Modeling and Decomposing Multi-Stage and Multi-Scale Stochastic Optimization Problems

Ramteen Sioshansi
Integrated Systems Engineering
The Ohio State University
Institute for Pure and Applied Mathematics
University of California, Los Angeles
12 January, 2016

Work supported by the National Science Foundation under Grant No. CBET-1029337
Overview

1. Introduction
2. Illustrative Formulation
3. Representative Operating-Stage Periods
4. Decomposition Method
5. Conclusions and Future Work
Background

How to satisfy electricity demands with minimum costs?

- **Scope**: consider long investment periods, multiple electricity-generating technologies, and uncertainties
- **Policy**: renewable portfolio standards, carbon limits, *etc.* may necessitate the use of variable renewable technologies
- **Perspective**: a centralized model where generation and transmission investments are planned together
Supply Uncertainty

Each Day is a different color.

Day 29

Day 5

Day 26

Average

Megawatts

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Supply Variability

![Graph showing normalized wind power variation from 7-Dec-06 to 17-Dec-06. The graph compares:
- Single Turbine
- Group of Wind Plants
- All German Wind Power

The graph visually represents the fluctuation of wind power over the given period.]
Operational Flexibility
Challenges

- Multistage forward-looking investments with recourse
- Multiscale short- and long-run uncertainties
  - **Long-run**: investment costs, technology development, policy changes, fuel prices, demand growth
  - **Short-run**: demand pattern, generation availability, wind speeds, solar insolation
- Problem size explodes if we model multiple stages and all of the uncertainties explicitly
Model Structure

Stochastic, multistage, multiscale model

- Investment decisions: coarse timescale
- Operating decisions: fine timescale

Large scale uncertainties:
- Investment cost
- Electricity Demand
- Fuel cost
Outline

1. Introduction
2. Illustrative Formulation
3. Representative Operating-Stage Periods
4. Decomposition Method
5. Conclusions and Future Work
Objective Function

\[
\min \sum_{\omega} \alpha_{\omega} \left( \sum_{\tau} \sum_{t} \sum_{n} c_{t,n,\omega}^{\tau} X_{t,n,\omega}^{\tau} \right) \quad \text{// gen. invest. cost}
\]

\[
\quad + \sum_{t} \sum_{l} c_{l,t,\omega}^{L} Y_{l,t,\omega}^{L} \quad \text{// trans. invest. cost}
\]

\[
\quad + \sum_{\tau} \sum_{t} \sum_{n} \sum_{r} N_{r} \cdot \sum_{d} \bar{c}_{t,n,\omega}^{\tau} P_{t,n,r,d}^{\tau} \quad \text{// operating cost}
\]
Illustrative Formulation

Investment-Stage Constraints

\[ 0 \leq X_{t,n,\omega}^{\tau} \leq X_{n,\max}^{\tau}, \forall \omega, \tau, t, n \]  
// investment

\[ 0 \leq Y_{t,l,\omega} \leq Y_{l,\max}, \forall \omega, t, n \]  
// limit

\[ \sum_{\tau} \sum_{n} c_{t,n,\omega}^{\tau} X_{t,n,\omega}^{\tau} + \sum_{l} c_{l,t,\omega}^{L} Y_{t,l,\omega} \leq c_{t,\max}^{\max}, \forall \omega, t \]  
// investment budget

\[ X_{t,n,\omega_{k}}^{\tau} = X_{t,n,\omega_{\bar{k}}}^{\tau} : \Omega_{m}(\omega_{k}) = \Omega_{m}(\omega_{\bar{k}}), \forall m < t, \tau, \omega, t, n \]  
// investment

\[ Y_{t,l,\omega_{k}} = Y_{t,l,\omega_{\bar{k}}} : \Omega_{m}(\omega_{k}) = \Omega_{m}(\omega_{\bar{k}}), \forall m < t, \omega, t, n \]  
// nonanticipativity
Operating-Stage Constraints

