Scenario Generation for Two-Stage Stochastic Economic Dispatch

I. Satkauskas, M. Reynolds, D. Sigler, J. Maack, W. Jones
FERC Technical Conference, June 27, 2019
Table of Contents

• Problem formulation
• WIND Toolkit data
• Importance sampling
• Experimental results
Problem Formulation
The Exascale Computing Project (ECP) was established to deliver exascale-ready applications and solutions that address currently intractable problems of strategic importance and national interest. 

**ECP Applications Target National Problems:**

- National security, energy security (e.g. ExaWind)
- Scientific discovery, earth system, healthcare
- Economic security
  - ExaSGD: Reliable and efficient planning of the power grid

https://www.exascaleproject.org
Two-stage linear stochastic programming problem

\[
\min_{x} \quad c^T x + \mathbb{E}_\xi [L(x, \xi)] \\
\text{s. t.} \quad Ax \leq b
\]

where \(L(x, \xi)\) is the optimal value of the second-stage problem,

\[
\min_{y} \quad q_\xi^T y \\
\text{s. t.} \quad T_\xi x + W_\xi y \leq h_\xi, \ y \geq 0
\]

and

- \(x\) - first stage decision vector (e.g. thermal dispatch)
- \(y\) - second stage decision vector (e.g. wind dispatch, wind spilled, load shed)
- \(\xi\) - uncertain data (e.g. deviation in wind power from forecast).
Cost functions

First stage:
- thermal generation costs plus expectation of second stage costs

\[
\min_x \sum_{g \in G} c_g x_g + \mathbb{E}_\xi [L(x, \xi)]
\]

Second stage:
- Wind generation, spilling wind, and slack variable costs (i.e. overload and loss-of-load) at buses \(i \in \mathcal{I}\) and wind plants \(w \in \mathcal{W}\)

\[
\min_{y^+, y^-, \omega, \omega^{spl}} \sum_{w \in \mathcal{W}} (c_w \omega_w + c_w^{spl} \omega^{spl}_w) + \sum_{i \in \mathcal{I}} (c_i^+ y_i^+ + c_i^- y_i^-)
\]
Constraints

First stage:

- constraints on output of thermal generators $g \in G$

\[ x_g^{\text{min}} \leq x_g \leq x_g^{\text{max}} \quad \forall g \]

- ramping constraints on thermal generators

\[ -R_g^{\text{down}} \leq x_g - l_g \leq R_g^{\text{up}} \quad \forall g \]

Second stage:

- slack variable (loss-of-load and overload) constraints on buses

\[ 0 \leq y_i^\pm \quad \forall i \in I \]

- wind power constraints

\[ 0 \leq \omega_w \leq \omega_w^{\text{fcst}} + \xi_w \quad \forall w \]

\[ \omega_w^{\text{spl}} = (\omega_w^{\text{fcst}} + \xi_w) - (\omega_w) \quad \forall w \]
• power balance constraints at every node $i$

$$y_i^+ - y_i^- + \sum_{w \in W_i} \omega_w + \sum_{g \in G_i} x_g = \sum_{q \in D_i} d_q - \sum_{e=(j,i)} f_e + \sum_{e=(i,j)} f_e$$

• line flow constraints

$$F_e \leq f_e \leq \overline{F}_e \quad \forall e \in \mathcal{E}$$

• power flow physics, e.g. ACOPF and DCOPF (below)

$$B_e (\theta_i - \theta_j) - f_e = 0 \quad \forall e \in \mathcal{E}$$
Solving the SAA: matrix structure

\[
\min_{x, y_1, \ldots, y_N} \left( c^T x \right) + \frac{1}{N} \sum_{s=1}^{N} \left( c_{x_s}^T y_s \right)
\]

such that

\[
Ax + \sum_{s=1}^{N} \left( T_{x_s} x + W_{x_s} y_s \right) = h
\]

\[
T_{x_1} x + W_{x_1} y_1 = h_1
\]

\[
T_{x_2} x + W_{x_2} y_2 = h_2
\]

\[
T_{x_3} x + \cdots = h
\]

\[
T_{x_N} x + W_{x_N} y_N = h_N
\]

\[
x \geq 0, \ y_1 \geq 0, \ y_2 \geq 0, \ldots, \ y_N \geq 0.
\]

There are specialized algorithms for solving these optimization problems:

- Schur complement approaches to solving updates in interior point methods, e.g. PIPS\(^1\).
- Progressive hedging algorithm, (e.g. http://www.pyomo.org/)

\(^1\)https://github.com/Argonne-National-Laboratory/PIPS
Uncertainty

• In this problem formulation the uncertainty is included by taking \( \mathbb{E}_{\xi} [L(x, \xi)] \), that we approximate by \( \frac{1}{N} \sum_{s=1}^{N} L(x, \xi_s) \)

• Since we are solving economic dispatch problem for 5-minute periods, \( \xi \) represents power deviation from persistence.

