One Day Ahead Hourly Load Forecasting

An Introduction

Tao Hong, PhD

Principal Engineer, Quanta Technology

tao.hong@quanta-technology.com

Guest Lecture at NC State University
March 31st , 2011

Tao Hong

- Education
  - B.Eng, Automation, Tsinghua University, Beijing
  - M.S. E.E., NC State University, Raleigh
  - M.S. O.R., I.E. NC State University, Raleigh
  - Ph.D. E.E., O.R., NC State University, Raleigh
- Experience
  - *Engineer, Sr. Engineer, Principal Engineer,* Quanta Technology, LLC.
  - *Business Knowledge Series Instructor,* SAS Institute, Inc.
- Interest
  - Load / price forecasting, energy market
  - System planning, power engineering

Objectives

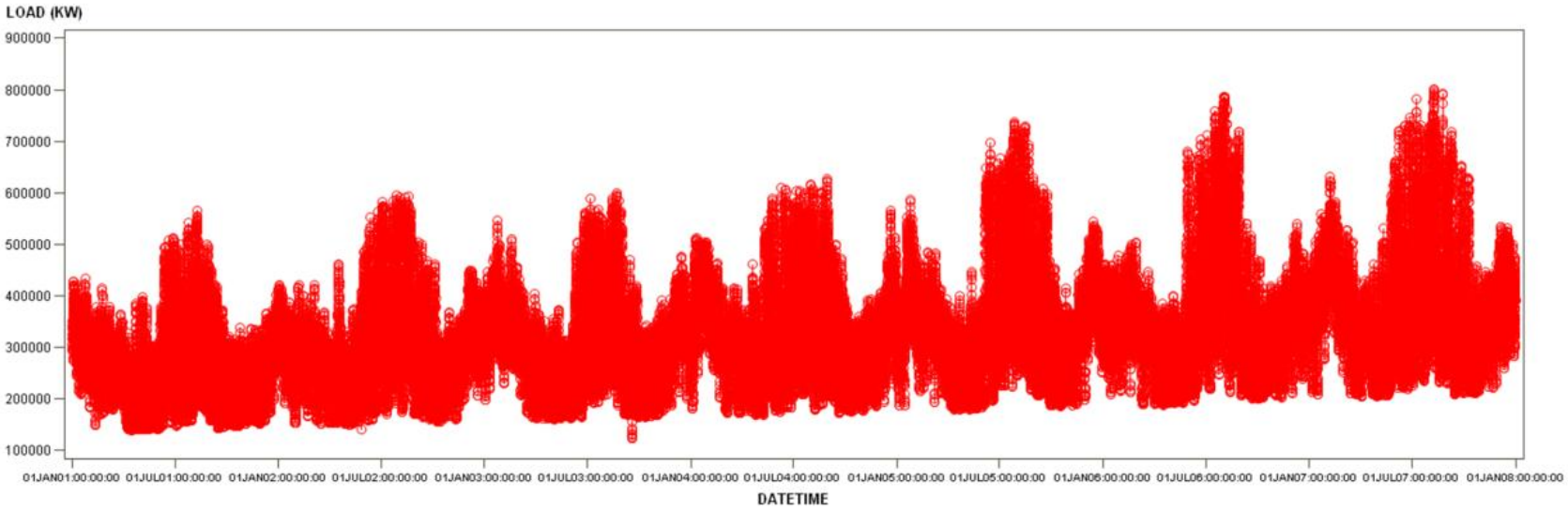
- Taste the flavor of load forecasting
- Understand basic concepts of short term load forecasting
- Learn how to
  - discover salient features of electric load series using graphic methods
  - capture the salient features using regression models
  - build regression models using SAS Enterprise Guide
  - develop a naïve model for one day ahead hourly load forecasting

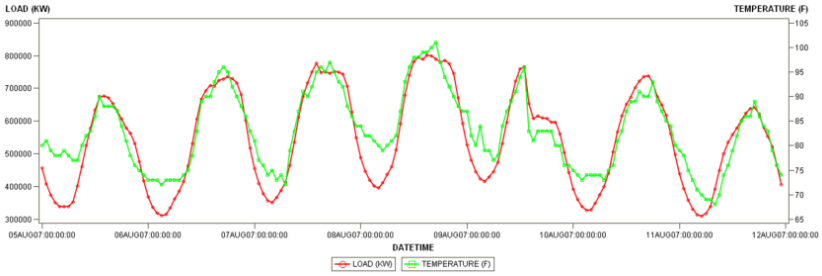
Notes, data, and reading materials can be downloaded from the lecture webpage:
http://courses.drhongtao.com/dahlf

