Integration of Large Data Sets for Improved Decision-Making in Bulk Power Systems: Two Case Studies

Le Xie
Texas A&M University
(le.xie@tamu.edu)

PSERC Webinar
October 20, 2015
Presentation Outline

• Brief Overview of PSERC T-51

• *Spatio-temporal* Correlations among Data Sets

• Case Study 1: Renewable Forecast

• Case Study 2: Synchrophasor dimensionality reduction and early anomaly detection

• Concluding Remarks
T-51 Project Objectives

• Improved decision-making by utilization of large data sets – “Big Data”

• The integration of emerging large-data sets to support advanced analytics, to enhance electricity system management (planning, operations and control)

• Correlation of data in time and space and assuring consistent data semantic and syntax
Big Data in Bulk Power Systems: Opportunities and Challenges

T-51 Project Summary

- Project duration: 2013-2015, final report available at PSERC
- Use of Big Data for Outage and Asset Management (Kezunovic, lead PI)
- Distributed Database for Future Grid (Grijalva and Chau)
- Spatio-Temporal Analytics for Renewables (Xie)
  - Wind power prediction and its economic benefits
  - Solar power prediction and quantified economic benefits
In 2014, renewable energy sources account for 16.28% of total installed U.S. operating generating capacity.

Solar, wind, biomass, geothermal, and hydropower provided 55.7% of new installed U.S. electrical generating capacity during the first half of 2014 (1,965 MW of the 3,529 MW total installed).

http://www.renewableenergyworld.com
http://www.eia.gov/
http://www.triplepundit.com/
Growth of Synchrophasors (PMU)

North America

Reported by NASPI*

• By March 2012, 500 networked PMUs installed.
• >1700 PMUs installed by 2015.

China

• More than 2000 PMU [Beijing Sifang, 2013].

PMU map in North America as of Oct. 2014.


• http://www.eia.gov/todayinenergy/detail.cfm?id=5630
Spatio-temporal Correlations at Multi Scales

Renewable Data

Improved Forecast & Dispatch
(5 minutes to hours ahead)
Spatio-temporal Correlations at Multi-scale

Early Anomaly Detection & Mitigation (milliseconds to seconds)
Wind Variability and Spatial Correlation

Total California Wind Generation

- Source: http://www.caiso.com/1c9b/1c9bd3a394f0.pdf
Communication and Information exchange:

- Produce good wind generation forecast
- Reduce the system-wide dispatch cost
- Lower the ancillary services cost
- Incorporate more wind generation

Existing Methods

• Persistence (PSS) Model
  • Assume the future wind speed is the same as the current one:
    \[ \hat{y}_{s,t+k} = y_{s,t} \]

• Autoregressive (AR) Model:
  • Estimate \( \mu_{s,t+k}^r \) as a linear combination of the previous wind speed at the same location
    \[ \mu_{s,t+k}^r = \alpha_0 + \sum_{i=0}^{p} \beta_i \mu_{s,t-i}^r \]

where \( \mu_{s,t+k}^r \) is the residue term of center parameter of wind speed.
Spatio-Temporal Wind Speed Forecast

• The scale parameter $\sigma_{s,t+k}$ is modeled as

$$\sigma_{s,t+k} = b_0 + b_1 v_{s,t}$$

• Where $b_0, b_1 > 0$ and $v_{s1,t}$ is the volatility value:

$$v_{s,t} = \sqrt{\frac{1}{2S} \sum_{s=1}^{S} \sum_{i=0}^{1} \left( \mu_{s,t-i}^r - \mu_{s,t-i-1}^r \right)^2}$$

• The residual term modeled as a linear function of current and past (up to time lag $h$) wind speed and trigonometric functions of wind direction.

$$\mu_{s,t+k}^r = \alpha_0 + \sum_{s=1}^{S} \sum_{j=0}^{P} \alpha_{s,j} \mu_{s,t-j}^r$$

$$+ \sum_{s=1}^{S} \sum_{j=0}^{P} \beta_{s,j} \cos(\theta_{s,t-j}^r) + \sum_{s=1}^{S} \sum_{j=0}^{P} \gamma_{s,j} \sin(\theta_{s,t-j}^r)$$
Spatio-Temporal Wind Forecast
West Texas Case Study [1]