• \( \xi = (\xi_1, \xi_2, \ldots, \xi_m) \), where \( \xi_i \) is power deviation at wind farm \( i \).

• We use high-fidelity data to produce realistic wind scenarios that respect physics and spatio-temporal relations.
WIND Toolkit\(^2\): domains

Figure: WRF simulation domains. (from “Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit”)

Wind Integration National Dataset (WIND) Toolkit

- Data sets: 2-TB, 50-TB, 0.5 PB
- 2 km x 2 km grid with 20 m vertical resolution.
- 5 min time resolution: pressure, wind speed, direction, humidity, temperature, and density.
- Techno-Economic data set: 5 min time series at 120 000 wind sites.

\(^2\)https://www.nrel.gov/grid/wind-toolkit.html
Figure: Power curves used to convert modeled wind speed to power.

- Power curves were chosen according to the estimated long-term wind conditions at each site.
- Each turbine is assumed to have a rated power of 2 MW (at 100 m)
TAMU\(^3\) 200 grid (Illinois)

Figure: TAMU 200 bus grid.

- Wind capacity 760 MW (19.42%).
- 6 wind farms, 50 wind sites (NREL Wind Toolkit).

\(^3\)https://electricgrids.engr.tamu.edu/electric-grid-test-cases/
- Wind capacity 760 MW (19.42%).
- 6 wind farms, 50 wind sites (NREL Wind Toolkit).
- 437 real turbines (USWTDB\textsuperscript{4}).

\textsuperscript{4}https://eerscmap.usgs.gov/uswtdb/
TAMU 2000 grid (Texas)

- 2000 buses
- Wind capacity 12574 MW (12.56%).
- 87 wind farms, 859 wind sites (NREL Wind Toolkit).
TAMU 2000 grid

- 2000 buses
- Wind capacity 12574 MW (12.56%).
- 87 wind farms, 859 wind sites (NREL Wind Toolkit).
- 7312 real turbines (USWTDB).
TAMU 2000 grid

- 2000 buses
- Wind capacity 12574 MW (12.56%).
- 87 wind farms, 859 wind sites (NREL Wind Toolkit).
- 7312 real turbines (USWTDB).
• 7 years (2007-2013) of power output data at 5-min resolution.
• One year we leave as “actuals” for experiments; 6 years are used to generate scenarios $\xi$.
• For each wind farm on the grid we aggregate WIND Toolkit wind sites up to farm’s capacity.
• Additionally, we split our data set based on total wind power in the network: low, medium, high.
TAMU 200 grid

Figure: TAMU 200 bus grid.

- Power (760 MW) conditioning splits data into 3 sets:
  1. low < 11MW
  2. 11MW < medium < 717MW
  3. high > 717MW
Distribution of deviations: TAMU 200

Figure: Power conditioned distributions of deviations
Examples of scenario tables for TAMU 200 grid

### Table: Low power

<table>
<thead>
<tr>
<th>IssueTime</th>
<th>65</th>
<th>104</th>
<th>105</th>
<th>114</th>
<th>115</th>
<th>147</th>
<th>TotalPower</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008-01-01 02:00:00</td>
<td>0.175488</td>
<td>-0.245543</td>
<td>-0.385107</td>
<td>0.025132</td>
<td>0.728723</td>
<td>1.064814</td>
<td>695.268688</td>
<td>1.364507</td>
</tr>
<tr>
<td>2008-01-01 02:05:00</td>
<td>0.041986</td>
<td>-0.255817</td>
<td>-0.464536</td>
<td>-0.003527</td>
<td>0.049244</td>
<td>0.964228</td>
<td>696.633194</td>
<td>0.331578</td>
</tr>
<tr>
<td>2008-01-01 02:10:00</td>
<td>-0.237243</td>
<td>0.049380</td>
<td>-0.520757</td>
<td>-0.022131</td>
<td>-0.759113</td>
<td>1.072492</td>
<td>696.964772</td>
<td>-0.417391</td>
</tr>
<tr>
<td>2008-01-01 02:15:00</td>
<td>-0.755391</td>
<td>0.021509</td>
<td>-0.258561</td>
<td>-0.031425</td>
<td>-0.657219</td>
<td>0.746616</td>
<td>696.547381</td>
<td>-0.934471</td>
</tr>
<tr>
<td>2008-01-01 02:20:00</td>
<td>-1.745687</td>
<td>-0.327942</td>
<td>-0.588913</td>
<td>-0.033362</td>
<td>-1.146570</td>
<td>-0.352600</td>
<td>695.612909</td>
<td>-4.195075</td>
</tr>
</tbody>
</table>