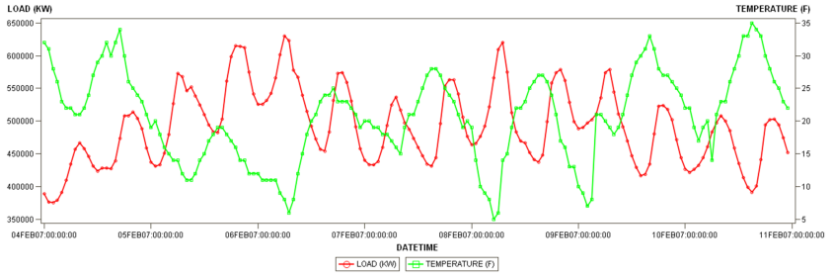
Outline

- Load Forecasting: What, Why, and How?
- General Linear Models
- A Naïve Multiple Linear Regression Based Benchmark
- Demonstration (SAS)
- 10 Ways to Screw up Your Load Forecasts

What's Electric Load Forecasting?





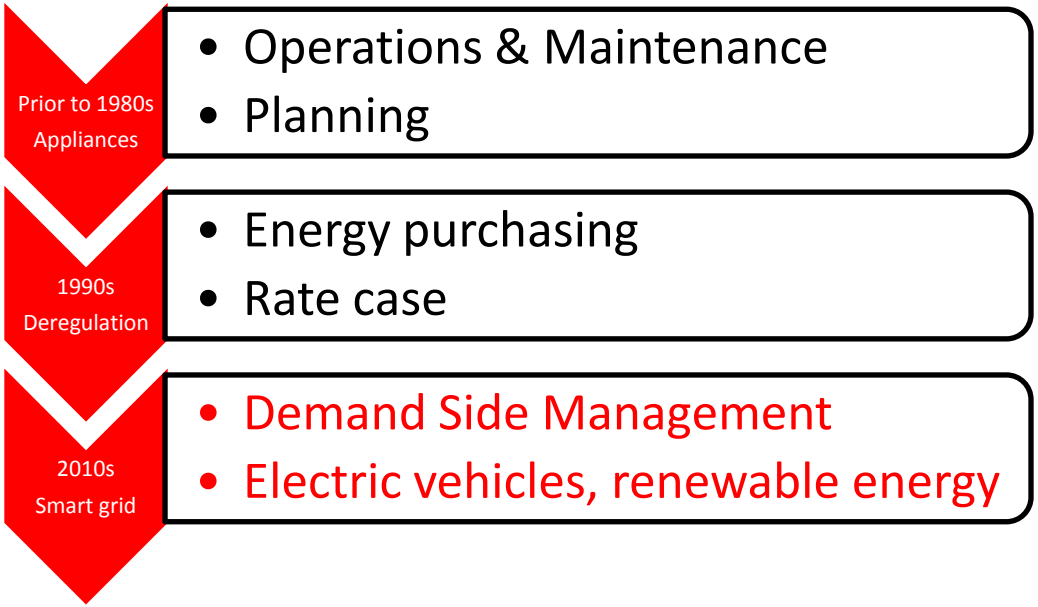


Why do We Need Load Forecasts?

- BSE is currently seeking a Senior Load Forecasting Specialist. The ideal candidate is prepared to meet the following expectations:
  - This position directs the processes necessary to produce the forecast of the total demand for electricity within the BSE balancing area along with forecasts of the sale of electricity by BSE and others and BSE's transmission demand requirements as part of BSE's overall Integrated Planning Process.
  - The incumbent serves as an expert in the methodologies used in the in load forecasting and is responsible for directing from beginning to end the one or more of the forecasting processes. The incumbent plans and insures that the appropriate research and design is completed to assure that BSE has load forecasting procedures and methodology which are equal to or exceed current standard forecasting processes when compared to other major electric utilities.
  - Accurate forecasts are essential to decisions impacting BSE's operating practices and the type/timing of resource acquisitions such as decisions to build, lease, or sell generation and transmission assets and the decisions to purchase or sell wholesale power.

