incorporate historical sample information in $\boldsymbol{P}_0$
- Rule of thumb for forgetting factor:

$$\textit{time window size} \approx \frac{1}{1-\omega}$$

- Hence, for a time window of *1 year* = *365*48 data points*:

$$\omega = 0.9999429$$

Don't forget
during summer what
happened in winter!

- Another potential issue: divergence of $\boldsymbol{P}_K$
- “Blowing-up” of Kalman gain

© 2014 IBM Corporation

Energy Demand Forecasting
Online learning

Stability:

- $\boldsymbol{P}_K$ is the inverse of the sum of discounted matrix terms

$$b(x_k)^T b(x_k), \qquad k = 1, 2, \ldots, K.$$

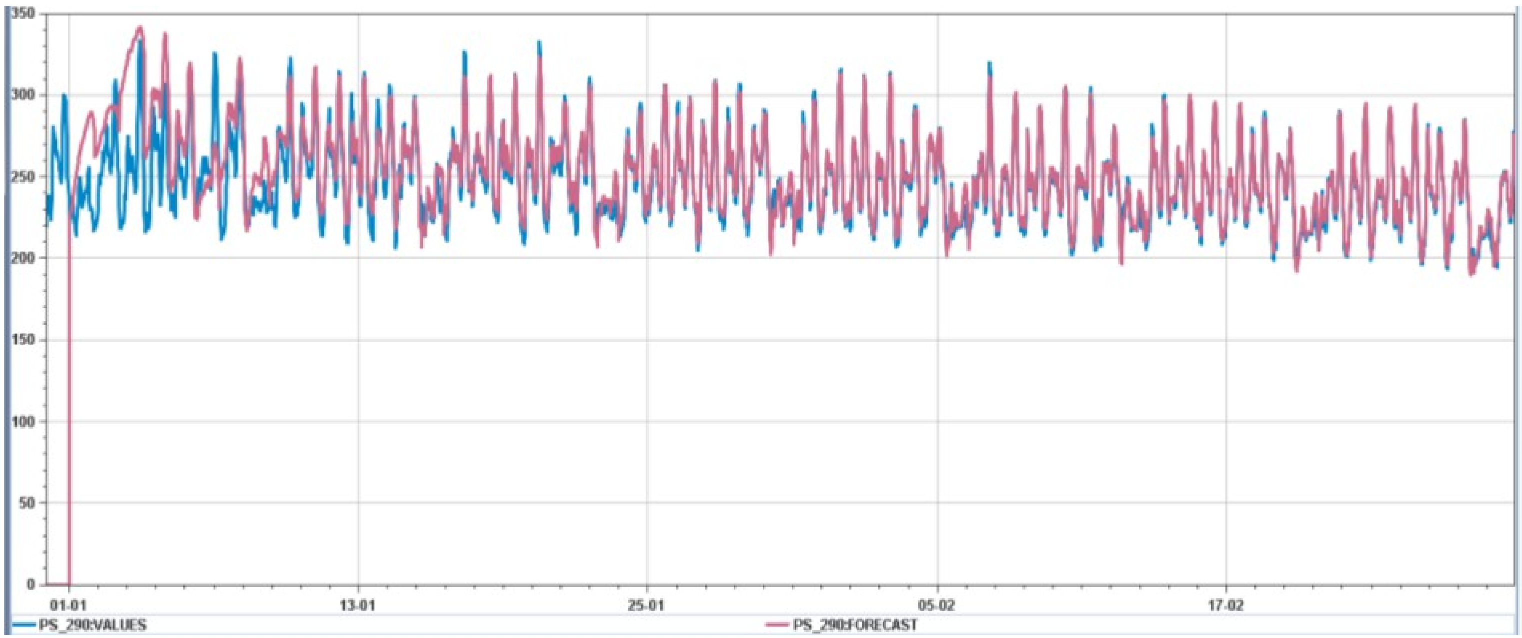
Outer product of
spline basis functions

- Divergence can occur, e.g.,
  - if subset of basis functions is (almost) collinear
  - if subset of basis functions is (almost) always zero
- Solution: **Adaptive regularizer**
  - Monitor matrix norm of $\boldsymbol{P}_K$
  - If norm exceeds threshold, then add diagonal matrix to the inverse of $\boldsymbol{P}_K$
  - Complexity: $O(p^3)$