System Constraints

\[
\sum_{\tau} P^{\tau}_{t,n,r,d,\omega} + P^{STD}_{t,n,r,d,\omega} - P^{STC}_{t,n,r,d,\omega} + UD_{t,n,r,d,\omega} \quad \text{// load}
\]

\[
- \sum_{l|O(l)=n} f_{t,l,r,d,\omega} + \sum_{l|D(l)=n} f_{t,l,r,d,\omega} = D_{t,n,r,d,\omega}, \forall \omega, t, n, r, d \quad \text{balance}
\]

\[
f_{t,l,r,d,\omega} = B_l \cdot (\theta_{t,O(l),r,d,\omega} - \theta_{t,D(l),r,d,\omega}), \forall \omega, t, l, r, d \quad \text{// flow def.}
\]

\[
- f_{l,ES}^{max} - \sum_{m=0}^{t} Y_{m,l,\omega} \leq f_{t,l,r,d,\omega} \leq f_{l,ES}^{max} \quad \text{// flow}
\]

\[
+ \sum_{m=0}^{t} Y_{m,l,\omega}, \forall \omega, t, l, r, d \quad \text{limit}
\]

\[
- \pi \leq \theta_{t,n,r,d,\omega} \leq \pi, \forall \omega, t, n, r, d \quad \text{// phase angle}
\]
Operating-Stage Constraints

Generator Constraints

\[
0 \leq P_{t,n,r,d,\omega}^\tau \leq F_{t,n,r,d,\omega}^\tau \cdot \left( X_{ES,n}^\tau + \sum_{m=0}^{t} X_{m,n,\omega}^\tau \right), \quad \forall \omega, \tau, t, n, r, d \quad \text{// gen. limit}
\]

\[
- R_{\tau} \cdot \left( X_{ES,n}^\tau + \sum_{m=0}^{t} X_{m,n,\omega}^\tau \right) \leq P_{t,n,r,d,\omega}^\tau - P_{t,n,r,d-1,\omega}^\tau \quad \text{// ramp}
\]

\[
\leq R_{\tau} \cdot \left( X_{ES,n}^\tau + \sum_{m=0}^{t} X_{m,n,\omega}^\tau \right), \quad \forall \omega, \tau, t, n, r, d
\]

\text{limit}
Operating-Stage Constraints

Storage Constraints

\[
P_{t,n,r,d,\omega}^{ST} = P_{t,n,r,d-1,\omega}^{ST} - P_{t,n,r,d,\omega}^{STD} + \eta P_{t,n,r,d,\omega}^{STC}, \forall \omega, t, n, r, d \quad // \text{SoC balance}
\]

\[
0 \leq P_{t,n,r,d,\omega}^{ST} \leq h \cdot \left( X_{ES,n}^{ST} + \sum_{m=0}^{t} X_{m,n,\omega}^{ST} \right), \forall \omega, t, n, r, d \quad // \text{SoC limit}
\]

\[
0 \leq P_{t,n,r,d,\omega}^{STC} \leq X_{ES,n}^{ST} + \sum_{m=0}^{t} X_{m,n,\omega}^{ST}, \forall \omega, t, n, r, d \quad // \text{charge limit}
\]

\[
0 \leq P_{t,n,r,d,\omega}^{STD} \leq X_{ES,n}^{ST} + \sum_{m=0}^{t} X_{m,n,\omega}^{ST}, \forall \omega, t, n, r, d \quad // \text{discharge limit}
\]
Two Challenges

1. Many operating stages to capture fine-scale uncertainties
2. Many investment-stage scenarios to capture coarse uncertainties
Outline

1. Introduction
2. Illustrative Formulation
3. Representative Operating-Stage Periods
4. Decomposition Method
5. Conclusions and Future Work
Representative Operating-Stage Periods

Problem: Model is intractable if solving dispatch decisions for every hour in the operating stage

Standard Solution: Use representative hours, based on LDC, to represent operating stage
- Loses correlations between load, wind, and solar
- Cannot model intertemporal constraints (e.g., storage, ramping)

Our Solution: Use representative days with intact correlation structures and intertemporal constraints in operating stage
Representative Days

- Each representative day contains one day’s hourly load, solar, and wind data in each region.
- Cluster to generate representative days that respects the correlation among variables, locations, and time.