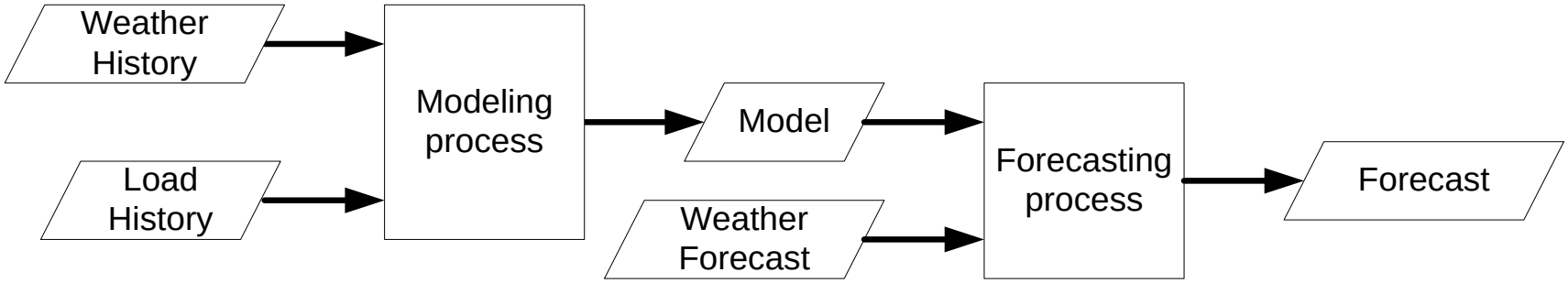
Why do We Need Load Forecasts?

- Generation
- Transmission
- Distribution
- Planning
- Operations & maintenance
- Market related activities



How to Forecast Electricity Demand?

- A load forecasting process



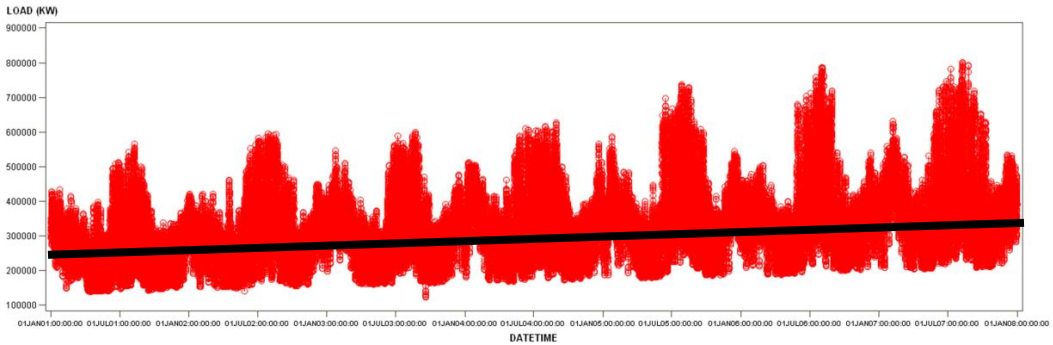
- Modeling process
  - Statistics: linear regression, Box-Jenkins, nonparametric regression
  - A.I.: ANN, fuzzy logic, neuro-fuzzy system, SVM
  - Others: hybrid approach, multi-stage approach

How to Forecast Electricity Demand?

- Prediction vs. explanation
- Regression analysis
  - Plot the data
  - Select variables
    - Dependent variables
    - Independent variables
  - Build the model
  - Evaluate the model
    - Statistical tests
    - Goodness-of-fit
    - Forecasting accuracy
    - Intuition
  - Document
    - Model
    - Procedure
    - Results
    - Reasoning
    - Any judgmental changes
    - Limitations
  - Report
    - Written report
    - Oral presentation

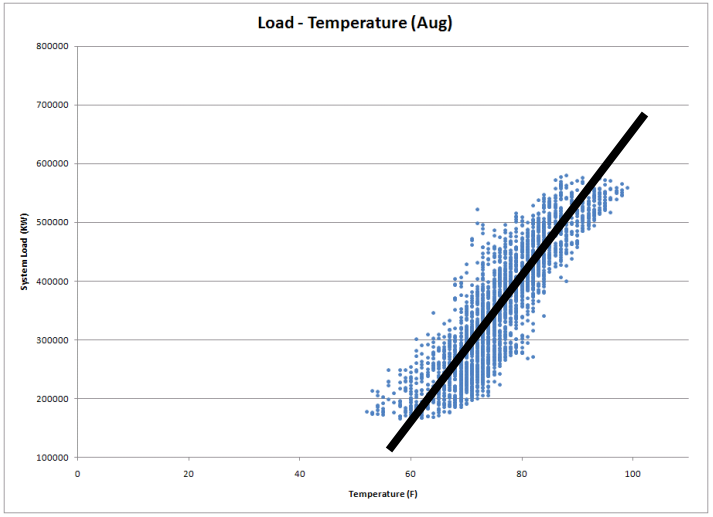
Simple Linear Regression

- Elements
  - Intercept
  - Slope
  - Random error term
- Formal assumptions
  - Linearity
  - Same variance
  - Independent
  - Normally distributed
- Parameter estimation