### Forecast Model Performance

**MAE VALUES OF THE 10-MINUTE-AHEAD, 20-MINUTE-AHEAD AND UP TO 1-HOUR-AHEAD FORECASTS ON 11 DAYS’ IN 2010 FROM THE PSS, AR, TDD AND TDDGW MODELS AT THE FOUR LOCATIONS (SMALLEST IN BOLD)**

<table>
<thead>
<tr>
<th>Location</th>
<th>Model</th>
<th>10 min</th>
<th>20 min</th>
<th>30 min</th>
<th>40 min</th>
<th>50 min</th>
<th>60 min</th>
</tr>
</thead>
<tbody>
<tr>
<td>PICT</td>
<td>PSS</td>
<td>0.56</td>
<td>0.72</td>
<td>0.84</td>
<td>0.92</td>
<td>1.00</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>0.55</td>
<td>0.70</td>
<td>0.80</td>
<td>0.87</td>
<td>0.94</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>TDD</td>
<td>0.54</td>
<td>0.68</td>
<td>0.77</td>
<td>0.84</td>
<td>0.90</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>TDDGW</td>
<td>0.54</td>
<td>0.68</td>
<td>0.77</td>
<td>0.83</td>
<td>0.89</td>
<td>0.94</td>
</tr>
<tr>
<td>JAYT</td>
<td>PSS</td>
<td>0.50</td>
<td>0.63</td>
<td>0.71</td>
<td>0.78</td>
<td>0.83</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>0.48</td>
<td>0.60</td>
<td>0.68</td>
<td>0.75</td>
<td>0.8</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>TDD</td>
<td>0.47</td>
<td>0.57</td>
<td>0.64</td>
<td>0.69</td>
<td>0.73</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>TDDGW</td>
<td>0.47</td>
<td>0.57</td>
<td>0.64</td>
<td>0.68</td>
<td>0.71</td>
<td>0.75</td>
</tr>
<tr>
<td>SPUR</td>
<td>PSS</td>
<td>0.51</td>
<td>0.64</td>
<td>0.73</td>
<td>0.81</td>
<td>0.86</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>0.49</td>
<td>0.61</td>
<td>0.69</td>
<td>0.76</td>
<td>0.80</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>TDD</td>
<td>0.48</td>
<td>0.59</td>
<td>0.67</td>
<td>0.72</td>
<td>0.76</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>TDDGW</td>
<td>0.49</td>
<td>0.59</td>
<td>0.67</td>
<td>0.71</td>
<td>0.75</td>
<td>0.79</td>
</tr>
<tr>
<td>ROAR</td>
<td>PSS</td>
<td>0.55</td>
<td>0.71</td>
<td>0.82</td>
<td>0.92</td>
<td>0.98</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>0.54</td>
<td>0.68</td>
<td>0.78</td>
<td>0.86</td>
<td>0.92</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>TDD</td>
<td>0.54</td>
<td>0.67</td>
<td>0.77</td>
<td>0.85</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>TDDGW</td>
<td>0.54</td>
<td>0.67</td>
<td>0.76</td>
<td>0.82</td>
<td>0.87</td>
<td>0.90</td>
</tr>
</tbody>
</table>
Wind Generation Forecast Distribution

One Hour-ahead Wind Generation Forecast Uncertainty

Hour ahead Wind generation forecast uncertainties of Jayton (JAYT), Texas under various days
Total Operating Cost
Spatio-temporal Solar Forecast [2]

ARX Model for Solar Irradiance Forecast [2]

\[
y[t] = f(y[t-1], \ldots, y[t-n], u_1[t-d_1], \ldots, u_1[t-d_1-m_1+1], \ldots, u_i[t-d_i], \ldots, u_i[t-d_i-m_i+1])
\]

We compared our model (ST) to persistence (PSS), auto regression (AR), and back-propagation neural network (BPNN) forecast models.