<table>
<thead>
<tr>
<th></th>
<th>h1</th>
<th>h2</th>
<th>h3</th>
<th>...</th>
<th>h22</th>
<th>h23</th>
<th>h24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wind</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

_Figure:_ There are 72 data points in a representative day for each region.
Clustering Methods

Method 1

- Hierarchical Clustering using Dynamic Time Warping
- Dynamic Time Warping: measures similarity between two time series, which may vary in time

Method 2

- Step 1: Use $k$-means clustering, with Euclidean distance as a metric, to find a starting set of clusters
- Step 2: Apply Method 1 within each cluster to find representative days
Clustering Test

- **Dataset**: One year’s hourly wind, solar, and load data for three cities in Texas
- **Model**: An investment model with one investment stage and 20 years’ operations
- **Method**:
  - Run the model with original dataset and with representative days from Methods 1 and 2
  - Compare investment decisions and total cost
  - Investigate how decisions change as a function of model inputs
$k$-Means Clustering

- $k$-means clustering is fast, but provides no representative days for the clusters.
- Using cluster centroids gives poor results—variable renewables are overbuilt because they are modeled as having ‘average’ performance.
- Percentiles within clusters could overcome this.
- Hierarchical clustering within each $k$-means cluster provides an actual day from the underlying data.
Investment Cost

![Graph showing Total Cost vs Wind Investment Cost Reduction for different methods.
- No Clustering
- Method 2
- Method 1

The graph illustrates the relationship between the total cost and the wind investment cost reduction for different methods. The cost decreases as the wind investment cost reduction increases. Method 1 shows the highest cost reduction compared to the other methods.]
Investment Capacities

- **Coal Capacities**
  - Graph showing the impact of wind investment cost reduction on coal power capacities.
  - Three methods: No Clustering, Method 2, Method 1.

- **Wind Capacities**
  - Graph showing the impact of wind investment cost reduction on wind power capacities.
  - Three methods: No Clustering, Method 2, Method 1.

- **Natural Gas Capacities**
  - Graph showing the impact of wind investment cost reduction on natural gas capacities.
  - Three methods: No Clustering, Method 2, Method 1.

- **Storage Capacities**
  - Graph showing the impact of wind investment cost reduction on storage capacities.
  - Three methods: No Clustering, Method 2, Method 1.
Clustering Results

- The two clustering methods perform similarly well overall.
- Method 2 takes less time to implement: Method 1 takes about 15 minutes in R studio as opposed to 2 minutes for Method 2.
- 30 clusters (representative days) gives a good approximation of original dataset.
Outline

1. Introduction
2. Illustrative Formulation
3. Representative Operating-Stage Periods
4. Decomposition Method
5. Conclusions and Future Work
Problem: Model may need many investment-stage scenarios to capture coarse-grain uncertainties

Some Solutions:

1. Lagrangian relaxation: multiplier updates highly sensitive to problem data
2. Progressive hedging algorithm
3. Linear decision rules (Kuhn): Fallback option—model decisions as being a linear function of problem data:

\[ x = A\xi \]

need to find coefficients, \( A \)

\(^1\)Rockafellar and Wets (1991): “Scenario and policy aggregation in optimization under uncertainty”.
Progressive Hedging

- Suppose we have the following stochastic problem:

\[
\min_{x} \sum_{s \in S} p_s f_s(x_s) \tag{1}
\]

s.t. \( x_s \in C_s \)