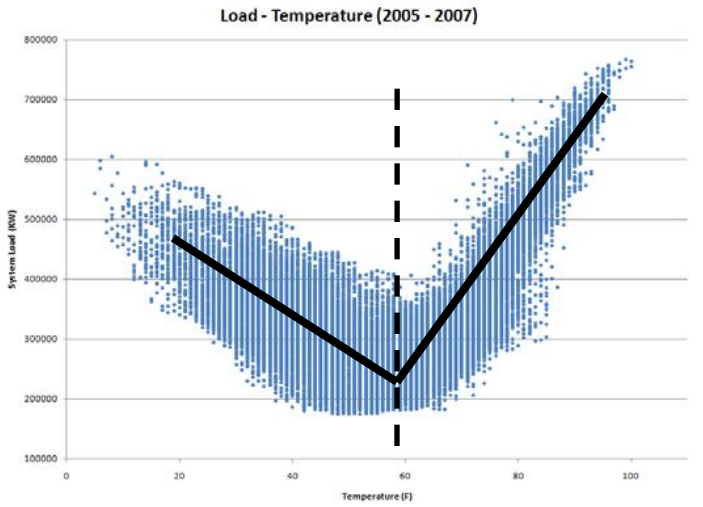
Multiple Linear Regression

- Quantitative variables
  - Trend
  - Temperature



Multiple Linear Regression

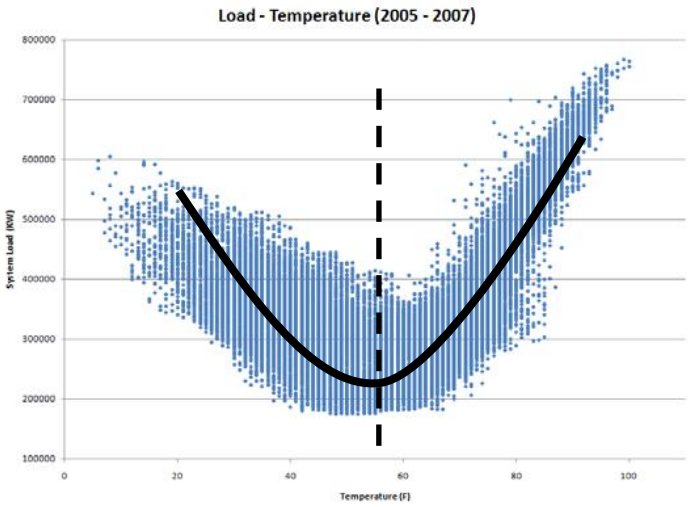
- Qualitative variables
  - 0-1 indicator
    - Piecewise linear regression
  - Calendar variables
    - Months of a year
    - Days of a week
    - Hours of a day



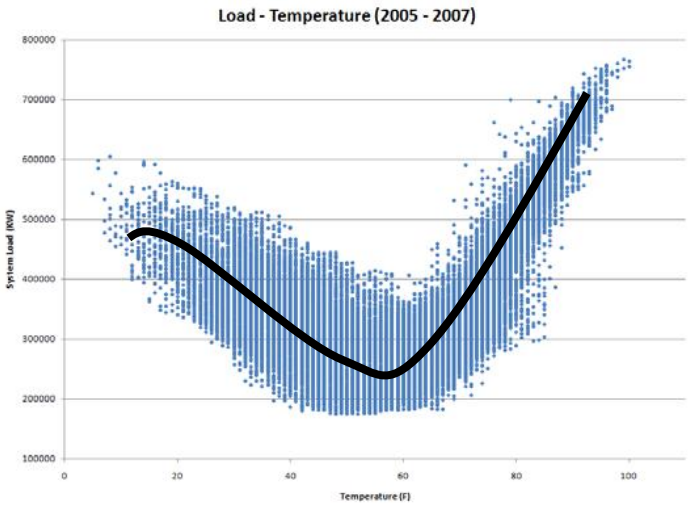
Practice: using SAS Enterprise Guide to plot average daily load profile for 2006.