## Results [2]

### Performance for 1 hour ahead

<table>
<thead>
<tr>
<th>Case</th>
<th>Number</th>
<th>Training Period</th>
<th>Validation Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>January, March, May, July,</td>
<td>February, April</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>April, June</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>September,</td>
<td>June, August</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>November</td>
<td>August, October</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Case 1</th>
<th></th>
<th>Case 2</th>
<th></th>
<th>Case 3</th>
<th></th>
<th>Case 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>ST</td>
<td>AR</td>
<td>BPNN</td>
<td>PSS</td>
<td>ST</td>
<td>AR</td>
<td>BPNN</td>
<td>PSS</td>
</tr>
<tr>
<td>MAE</td>
<td>58.3</td>
<td>62.0</td>
<td>99.1</td>
<td>92.7</td>
<td>52.1</td>
<td>54.4</td>
<td>97.8</td>
<td>102.0</td>
</tr>
<tr>
<td>RMSE</td>
<td>81.5</td>
<td>87.0</td>
<td>137.6</td>
<td>116.2</td>
<td>78.3</td>
<td>84.0</td>
<td>139.3</td>
<td>124.8</td>
</tr>
</tbody>
</table>

### Performance for 2 hour ahead

<table>
<thead>
<tr>
<th></th>
<th>Case 1</th>
<th></th>
<th>Case 2</th>
<th></th>
<th>Case 3</th>
<th></th>
<th>Case 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>ST</td>
<td>AR</td>
<td>BPNN</td>
<td>PSS</td>
<td>ST</td>
<td>AR</td>
<td>BPNN</td>
<td>PSS</td>
</tr>
<tr>
<td>MAE</td>
<td>100.0</td>
<td>107.9</td>
<td>148.1</td>
<td>160.3</td>
<td>83.4</td>
<td>90.5</td>
<td>145.9</td>
<td>173.5</td>
</tr>
<tr>
<td>RMSE</td>
<td>136.3</td>
<td>143.7</td>
<td>198.7</td>
<td>200.3</td>
<td>122.9</td>
<td>132.2</td>
<td>200.0</td>
<td>207.3</td>
</tr>
</tbody>
</table>

Shorter time-scale: Spatio-temporal is worse than PSS
Part I: Summary

- Spatio-temporal correlation among renewable generation sites (wind and photovoltaic) could be leveraged for improved near-term forecast.
- The economic benefit from spatio-temporal forecast vary at different time scales.
- Possible extensions:
  - Large ramp forecast
  - Distributed PV forecast
Barriers to Using PMU for Real-time Operations

Large sets of PMU data

- Efficient real-time analysis
- Data storage
- Missing data prediction

Dimensionality Reduction

Related Work:
Raw PMU Data from Texas

No system topology, no system model.
Total number of PMUs: 7.
Total number of PMUs: 14 for frequency analysis
8 for voltage magnitude analysis.
PCA for Texas Data

\[ Y_B^\omega = \begin{bmatrix} \omega_5, \omega_3 \end{bmatrix} \]

\[ Y_B^V = \begin{bmatrix} V_2, V_1, V_7 \end{bmatrix} \]

Cumulative variance for bus frequency and voltage magnitude for Texas data.
PCA for Eastern Interconnection

\[ Y_B^\omega = [\omega_{11}, \omega_6] \]

\[ Y_B^V = [V_8, V_5, V_2] \]

Cumulative variance for bus frequency and voltage magnitude for PJM data.
Scatter Plot of Bus Frequency

2D Scatter plot for bus frequency.

3D Scatter plot for bus frequency.
Scatter Plot of Voltage Magnitude

2D Scatter plot for voltage magnitude.

3D Scatter plot for voltage magnitude.

Normal Condition  |  Abnormal Condition  |  Back to Normal Condition
Observations

• High dimensional PMU raw measurement data lie in an much lower subspace (even with linear PCA)

• Scattered plots suggest that
  Change of subspace -> Occurrence of events !

• But, what is the way to implement it?

• Is there any theoretical justification?

  Data-driven subspace change ⇔ Indication of physical events in wide-area power systems
Early Event Detection

- **Synchrophasor Data Dimensionality Reduction**
  - Corporate PDC
  - Local PDC

- **Data Storage**
  - Corporate PDC
  - Local PDC

- **Pilot PMUs**

**Key Terms**:
- PMU: Phasor measurement unit
- PDC: Phasor data concentrator
- Preprocessed PMU data
- Raw measured PMU data
Dimensionality Reduction Algorithm [3]

1. PMU Data Collection: Measurement matrix

\[ Y_{n \times N} = [y^{(1)}, \ldots, y^{(N)}], \quad y^{(i)} = [y_1^{(i)}, \ldots, y_n^{(i)}]^T \]

\( N \) measurements. Each has \( n \) samples in the time history.