// \( x_s \) is admissible

\( x_s \) is implementable

- Relax nonanticipativity to get scenario-s problem:

\[
\min_{x_s} f_s(x_s) \tag{2}
\]

s.t. \( x_s \in C_s \)
Progressive Hedging

- Add penalty for nonanticipativity violations:

$$
\min_{x_s} f_s(x_s) + \left[ W^T x_s + \frac{\rho}{2} \|x_s - \hat{x}\|^2 \right]
$$

s.t. $x_s \in C_s$

- $W$: Lagrange multiplier vector
- $\rho$: positive penalty parameter, introduced to attain convergence stability in an algorithmic sense
- $\hat{x} = \sum_{s \in S} p_s x_s$: average of $x_s$'s
Algorithm

1: for $s \in S$ do
2: Solve Problem (2) for scenario $s$
3: end for
4: $\hat{x} \leftarrow \sum_{s \in S} p_s x_s$
5: $W_t \leftarrow \rho \cdot (x_t - \hat{x}_t)$
6: while $|x - \hat{x}| > \epsilon$ do
7: for $s \in S$ do
8: Solve Problem (3) for scenario $s$
9: end for
10: $\hat{x} \leftarrow \sum_{s \in S} p_s x_s$
11: $W_t \leftarrow \rho \cdot (x_t - \hat{x}_t)$
12: end while
A feasible solution gives an upper bound

The dual of the non-anticipativity constraints in two-stage stochastic MIPs define implicit lower bounds $^2$

$$\sum_s \rho_s [\min f_s(x_s) + W^T x_s]$$

We show a similar bound for multi-stage stochastic problems

Allows us to assess the quality of a progressive hedging solution

LB obtained with the same effort as one PH iteration

---

$^2$Gade et al. (2014): “Obtaining Lower Bounds from the Progressive Hedging Algorithm for Stochastic Mixed-Integer Programs”.
Performance of PHA

Table: Performance of PHA with Different Number of Representative Days in Operating Stage

<table>
<thead>
<tr>
<th>Days</th>
<th>Variables</th>
<th>Constraints</th>
<th>Full Problem</th>
<th>PHA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>CPU Time [s]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Objective [billion]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CPU Time [s]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1337856</td>
<td>2248967</td>
<td>2968</td>
<td>95.096</td>
</tr>
<tr>
<td>9</td>
<td>3992064</td>
<td>6727943</td>
<td>16535</td>
<td>100.988</td>
</tr>
</tbody>
</table>

- Tested on a three-region system
- Four investment periods, each operating stage lasts 10 years
- 128 scenarios in total
The investment decisions from the original model and the decomposed model are similar.

Maximum and average absolute differences < 2% and 0.3%.

**Table: Investment Capacities [MW]**

<table>
<thead>
<tr>
<th>Investment Period</th>
<th>No Decomposition</th>
<th>Decomposed Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bus 1</td>
<td>Bus 2</td>
</tr>
<tr>
<td>1</td>
<td>37717</td>
<td>3495</td>
</tr>
<tr>
<td>2</td>
<td>1985</td>
<td>2427</td>
</tr>
<tr>
<td>3</td>
<td>4509</td>
<td>2621</td>
</tr>
<tr>
<td>4</td>
<td>8402</td>
<td>949</td>
</tr>
</tbody>
</table>
Upper Bounds and Lower Bounds

3 Representative Days

9 Representative Days
Outline

1. Introduction
2. Illustrative Formulation
3. Representative Operating-Stage Periods
4. Decomposition Method
5. Conclusions and Future Work
Conclusions

- Decisions and uncertainties occur at different stages and scales
- Long-term uncertainties are explicitly modeled in the scenario tree
- Short-term uncertainty are implicitly modeled through different operating-stage problems
- The resulting multistage multiscale stochastic model can be effectively solved using PHA
- Representative days allows intertemporal constraints to be captured in long-term investment decisions
Future Work

- Comprehensive numerical case study
- Study effects of variable renewable energy sources on price signals and investment cost remuneration
Questions?