Multi-Area Reliability Assessment with Variable Energy Resources and Optimal Importance Sampling based on Monte Carlo Markov Chains

IEEE LOLE
- August 2019 –

- Daniela Bayma de Almeida
- Gerson Couto
- Carmen Tancredo Borges
Electrical System Planning

- **Motivation:** The worldwide insertion of variable renewable resources (VRE) such as solar and wind made supply reliability evaluation even more significant and increased the complexity of planning studies.

- **Planning process**

  1. **Investment Module**
     - Investment Cost
     - Minimize
     - Expands Plan candidate

  2. **Operating Module**
     - Expected Value of Operating Cost

- **Reliability Module**
  - Ensures that the expansion planning meets a reliability requirement

- Methodological challenges for reliability evaluation with VREs:
  - Use of stochastic models to produce VRE scenarios
  - Spatial and temporal correlation between renewable sources
  - Daily profiles with hourly resolution

- Renewable sources also increased the importance of interconnections.
Supply Reliability (Adequacy)

- **Importance of interconnections**
  - **Inside a country:** connect renewable sources that are far from the network grid
  - **Among countries:**

  10% Interconnection target by 2020
  
  for every Member State of Europe Union

  15% Interconnection target by 2030

- **Multi Area reliability assessment**
  - **Network flow model**

  ![Network Flow Diagram]

  \[ \text{Maximum flow in a network} = \text{minimum cut of the graph} \]

  - **Monte Carlo Simulation**

  \[ N \sim \frac{1}{\alpha^2 \text{LOLP}} \]

  - **Challenges:**

    - N inversely proportional to LOLP, which is very small
    - Very large number of samples to estimate LOLP
    - Most samples do not result in problems \( \Rightarrow \) not useful data
    - Challenge: How to get more information from each sample?
Proposed solution: Importance Sampling

- Change the probability distribution of available generation capacity to increase the likelihood of sampling the (rare) states that lead to supply problems
  - Less “wasted” samples $\Rightarrow$ smaller sample size $N$ (for the same accuracy)
  - The estimator expected value is determined by:

$$\overline{LOLP} = \frac{1}{S} \sum_{s=1}^{S} \frac{p(s)}{p^*(s)} \Phi(s)$$

$$p(s) = \prod_{i=1}^{n} p_i(x_i)$$

$$p^*(s) = \prod_{i=1}^{n} p^*_i(x_i)$$

$x_i$: available capacity of component $i$

$p(s)$: original joint distribution

$p^*(s)$: distorted joint distribution

$\Phi(s)$: indicator function

The “Goldilocks” challenge: if you change the distribution too little, very large $N$; if you change too much, very large $N$ again. How to change exactly the right amount?
How to find the optimal IS distribution?

**IS procedure**
1. Divide the probability distribution of generation capacity into K clusters (‘bins’); each bin contains a truncated subset of the distribution.
2. Change the probability of each bin (Importance Sampling).
3. Sample the bins from the changed probability of step 2.
4. Given a sampled bin of step 3, now sample a generation capacity state from the corresponding truncated distribution of that bin.

**IS optimization:** Suppose I have a sample of \( S^* \) infeasible generation capacity states (i.e. that lead to supply failure). The *optimum* IS distribution is the one that maximizes the likelihood of that sample.

**Chicken and egg problem!**
How to get a sample of infeasible states if they are rare events?
Proposed Approach to create a sample of infeasible states: MCMC

- **Markov Chain Monte Carlo (MCMC)**
  - Use Markov chains to transition from one load shedding event to another load shedding event
  - Similar to a “random walk” inside the “infeasible” region
  - For a two-area system:

  ![Diagram showing infeasibility cuts](image)

  
  \[
  \begin{align*}
  f_{S_0,A_1} + f_{S_0,A_2} &\leq f_{A_1,T} + f_{A_2,T} \\
  f_{S_0,A_2} &\leq f_{A_2,T} - f_{A_1,A_2} \\
  f_{S_0,A_1} &\leq f_{A_1,T} - f_{A_2,A_1}
  \end{align*}
  \]

  - Once loading shedding samples are obtained

  ➔ Importance Sampling is applied to the proxy marginal distribution (bin)

- **Monte Carlo Simulation:**

  \[
  \text{LOLP} = \frac{1}{S} \sum_{s=1}^{S} \frac{p(x_s^k)}{p^*(x_s^k)} \Phi(x_s^k)
  \]