© 2014 IBM Corporation

Energy Demand Forecasting
Online learning

Learning "from scratch" (initial parameters $\hat{\beta}_0$ all equal to zero):



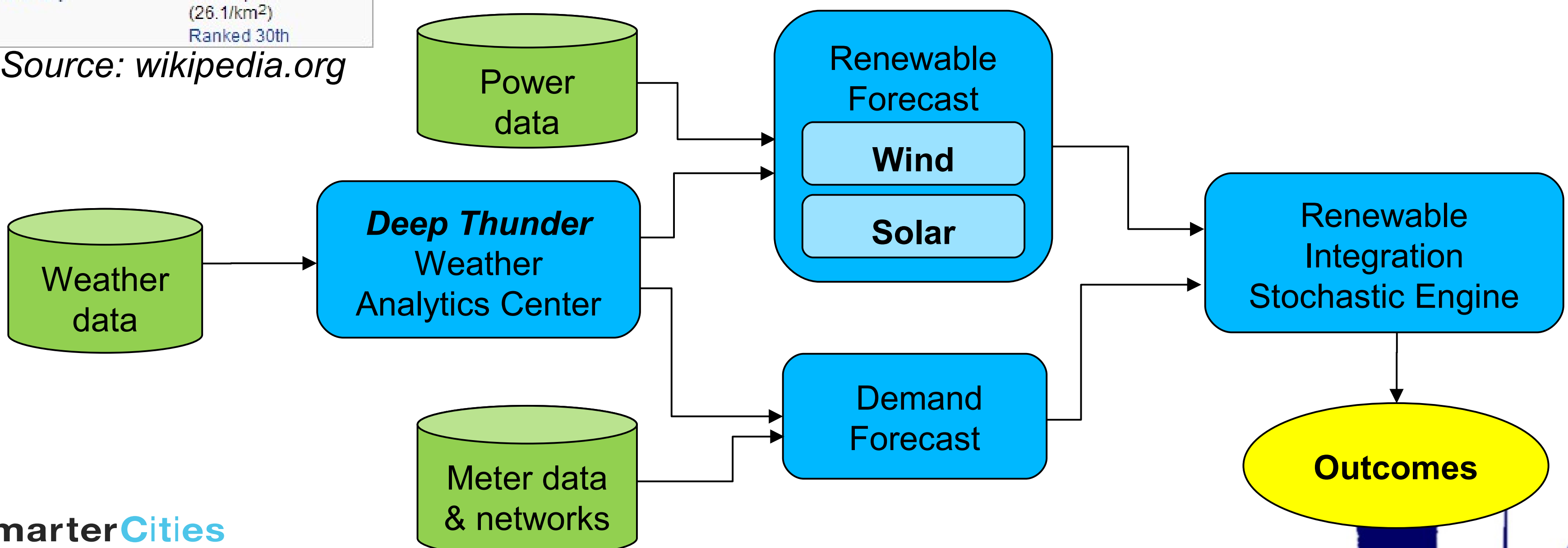
© 2014 IBM Corporation

Vermont Project
Scope

State of Vermont	

Area	Ranked 45th
- Total	9,620 sq mi (24,923 km²)
Population	Ranked 49th
- Total	626,630 (2013 est)[1]
- Density	67.7/sq mi (26.1/km²) Ranked 30th