Polynomial Regression

- Polynomials of independent variable(s)
  - Temperature



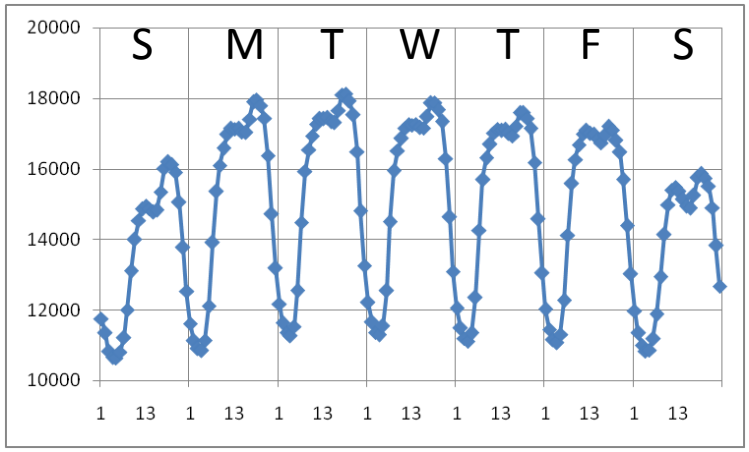
2nd order – symmetric



3rd order – asymmetric

Interaction Regression

- Interactions between/among
  - Quantitative variables
    - A new quantitative variable (*not covered*)
  - Quantitative variables and qualitative variables
    - Month and temperature (plots)
    - Hour and temperature (plots)
  - Qualitative variables
    - Weekday and hour



Linear Models vs. Linear Response Surfaces

- There should be no ambiguousness that the GLM can be used to generate a large variety of nonlinear response surfaces. In other words, linear models are not restricted to linear response surfaces. The term "linear" in GLM refers to the parameters. A regression model is "linear" in the parameters when it can be written as:

$$Y = XB + E$$

where Y is a vector of responses, B is a vector of parameters, X is a matrix of constants, and E is a vector of independent normal random variables with zero expectation and variance-covariance matrix $\sigma^2 I$.

Requirements for a Naïve Benchmark

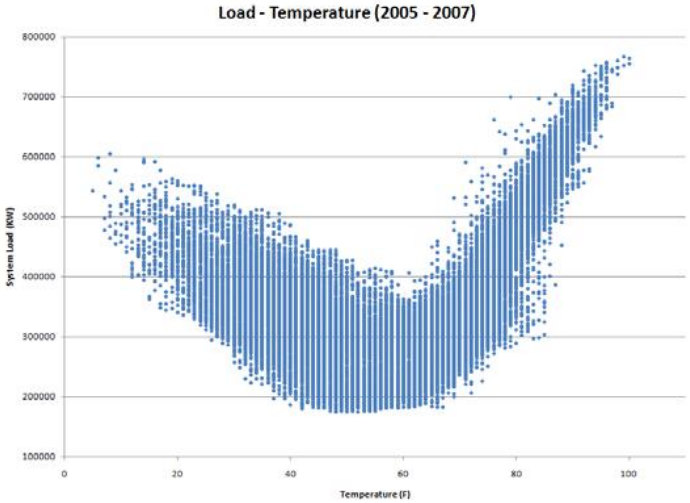
- Simple
  - Easy to implement
- Creditable
  - Fairly accurate and theoretically sound
- Widely applicable
  - Can be used by a wide range of utilities
- Interpretable
  - Can be understood by average electrical engineers, managers, etc.
- Reproducible
  - Can be reproduced based on documented procedures

