A Hybrid Data Filtering – Statistical Modeling Framework for Near-Term Forecasting

Frank A. Monforte, Ph.D.
Itron’s Forecasting Brown Bag Seminar
January 15, 2008
Please Remember

- In order to help this session run smoothly, your phones are muted.

- If you would like to make the presentation portion of the screen larger, press the expand button on the toolbar. Press it again to return to regular window.

- If you need technical assistance during the meeting, dial *0 and you will be connected to a Premiere Conferencing technician.

- If you need to give other feedback to the presenter during the meeting, such as, slow down or need to get the presenters attention for some other reason, use the pull down menu in the seating chart and we will address it right away.

- If you have general questions regarding the presentation, please type your question in the Q&A box in the bottom, right corner. We will try to answer as many questions as we can.
2008 Brown Bag Seminars

- Comparison of Load Research Expansion Methods - April 29, 2008
- Approaches to Constructing Forecast Bounds - August 26, 2008
- Using Neural Networks to Improve Regression Models - December 2, 2008

All at noon, Pacific Time
All will be recorded and available for review after the session.
Itron Forecasting Background

- Itron has been deploying forecasting solutions for over 25 years for a range of companies and government agencies.
  
  > Worldwide user base
  > Independent system and market operators
  > Regional transmission system operators
  > Retailers operating in one or more market regions
  > Wholesalers, municipalities and cooperatives
  > G&T utilities, municipalities, and cooperatives
A Hybrid Data Filtering-Statistical Modeling Framework for Near-Term Forecasting

A Hitch Hikers Guide to the Future
The Model Framework presented here is based on a collaborative effort. Special thanks go to:

> Arthur Maniaci of the New York ISO,
> Sen Li of the Midwest ISO,
> Mark Quan, David Fabiszak, Stuart McMenamin, and Jeff Fordham of Itron.
Load forecast at the 5-minute level of load resolution from 5 minutes ahead out 2 hours ahead

Forecast is to be updated every 5 minutes leveraging off real-time SCADA metering

With market restructuring nodal level forecasts are required

*Time is of the Essence*
Forecast Challenges

- Volatility in the 5 Minute SCADA Reads
- Hitting the Ramp Rates
- Hitting the Turning Points (i.e. troughs, peaks)

How do you decide how much useful information is in the most recent 5 minute read?

How useful is the most recent data read out 5 minutes, 15 minutes, 30 minutes, 60 minutes or more?
Forecasting the Morning Peak
Forecasting the Afternoon Peak
Forecasting the Afternoon Peak
Traditional Approach: Data Filtering Methods

- These methods project historical data trends into the future.
  - Kalman Filter
    - Exponential Smoothing Models
  - ARIMA Models
    - Moving Average Models
    - Polynomial Fits
**ARIMA Model**

\[ Load_t = \beta_0 + \beta_1 Load_{t-1} + \beta_2 Load_{t-2} + \ldots + \beta_j Load_{t-j} + \varepsilon_t \]

**Moving Average Model**

\[ Load_t = \frac{\sum_{j=1}^{J} Load_{t-j}}{J} \]

**Polynomial Fit**

\[ Load_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \alpha_4 t^4 \]
Kalman Filter

**Update Equations**

\[
\hat{L}_t = \hat{L}_{t-1}
\]

\[
P_t = P_{t-1} + Q_t
\]

**Measurement Equations**

\[
K_t = \frac{P_t}{P_t + R_t}
\]

\[
\hat{L}_t = \hat{L}_t + K_t (L_t - \hat{L}_t)
\]

\[
P_t = (1 - K_t)P_t
\]
Kalman Filter Example (t=1)

\[
\hat{L}_0 = 1000 \\
Q_t = Q = 10 \\
R_t = R = 10 \\
P_0 = 0.0
\]

**Update Equations**

\[
\hat{L}_1^\sim = \hat{L}_0 = 1000 \\
P_1^\sim = P_0 + Q_t = 0 + 10 = 10
\]

**Measurement Equations**

\[
K_1 = \frac{P_1^\sim}{P_1^\sim + R_1} = \frac{10}{10 + 10} \\
\hat{L}_1 = \hat{L}_1^\sim + K_1 (L_1 - \hat{L}_1^\sim) = 1000 + 0.5(1200 - 1000) = 1100 \\
P_1 = (1 - K_1)P_1^\sim = (1 - 0.5)10 = 5
\]
Kalman Filter Example (t=2)

**Update Equations**

\[ \hat{L}_2 = \hat{L}_1 = 1100 \]

\[ P_2^\sim = P_1 + Q_2 = 5 + 10 = 15 \]

**Measurement Equations**

\[ K_2 = \frac{P_2^\sim}{P_2^\sim + R_2} = \frac{15}{15 + 10} \]

\[ \hat{L}_2 = \hat{L}_2 + K_2 (L_2 - \hat{L}_2) = 1100 + 0.6(1500 - 1100) = 1340 \]

\[ P_2 = (1 - K_2)P_2^\sim = (1 - 0.6)15 = 6 \]
Data Filter Methods

- **Strengths**
  - Fast execution, can be applied to a large number of nodes
  - Highly leverages the most recent SCADA data
  - Strong 5 to 10 minutes ahead

- **Weaknesses**
  - Projects most recent trends into the future, subject to missing turning points
  - Load volatility will shadow itself into the forecast

Data Filters are akin to driving a car while looking in the rear view mirror
Statistical Model Approaches