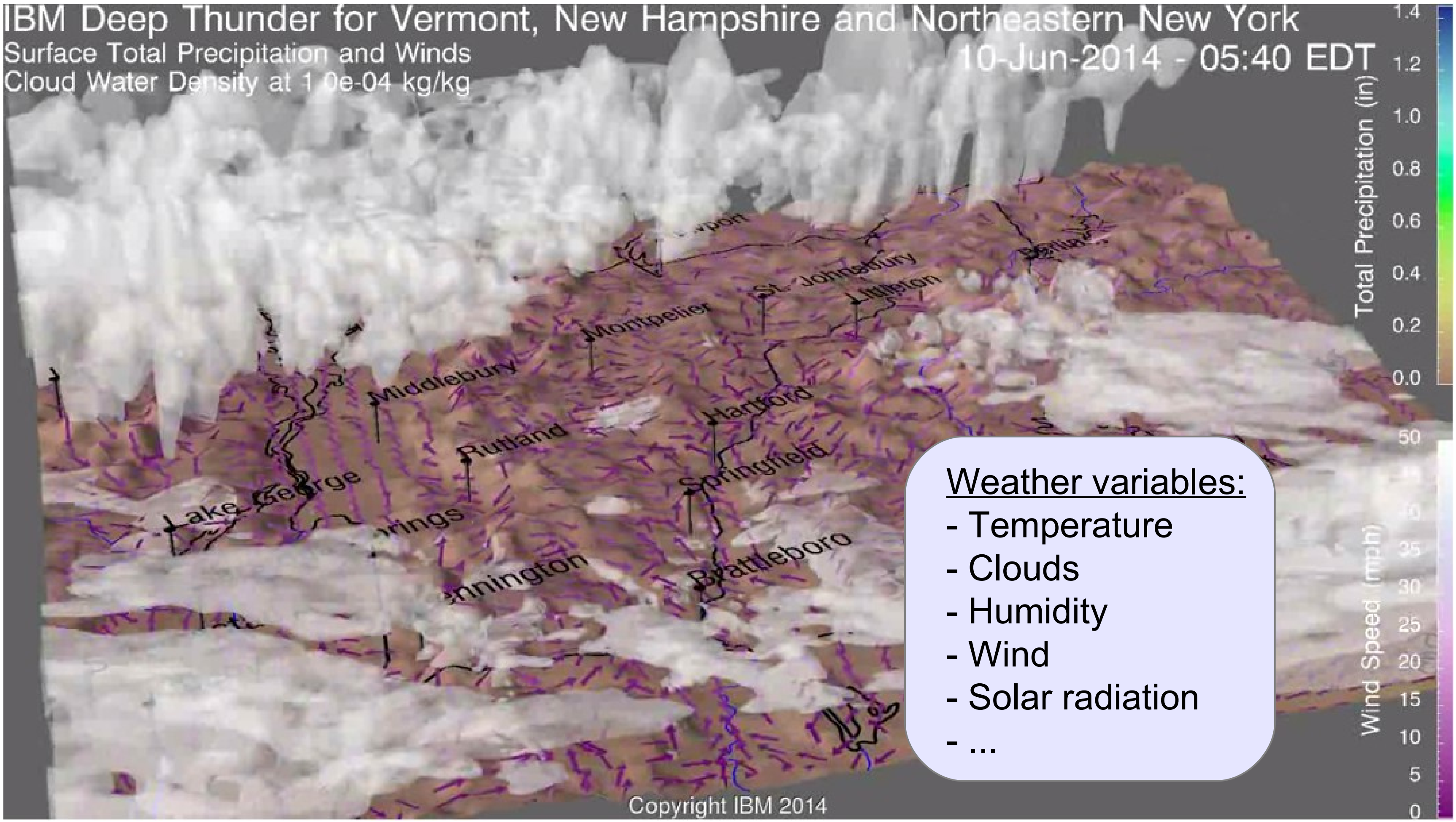
Source: wikipedia.org



SmarterCities
Technology Centre

© 2014 IBM Corporation

Vermont Project
Deep Thunder



<u>Weather variables:</u>
- Temperature
- Clouds
- Humidity
- Wind
- Solar radiation
- ...

© 2014 IBM Corporation

Vermont Project
Demand forecasting challenges

Modeling:
- Distributed renewables “behind the meter”
- Forecasting uncertainty

Variable selection & feature extraction:
- Spatial averages of weather variables
- Temporal features (e.g., heat waves)
- Formalization & automatization

Transfer learning:
- How to integrate information from older, lower-resolution data sets?

Transparent analytics:
- GUI which allows users to “drive” analytics withouth in-depth statistical knowledge

© 2014 IBM Corporation

Conclusions

- Smarter Energy Research at IBM
- Energy demand forecasting
  - Current practices & future challenges
  - Methodology
  - Insights from two projects

Thank you!

© 2014 IBM Corporation