  \(x_s^k\): capacity of the components in bin \(k\) and sample \(s\)

  \(p(x_s^k)\): original joint distribution of bin \(k\)

  \(p^*(x_s^k)\): distorted joint distribution of bin \(k\)

  \(\Phi(x_s^k)\): indicator function
Example derived from the Saudi Arabia system

- The multi-area system:
  - 4 areas and 4 interconnections
  - Total installed capacity: 47,300 MW
  - Demand in each area proportional to the system total demand

- Using k-means:
  - $A_1 \rightarrow 15$ bins  $A_2 \rightarrow 9$ bins  $A_3 \rightarrow 14$ bins  $A_4 \rightarrow 11$ bins
  - Total demand $\rightarrow 20$ bins

- Results:
  - With 5000 samples of MCMC

<table>
<thead>
<tr>
<th></th>
<th>LOLP expected value</th>
<th>Number of MC samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Method</td>
<td>$10^{-4}$</td>
<td>5,000,000</td>
</tr>
<tr>
<td>Proposed Method: MCMC + IS+ MC</td>
<td>$10^{-4}$</td>
<td>15,000</td>
</tr>
</tbody>
</table>

- Same LOLP expected value
- Speedup: 335 times
Part 2: VRE modeling

- Renewables require a special treatment because the stochastic process changes along the day

\[ \text{Stratification} \]

- The optimal number of draws in each stratum is proportional to the variance of its LOLP
  \[ \Rightarrow \text{another chicken and egg problem!} \]

\[ \text{Solar Generation Profile} \]

**Strata**

- Night ①
- Sunrise ②
- Day ③
- Sunset ④

**MCMC and IS for each stratum**

How many samples should be drawn in each stratum?

**Variance of LOLP estimator is unknown**

**Solution: Estimate**

- LOLP Upper Bound
- LOLP Lower Bound

LOLP ≠ 0 if at least one feasibility cut is violated
Proposed solution: use the LOLP upper value as a proxy

**Upper Bound Estimator:**
- Hunter Inequality

\[
LOLP_{UB} = \sum_{S_1} P(A_i) - \max_{(i,j) \in \tau^*} \sum_{S_2} P(A_i \cap A_j)
\]

- **\(S_1\):** probability of each feasibility cut \(A_i\) be violated individually
- **\(S_2\):** probability of two feasibility cuts \((i,j)\) be violated at the same time
- **\(\tau^*\):** maximum spanning tree

**Lower Bound Estimator:**

\[
LOLP_{LB} = 1 + \left[ \frac{2S_2}{S_1} \right]
\]

With \(LOLP_{UB}\) and \(LOLP_{LB}\) for each stratum:
- If \(LOLP_{UB}\) and \(LOLP_{LB}\) of a stratum are close \(\rightarrow\) LOLP value is “already well estimated” and don’t need to apply MCMC and IS
- If \(LOLP_{UB}\) of one stratum is much smaller than the \(LOLP_{UB}\) of all strata \(\rightarrow\) value is “already well estimated” and don’t need to apply MCMC and IS

- Define more draws in stratum with higher \(LOLP_{UB}\)

**Reduced computational effort**
Example derived from the Chilean system

3 interconnected areas:

<table>
<thead>
<tr>
<th>Area</th>
<th>Thermal</th>
<th>Hydro</th>
<th>Renewable</th>
<th>Number of generators</th>
<th>Nominal Installed Capacity (MW)</th>
<th>Thermo</th>
<th>Hydro</th>
<th>Renewable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41</td>
<td>23</td>
<td>13</td>
<td></td>
<td>3,643</td>
<td>1,149</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>-</td>
<td>17</td>
<td></td>
<td>4,708</td>
<td>-</td>
<td>832</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>4</td>
<td>40</td>
<td></td>
<td>1,176</td>
<td>33</td>
<td>1,653</td>
<td></td>
</tr>
</tbody>
</table>

44 daily generation profiles for each renewable plant
- For instance, for January there were 1374 generation values (44 profiles x 31 days)
- Sum for each hour the generation of all sources.
- Get the average value and CVaR for the 44 scenarios and through Convex Combination get daily generation profile.