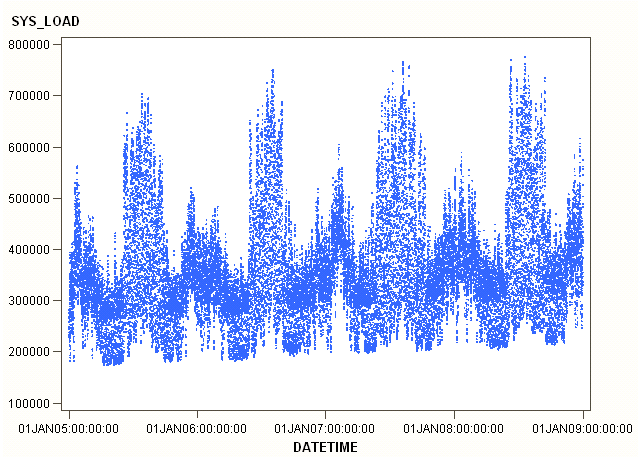
Trend and Seasonality

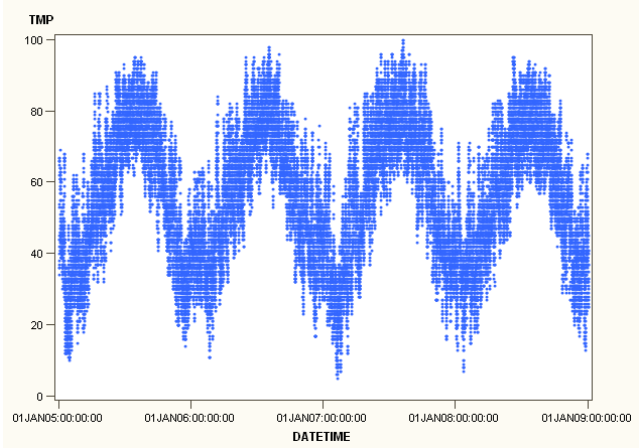
- Trend
  - Linear, macroeconomic indicators (GDP, GSP, employment, etc.)
- In a year
  - 4 seasons, 12 months, etc.
- In a week
  - Weekday/weekend, etc.
- In a day
  - 24 hours, etc.

Temperature

- Piecewise linear
- Quadratic
- 3rd order polynomial
- Etc.







Interactions

- Temperature vs. calendar variables
  - High during the day, low at night
  - High in summer, low in winter
  - No apparent interactions across 7 days of a week
- Interactions (or cross effects)
  - "Temperature" × Hour
  - "Temperature" × Month

  "Temperature": T, T^2, T^3

$$E(Load) = \beta_0 + \beta_1 \times Trend + \beta_2 \times Day \times Hour + \beta_3 \times Month + \beta_4 \times Month \times T + \beta_5 \times Month \times T^2 + \beta_6 \times Month \times T^3 + \beta_7 \times Hour \times T + \beta_8 \times Hour \times T^2 + \beta_9 \times Hour \times T^3$$

Demonstration

- SAS EG 101
- GLM
  - Simple linear regression (on trend)
  - Class variables (on hour)
  - Polynomial regression (on temperature)
  - Interaction (on hour and temperature)
  - Naïve MLR based benchmarking model

Expectation

- Every forecast is wrong
- Some are useful – *"capture the salient features"*
  - Covering business need(s)
  - Accurate
  - Interpretable
  - Documentable
  - Reproducible
  - Defensible
- Accuracy vs. defensibility

Data

1. kW, kVar, and kVA
   - ☹ $kW = kVA$; $kW^2 - kVar^2 = kVA^2$
2. Time zones & DST
   - ☹ Not checking time zones, DST, etc.
3. Missing data, redundant data, & outliers
   - ☹ Incorrect treatment to missing data, redundant data & outliers
4. T&D losses
   - ☹ Not (or incorrectly) accounting transformer/line/… losses
5. Load transfer
   - ☹ Not paying attention to load transfer

Models

6. Design
   - ☹ The newer/more complicated/etc., the better
   - ☹ Not covering business needs
7. Model building
   - ☹ Verbose models: high order polynomial; over-parameterized ANN
8. Model assumptions
   - ☹ "Sampling" times series data incorrectly
   - ☹ AR w/o unit root tests
9. Comparing models
   - ☹ Models from different utilities
   - ☹ MAPE and other measures

Decisions

10. Judgmental forecast w/o DOCUMENTATION

☹ Not documenting the justifications when alternating the forecasting results.

- Not one of the "10 ways"

☹ Fraud forecast

References

- Tao Hong, Pu Wang and H. Lee Willis, "*A Naïve Multiple Linear Regression Benchmark for Short Term Load Forecasting*", 2011 IEEE PES General Meeting
- Ott and Longnecker, "*An Introduction to Statistical Methods and Data Analysis 5ed*"

Takeaways

- Weather, human activities, and the interactions of the two drive the load profile
  - Discover, model, and analyze trend, seasonality and interactions
- Keep in mind some frequently made mistakes

Win a Free Trip to Orlando?

- SAS Analytics Conference 2011, Orlando, FL
- I won a free trip to Las Vegas two years ago ☺
  - "*Behavior Mining of Electric Load Consumption: A Regression Approach*", SAS' 12th Annual Data Mining Conference, M2009, Las Vegas, Oct 26-27, 2009
- Poster contest detail:

  http://www.sas.com/events/analytics/us/poster.html

Thank You

- Questions / comments?

Notes, data, and reading materials can be downloaded from the lecture webpage:
http://courses.drhongtao.com/dahlf