### Table: Medium power

<table>
<thead>
<tr>
<th>IssueTime</th>
<th>65</th>
<th>104</th>
<th>105</th>
<th>114</th>
<th>115</th>
<th>147</th>
<th>TotalPower</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008-01-01 09:40:00</td>
<td>-0.415083</td>
<td>-0.164326</td>
<td>-0.331099</td>
<td>-0.032976</td>
<td>-0.451123</td>
<td>-0.037300</td>
<td>10.955064</td>
<td>-1.431906</td>
</tr>
<tr>
<td>2008-01-01 09:45:00</td>
<td>-0.524133</td>
<td>-0.146091</td>
<td>-0.220279</td>
<td>-0.032994</td>
<td>-0.329175</td>
<td>-0.043634</td>
<td>9.523158</td>
<td>-1.296306</td>
</tr>
<tr>
<td>2008-01-01 09:50:00</td>
<td>-0.290003</td>
<td>-0.071754</td>
<td>-0.305224</td>
<td>-0.032987</td>
<td>-0.300771</td>
<td>-0.035274</td>
<td>8.226852</td>
<td>-1.036013</td>
</tr>
<tr>
<td>2008-01-01 09:55:00</td>
<td>-0.213800</td>
<td>-0.042937</td>
<td>-0.225679</td>
<td>-0.015872</td>
<td>-0.246981</td>
<td>-0.035779</td>
<td>7.190839</td>
<td>-0.781049</td>
</tr>
<tr>
<td>2008-01-01 10:00:00</td>
<td>-0.305855</td>
<td>-0.037831</td>
<td>-0.086708</td>
<td>0.004084</td>
<td>-0.036173</td>
<td>-0.029725</td>
<td>6.409790</td>
<td>-0.492209</td>
</tr>
</tbody>
</table>
Distribution of deviations: TAMU 200

Figure: Low power deviations
Distribution of deviations: TAMU 2000

Figure: Medium power deviations
Basic idea

Importance sampling, is based on the following idea:

$$\mathbb{E}_p \left[ L(\xi) \right] = \int L(\xi) p(\xi) \, d\xi = \int L(\xi) \frac{p(\xi)}{q(\xi)} q(\xi) \, d\xi = \mathbb{E}_q \left[ L(\xi) R(\xi) \right],$$

where $R(\xi) = \frac{p(\xi)}{q(\xi)}$, $p(\xi)$ is called nominal and $q(\xi)$ is called importance distribution.
Basic idea

Low fidelity approach: surrogate model for $L(x, \xi)$ using loss-of-load and costs of spilling wind

$$\tilde{L}(\xi) = \begin{cases} 
  c^- \sum_w \xi_w & \sum_w \xi_w > 0 \\
  c^+ \sum_w \xi_w & \sum_w \xi_w \leq 0 
\end{cases}$$

where $c^-$ is cost of spilling wind, and $c^+$ is the cost of loss-of-load.
Nominal and importance distribution

Figure: Histograms of deviations (TAMU 200 grid, medium power)
Example scenarios: nominal

Figure: TAMU 200 medium power scenarios

Figure: TAMU 200 scenarios drawn from nominal distribution
Example scenarios: importance