Define Strata

4 strata in each one of the 3 areas

Available Power Distribution is obtained (through numerical convolution)

Define bins for the demand inside each stratum

<table>
<thead>
<tr>
<th>Strata</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8 A.M – 9 A.M</td>
</tr>
<tr>
<td>2</td>
<td>9 A.M – 7 P.M</td>
</tr>
<tr>
<td>3</td>
<td>8 P.M – 9 P.M</td>
</tr>
<tr>
<td>4</td>
<td>9 P.M – 7 A.M</td>
</tr>
</tbody>
</table>

Keeps temporal and/or spatial correlation between demand and renewable generation

<table>
<thead>
<tr>
<th>Strata</th>
<th>#Bins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
</tr>
</tbody>
</table>
Example derived from the Chilean system

- Defining the LOLP\(_{UB}\) and LOLP\(_{LB}\) for each strata:
  - 3 areas → 7 feasibility cuts (2\(^3\) − 1)

<table>
<thead>
<tr>
<th>Strata</th>
<th>LOLP(_{LB})</th>
<th>LOLP(_{UB})</th>
<th>Duration (hours)</th>
<th>LOLP(<em>{UB}) / LOLP(</em>{LB})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.41E-04</td>
<td>2.42E-04</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>2</td>
<td>6.00E-04</td>
<td>6.01E-04</td>
<td>11</td>
<td>0%</td>
</tr>
<tr>
<td>3</td>
<td>5.40E-04</td>
<td>1.50E-03</td>
<td>1</td>
<td>178%</td>
</tr>
<tr>
<td>4</td>
<td>1.71E-05</td>
<td>1.72E-05</td>
<td>11</td>
<td>1%</td>
</tr>
</tbody>
</table>

- MCMC, IS and MC are applied for stratum 3
- Total Available Power of each area is divided into bins

<table>
<thead>
<tr>
<th>Area</th>
<th>#Bins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

- 5,000 MCMC samples are drawn from the joint distribution of the areas in stratum 3.

- These samples are used to “skew” the probability distribution of the bins in each area. (IS optimization)
Example derived from the Chilean system

For stratum 3

For each area

Sample a bin

Sample:
- A total available power from the “skewed” capacity distribution
- A demand value from the “skewed” capacity distribution

Adequacy Analysis

Accumulate indexes

\[
LOLP = \frac{1}{S} \sum_{k=1}^{S} p(x^k) \Phi(x^k)
\]

- 50,000 MC samples were drawn from stratum 3
- Relative uncertainty of 1.8%

Results

<table>
<thead>
<tr>
<th>Strata</th>
<th>Duration</th>
<th>LOLP</th>
<th>LOLP x Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.042</td>
<td>2.42E-04</td>
<td>1.01E-05</td>
</tr>
<tr>
<td>2</td>
<td>0.458</td>
<td>6.01E-04</td>
<td>2.75E-04</td>
</tr>
<tr>
<td>3</td>
<td>0.042</td>
<td>1.52E-03</td>
<td>6.33E-05</td>
</tr>
<tr>
<td>4</td>
<td>0.458</td>
<td>1.72E-05</td>
<td>7.87E-06</td>
</tr>
</tbody>
</table>

LOLP = 3.56 E-04

Determined with the probability of each stratum

Comparison with the Standard MC Method

- For the same Relative uncertainty = 1.8%

<table>
<thead>
<tr>
<th>Method</th>
<th>LOLP expected value</th>
<th>Number of MC samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Method</td>
<td>3.56 E-04</td>
<td>8,700,000</td>
</tr>
<tr>
<td>Proposed Method: MCMC + IS + MC</td>
<td>3.56 E-04</td>
<td>10,237</td>
</tr>
</tbody>
</table>

Speedup: 850 times
Conclusions

- **Importance Sampling** is the best Variance Reduction Technique (VRT) for reducing the computational effort of the Standard Monte Carlo Simulation → However requires load shedding samples.

- **Markov Chain Monte Carlo** provides loading shedding samples.

- **Stratification** of the daily profiles of VRE lets considering the variability of renewables within each stratum.

- Defining bins for the demand profile inside each stratum lets keeping **spatial and temporal correlation between renewables and the demand**

- It is possible to avoid applying MCMC and IS in some strata (according to their LOLP upper and lower bounds) which increase even more the method’s efficiency.
Thank you!

Daniela Bayma de Almeida
daniela@psr-inc.com