- Generate load forecasts given forecasts of calendar, solar, and weather conditions, as well as incorporate autoregressive terms.
  - Structural Regression/Neural Network Models with or without autoregressive terms

Diagram:
- Calendar
- Weather
- Solar
- 5 Minute Actuals
- Statistical Model
- 5 Minute Forecast
Statistical Model Approaches

Strengths

> Forward looking, incorporates future conditions (i.e. weather, solar, calendar)
> Can catch future turning points/ramp rates
> Can leverage the most recent SCADA data

Statistical Models Map the Future Road
Statistical Model Approaches

- **Strengths**
  - Forward looking, incorporates future conditions (i.e. weather, solar, calendar)
  - Can catch future turning points/ramp rates
  - Can leverage the most recent SCADA data

- **Weaknesses**
  - Heavy IT footprint making it difficult to generate multiple nodal forecasts every 5 minutes
  - If autoregressive terms are included then subject to the same problem with load volatility shadowing itself into the forecast
A Hybrid Approach

- Calendar
- Weather
- Solar
- 5 Minute Actuals

5-Minute Forecast
Where We Were & Where We Are Going

The Intent is to combine Data on Where We Were with Forecasts About Where We Are Going
Fit a Polynomial to Both Information Sets

- 5 Minute Actuals
- 15 Minute Forecast
- Polynomial Forecast

Where We Were
Where We Are Going
Poly.
Final Forecast Compared with Actuals

![Graph showing forecast vs actuals for Brownbag](image)

<table>
<thead>
<tr>
<th>Time</th>
<th>Forecast</th>
<th>15-Min Forecast</th>
<th>Sim Actual</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>7.151 86</td>
<td>7.151 89</td>
<td>7.152 00</td>
<td>7.152 0</td>
</tr>
<tr>
<td>14:35</td>
<td>7.153 33</td>
<td>7.153 33</td>
<td>7.155 00</td>
<td>7.155 0</td>
</tr>
<tr>
<td>14:40</td>
<td>7.156 00</td>
<td>7.156 00</td>
<td>7.154 00</td>
<td>7.154 0</td>
</tr>
<tr>
<td>14:45</td>
<td>7.160 46</td>
<td>7.160 44</td>
<td>7.159 00</td>
<td>7.159 0</td>
</tr>
<tr>
<td>14:50</td>
<td>7.165 82</td>
<td>7.165 87</td>
<td>7.182 00</td>
<td>7.182 0</td>
</tr>
<tr>
<td>14:55</td>
<td>7.170 67</td>
<td>7.170 67</td>
<td>7.163 00</td>
<td>7.163 0</td>
</tr>
<tr>
<td>15:00</td>
<td>7.175 37</td>
<td>7.175 37</td>
<td>7.167 00</td>
<td>7.167 0</td>
</tr>
<tr>
<td>15:05</td>
<td>7.180 72</td>
<td>7.180 72</td>
<td>7.184 00</td>
<td>7.184 0</td>
</tr>
<tr>
<td>15:10</td>
<td>7.184 67</td>
<td>7.184 67</td>
<td>7.184 00</td>
<td>7.184 0</td>
</tr>
<tr>
<td>15:15</td>
<td>7.185 72</td>
<td>7.185 72</td>
<td>7.186 00</td>
<td>7.186 0</td>
</tr>
<tr>
<td>15:20</td>
<td>7.186 11</td>
<td>7.186 11</td>
<td>7.183 00</td>
<td>7.183 0</td>
</tr>
<tr>
<td>15:25</td>
<td>7.185 67</td>
<td>7.185 67</td>
<td>7.178 00</td>
<td>7.178 0</td>
</tr>
<tr>
<td>15:30</td>
<td>7.182 11</td>
<td>7.182 11</td>
<td>7.186 00</td>
<td>7.186 0</td>
</tr>
<tr>
<td>15:35</td>
<td>7.176 43</td>
<td>7.176 43</td>
<td>7.172 00</td>
<td>7.172 0</td>
</tr>
<tr>
<td>15:40</td>
<td>7.172 67</td>
<td>7.172 67</td>
<td>7.171 00</td>
<td>7.171 0</td>
</tr>
<tr>
<td>15:45</td>
<td>7.172 67</td>
<td>7.172 67</td>
<td>7.171 00</td>
<td>7.171 0</td>
</tr>
</tbody>
</table>

Simulation Date: 6/9/2006  | Clock Time: 14:50  | Step  | Reset
Questions?

Press *1 to ask a question on the phone or type in the box at the bottom, right corner.

HANDS-ON WORKSHOPS
- Fundamentals of Sales and Demand Forecasting - March 3-5 - Orlando
- Forecasting 101 - April 7-9 - Washington, DC
- Fundamentals of MetrixND - May 19-20 - Washington, DC
- Fundamentals of Short-term and Hourly Forecasting - September 17-19 - Las Vegas
- Forecasting 101 - November 12-14 - San Diego

USER MEETINGS
- European Forecasting User Meeting - 7-8 February - Brussels
- Australian Forecasting User Meeting - 12 March - Sydney
- Annual Energy Forecasting Group (EFG) Meeting - May 15-16 - Las Vegas
- 2nd Annual ISO/RTO Forecasting Summit - May 13-14 - Las Vegas
- Itron Users Conference - October 19-21 - Dallas

For more information and registration:  www.itron.com/forecastingworkshops

Contact us at:  1.800.755.9585, 1.858.724.2620 or forecasting@itron.com