2. PCA-based Dimensionality Reduction

(1) Eigenvalues and eigenvectors of covariance matrix \( \text{Cov}(Y) \).
(2) Rearrange eigenvalues in decreasing order to find principal components (PCs).
(3) Select top \( m \) out of \( N \) PCs based on predefined threshold.
(4) Project original \( N \) variables in the \( m \) PC-formed new space.
(5) \( m' \) variables are kept and chosen as orthogonal to each other as possible.
(6) Basis matrix \( Y_B := [y_b^{(1)}, \ldots, y_b^{(m')}]. \subseteq Y \)
(7) Predict \( y^{(i)} \) in terms of \( Y_B \)

\[ \hat{y}^{(i)} \approx \sum_{j=1}^{m'} v_j^{(i)} y_b^{(j)} = Y_B v^{(i)}. \quad \quad v^{(i)} := (Y_B^T Y_B)^{-1} Y_B^T y^{(i)}. \]

PCA for Bus Frequency

Cumulative variance for bus frequency.

2D Loading plot for bus frequency.

2 PCs. Basis matrix $Y_B^{\omega} = \begin{bmatrix} \omega_{206} & \omega_{102} \end{bmatrix}$
Possible Implementation

Early Event Detection Algorithm

Adaptive Training
PCA-based Dimensionality Reduction
- PMU Measurement $Y_{n \times N}(t_0)$
- Covariance Matrix $C_Y$
- Reorder $N$ Eigenvalues
- Select $m$ PCs, $m \ll N$
- Project $Y$ in $m$-D Space
- Define Base Matrix $Y_B$
- Calculate $v^{(i)}$

Robust Online Monitoring
Online Detection
- Approximate $\hat{y}(t)^{(i)}$
- Approximation error $e(t)^{(i)}$
- Event indicator $\eta(t)^{(i)}$
- $\eta(t)^{(i)} \geq \gamma$?
- $T_{t-t_0} \geq T_U$?
- Alert to System Operators
- NEED
- Update
- $t = t + 1$

Update
Theorem for Early Event Detection

Using the proposed event indicator $\eta$, a system event can be detected within 2-3 samples of PMUs, i.e., within 100 ms, whenever for some selected non-pilot PMU $i$, the event indicator satisfies

$$\left| \eta(t)^{(i)} \right| \geq \gamma$$

where $\gamma$ is a system-dependent threshold and can be calculated using historical PMU data.

**Theoretically justified.**

$$\eta(t)^{(i)} := \frac{e(t)^{(i)}}{e_{normal}} \quad e(t)^{(i)} := \left| \frac{\tilde{y}(t)^{(i)}}{y(t)^{(i)}_{meas}} \right| \times 100\%.$$
Sketch of the Proof

- **Power system DAE model**
  \[
  \dot{x}(t) = f(x(t), u(t), h(t), q(t)), \\
  0 = g(x(t), u(t), h(t), q(t)), \\
  \]

- **Discretization**
  \[
  x[k + 1] = A_dx[k] + B_du[k] + \alpha[k], \\
  y[k] = C_dx[k] + D_du[k] + \varepsilon[k], \\
  \]

- **Using back substitution, explicitly express output (measurement) \( y[k] \) in terms of initial condition \( x[1] \), control input \( u[k] \), noise \( e[k] \)**
  \[
  y[k] = C(e^{AT})^{k-1}x[1] + \sum_{l=1}^{k-1} C(e^{AT})^{l-1}A^{-1}(e^{AT} - I)Bu[k - l] + \varepsilon[k] \\
  = y_x[k] + y_u[k] + y_\varepsilon[k], \\
  \]
Sketch of the Proof (cont.’d)

• Normal conditions: training errors are small

\[
[c_x(i) - \sum_{j=1}^{n_l} v_j^{(i)} c_x^{(j)}] x[1] + [y_e(i) - \sum_{j=1}^{n_l} v_j^{(i)} y_e^{(j)}] + [c_u(i) - \sum_{j=1}^{n_l} v_j^{(i)} c_u^{(j)}] U_0
= \Delta c_x x[1] + \Delta y_e + \Delta c_u U_0 \approx 0.
\]

• \(U_0\) and \(x[1]\) can be theoretically calculated by TRAINING data.