<table>
<thead>
<tr>
<th>IssueTime</th>
<th>65</th>
<th>104</th>
<th>105</th>
<th>114</th>
<th>115</th>
<th>147</th>
<th>TotalPower</th>
<th>Deviation</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012-10-12 14:25:00</td>
<td>-0.886548</td>
<td>-0.656891</td>
<td>-1.159231</td>
<td>0.009730</td>
<td>-0.049452</td>
<td>-0.646495</td>
<td>88.249340</td>
<td>-3.388887</td>
<td>1.855446e-06</td>
</tr>
<tr>
<td>2013-02-03 15:30:00</td>
<td>-0.293306</td>
<td>2.029175</td>
<td>3.345293</td>
<td>-0.379425</td>
<td>-2.351285</td>
<td>0.460209</td>
<td>219.020447</td>
<td>2.830860</td>
<td>4.649437e-12</td>
</tr>
<tr>
<td>2013-02-22 15:35:00</td>
<td>-3.879552</td>
<td>-0.951475</td>
<td>-3.166533</td>
<td>-0.138267</td>
<td>-2.658777</td>
<td>-1.502286</td>
<td>520.329205</td>
<td>-17.299900</td>
<td>9.470211e-06</td>
</tr>
<tr>
<td>2012-08-17 15:00:00</td>
<td>0.527059</td>
<td>-0.888020</td>
<td>-0.662040</td>
<td>0.035835</td>
<td>-1.059429</td>
<td>-0.085468</td>
<td>31.279223</td>
<td>-2.152098</td>
<td>1.178277e-06</td>
</tr>
<tr>
<td>2011-10-10 16:35:00</td>
<td>-0.048058</td>
<td>0.276023</td>
<td>-0.153408</td>
<td>-0.012902</td>
<td>-0.095665</td>
<td>0.554852</td>
<td>76.247409</td>
<td>0.521732</td>
<td>8.568582e-13</td>
</tr>
</tbody>
</table>

Figure: TAMU 200 medium power table

<table>
<thead>
<tr>
<th>IssueTime</th>
<th>ScenarioNr</th>
<th>65</th>
<th>104</th>
<th>105</th>
<th>114</th>
<th>115</th>
<th>147</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007-07-04 02:00:00</td>
<td>1</td>
<td>-2.11405</td>
<td>-0.559398</td>
<td>-2.50884</td>
<td>-0.0236497</td>
<td>-1.83873</td>
<td>-9.94101</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-0.23966</td>
<td>-0.0533235</td>
<td>-2.34025</td>
<td>-0.20580</td>
<td>-0.538439</td>
<td>-1.5048</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-8.84335</td>
<td>-1.82405</td>
<td>-5.94129</td>
<td>-0.373219</td>
<td>-0.947498</td>
<td>-5.18819</td>
</tr>
<tr>
<td>2007-07-04 02:05:00</td>
<td>1</td>
<td>-9.67678</td>
<td>-7.37214</td>
<td>-14.9094</td>
<td>-0.422604</td>
<td>-5.41246</td>
<td>8.31575</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.411175</td>
<td>-0.33989</td>
<td>-0.030612</td>
<td>0.101074</td>
<td>0.743205</td>
<td>-1.95942</td>
</tr>
</tbody>
</table>

Figure: TAMU 200 scenrios drawn from importance distribution
Experimental Results
One week economic dispatch for TAMU 200 grid

First stage costs using 20 samples

Second stage costs using 20 samples
One week economic dispatch for TAMU 200 grid

Second stage costs using 20 samples

Second stage costs using 40 samples
One week economic dispatch for TAMU 200 grid

<table>
<thead>
<tr>
<th>Sampling method</th>
<th>First stage costs [$]</th>
<th>Second stage costs [$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>1,180,299</td>
<td>949,091</td>
</tr>
<tr>
<td>IS</td>
<td>1,209,337</td>
<td>124,102</td>
</tr>
</tbody>
</table>

Figure: 20 samples

<table>
<thead>
<tr>
<th>Sampling method</th>
<th>First stage costs [$]</th>
<th>Second stage costs [$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>1,185,960</td>
<td>614,669</td>
</tr>
<tr>
<td>IS</td>
<td>1,218,600</td>
<td>26,748</td>
</tr>
</tbody>
</table>

Figure: 40 samples
Future work

- High fidelity surrogate model for $L(x, \xi)$.
- Extend to multi-period scenarios.
- More realistic physics, e.g. ACOPF.
- Extending toolset to data sets other than WIND toolkit.
Thank you. Questions?

- TAMU Synthetic Grids [https://electricgrids.engr.tamu.edu/electric-grid-test-cases/](https://electricgrids.engr.tamu.edu/electric-grid-test-cases/)
- USWTDB [https://eerscmap.usgs.gov/uswtdb/](https://eerscmap.usgs.gov/uswtdb/)