• Any changes in control inputs and initial conditions will lead to large prediction error.

• If system topology changes, \(\Delta c_x\) and \(\Delta c_u\) will change, resulting in a large prediction error.
Case Study 1: PSS/E Data

- 23-bus system
- 23 PMUs.
- Outputs of PMUs: $\omega$, $V$.

## Oscillation Event

The figure illustrates the time evolution of oscillation during a training stage. The key events are marked on the time axis, with each event accompanied by the corresponding time points.

### Table of Events

<table>
<thead>
<tr>
<th>Time</th>
<th>Sampling Points</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-100s</td>
<td>1-3000</td>
<td>Normal Condition</td>
</tr>
<tr>
<td>100.03-150s</td>
<td>3001-45000</td>
<td>Bus Disconnection (206)</td>
</tr>
<tr>
<td>150.03-250s</td>
<td>4501-7500</td>
<td>Voltage set point changes (211)</td>
</tr>
</tbody>
</table>

The figure shows a point at 250 seconds indicating a change in voltage set point. The corresponding change in voltage reference is marked as $\Delta V_{ref}^{211} = -0.1$. The diagram also highlights the training stage starting at -250 seconds and ending at 0 seconds, with bus 206 disconnected at 100 seconds.
Early Event Detection

\[ \omega_{152} \] profile.\n\[ \eta_{152}^{\omega} \] during line tripping event.
Early Event Detection

• How EARLY is our algorithm?
  Our Method: potentially within a few samples (<0.1 seconds)

• Most Oscillation monitoring system (OMS) needs 10 sec to detect the oscillation.

Case Study 2: Unit Tripping Event

• No system topology, no system model.
• Total number of PMUs: 7.
• 2 unit tripping events.
• Sampling rate: 30 Hz.
PCA for Bus Frequency and Voltage Magnitude

$Y_B^\omega = [\omega_5, \omega_3]$

(a) Cumulative Variance for Bus Frequency in Texas Data

Cumulative variance for bus frequency and voltage magnitude for Texas data.
Early Event Detection

(a) $w_4$ Profile During Unit Tripping Events

(b) Zoomed-in $w_4$ Profile During 1st Unit Tripping Event

(c) Zoomed-in $w_4$ Profile During 2nd Unit Tripping Event

$w_4$ profile.

(a) $\eta_4^o$ During Unit Tripping Events

(b) Zoomed-In $\eta_4^o$ During 1st Unit Tripping Event

(c) Zoomed-In $\eta_4^o$ During 2nd Unit Tripping Event

$\eta_4^o$ during unit tripping event.
Part II: Summary

• Large-scale PMU data can be reduced to a space with much lower dimensionality (surprisingly well).

• Change of dimensionality could be leveraged for novel early anomaly detection

• Rich physical insights can be obtained from PMU data

• Possible extensions:
  • Event classification, specification and localization
  • Online PMU bad data processing

Concluding Remarks

• We investigated the integration of large data sets for improved grid operations:
  – Spatio-temporal analytics for improved renewable forecast & market operations
  – Dimensionality reduction of synchrophasor data for improved monitoring and anomaly detection

• Many open questions
  – Streaming data quality
  – Data analytics integration with EMS, DMS, MMS
  – How to teach “data sciences” for power systems? A first attempt in Fall 15 at TAMU
Acknowledgements

• Support from PSERC and BPA
• T-51 Team (Profs. M. Kezunovic, S. Grijalva, P. Chau) and Industry Advisors
• Collaborators:
  • Prof. Marc G. Genton (KAUST, Spatio-temporal Statistics)
  • Prof. P. R. Kumar (Texas A&M, System Theory)
• Students:
  • Yingzhong Gary Gu (GE)
  • Yang Chen (PJM)
  • Anupam Thatte (MISO)
  • Chen Nathan Yang (NYISO)
  • Meng Wu
Key References


Thank You!

Le Xie
Associate Professor
Department of Electrical and Computer Engineering
Texas A&M University
Le.xie@tamu.edu
www.ece.tamu.edu/~lxie