Wind Output Forecasts and Scenario Analysis for Stochastic Multiperiod Optimal Power Flow

Lindsay Anderson
Ray Zimmerman
PSERC Webinar
November 1, 2011
Overview

• Structure of the Multiperiod SOPF tool
• Representation of uncertainty in inputs
  – Selecting scenarios from data
  – Estimating transitions between scenarios
• Preliminary results with scenario inputs
• Conclusions
Need for Advanced OPF/UC Tools

Recent trends in power systems (increased renewables, controllable demand, storage, “smart grid”) are driving the need for advanced OPF and UC tools able to handle*:

– AC power flow models
– intermittent (renewable) sources
  • variability => costs and constraints on ramping, need for dispatch with multi-period look-ahead
  • uncertainty => impact on reserve requirements, stochastic cost
– energy storage
– controllable demand
– efficient allocation & proper valuation

Tools Require Data

• Important how uncertainty is characterized

• Two types of uncertainty:
  – wind, load (continuous, distribution)
  – contingencies (discrete, low probability)

• Three competing goals:
  – guard against worst cases
  – accurately represent stochastic cost
  – avoid exceeding problem size limitations
Co-optimization Structure

- $P_c$
- $P_0$
- $P^1$
- $P^2$
- $P^k$

- Power flow scenario "high probability", intact case
- Power flow scenario "low probability", contingency case
- Root variable set, deviations, limits (e.g. contract, incs/decs, reserves)
- Transition constraint (e.g. ramp limit)
Reserves
Reserves

MW injections

$p^3$
$p^1$
$p^0$
$p^k$
$p^2$

$0 \hspace{1cm} 1 \hspace{1cm} 2 \hspace{1cm} 3 \hspace{1cm} k$

contingencies

$r_{p+}$ upward reserve

$p_c$

$r_{p-}$ downward reserve
Extending for Wind Uncertainty
Extending to Multiple Periods
Ramping
Ramping
Ramping – Load Following Reserve
Ramping – Load Following Reserve

MW injections

central “high-probability” path
load following ramp up capacity
load following ramp down capacity
Storage
Storage
Storage
Storage

Stored Energy MWh

\( s_{\text{max}} \)

\( s^t \)

\( s^t_+ \)

\( s^t_- \)

\( s_{\text{min}} \)

\( t \quad t+1 \quad t+2 \quad t+3 \quad t+4 \quad \text{time} \)
Storage

Stored Energy MWh

$s_{\text{max}}$

$s^t$

$s^t_+$

$s^t_-$

$s^{t+1}$

$s^{t+1}_+$

$s^{t+1}_-$

$s_{\text{min}}$

$t$

$t+1$

$t+2$

$t+3$

$t+4$

time
Storage

Stored Energy MWh

$s_{\text{max}}$

$s_{\text{min}}$

$t$

$t+1$

$t+2$

$t+3$

$t+4$

$time$
Storage

Stored Energy MWh

$t$, $t+1$, $t+2$, $t+3$, $t+4$
Full Problem
DC Network Version
Decomposition of AC Version
Scenario Selection Goals

Find set of system states across all periods in planning horizon, with corresponding probabilities, that ...

1. approximates overall stochastic cost adequately,

2. includes credible, low probability, high impact events to ensure security,

3. minimizes number of scenarios to keep computational cost reasonable.
Representing Wind in the Multiperiod SOPF

Challenge:
How to represent the possible states of a non-trivial number of wind farms on a system?

• Maintain spatial characteristics, not just total output

• Estimate propensity for transitions between possible states
Wind Scenario Reduction

• Initially looking at a “typical summer day” as determined from historical load data
• A single state defines the output levels of $N$ wind sites

$$P_G = [P_{G1}, P_{G2}, \ldots, P_{GN}]$$

• Data provides a state for each hour of each day included in the data set
Data Clustering

• Unsupervised classification of (large) data sets
• Clusters are selected to maximize similarity of objects within the cluster
• Used look for similar behaviors or attributes in data, determine patterns
• Parameter estimation from historical data

*Common theme is reduction of large data sets to elucidate useful information*
Some Applications

• Gene expression research: How to classify data sets to determine which genes have similar function?
• Market Research: Clustering groups into market segments and target groups
• Health Care: Identification of risk groups or behaviors,
• Communications: Pattern recognition
Quantitative Clustering

• Various clustering techniques exist, specific to data characteristics
• Clustering algorithms can be fuzzy or hard
• Most common is k-means clustering,
  – the number of clusters $k$ is determined \textit{a priori}
  – assign points to clusters to minimize the in-cluster distances
• Reduction of all possible states at each hour to $k$ scenarios that best represents actual states
Algorithm for $k$-means

1. Assign each point $x$ to a cluster (randomly)
2. Calculate the centroid of each cluster
3. Calculate the distance between each point and each centroid
4. Reassign points to cluster with minimum distance to centroid
5. If any points are re-assigned to a new cluster, return to step 2
6. Otherwise stop.

Final centroids are used as “scenarios”
Effect of Increasing $k$
Improving Representation with Larger $K$

If $k$ is large enough, then the errors go to zero, but too many scenarios

If $k$ is one, simple computationally, but nothing is represented well

What is the “correct” $k$ for this application?

Other Metrics:

Largest Deviation in “extreme” scenarios
- System scale
- Single site deviation
Transition probabilities

• Can be estimated empirically from data/scenario outputs

• For each scenario in each hour, what is the probability of transitioning to each of the other states in the next hour?
Transition Diagram
Transition Diagram
Transition Diagram
Transition Matrix

• One transition matrix for each hour to the next:
• For example, hour 1 to hour 2:

\[
p_{ij}(1,2) = \begin{pmatrix}
0.750 & 0.250 & 0.000 & 0.000 \\
0.077 & 0.846 & 0.000 & 0.077 \\
0.014 & 0.000 & 0.873 & 0.113 \\
0.055 & 0.036 & 0.145 & 0.764
\end{pmatrix}
\]

• Where \(p_{ij}\) represents the transition probability from state \(i\) to state \(j\)
• Note that the tendency to persist in the current state is visible in the diagonal of the matrix
$k$-step transition probabilities

• If we consider the impact of a 24-hour horizon, we can see that there is significant uncertainty in the 24-step transition:

$$p(1,24) = \begin{bmatrix} 0.051 & 0.105 & 0.464 & 0.380 \\ 0.052 & 0.107 & 0.460 & 0.381 \\ 0.049 & 0.095 & 0.484 & 0.372 \\ 0.050 & 0.098 & 0.477 & 0.375 \end{bmatrix}$$

• The uncertainty is seen in the similarity of the values in each column – the likelihood of transitioning into each state is nearly independent of the current state.
Preliminary Results

• Implementation of four scenario case with transitions
• Comparison of expected system dispatch with and without wind
Expected Dispatch, No Wind
Expected Dispatch, with Wind
Dispatch of Wind Resources

Real Total Wind Output

Power, MW

Hour

E[dispatch]
E[available]
Min[dispatch]
Max[dispatch]
Min[available]
Max[available]
Conclusions

• A multiperiod planning tool with the capability of taking uncertainty into consideration
• Method for representing uncertainty in historical data
• Preliminary results show that the SOPF is hedging the uncertainty of the wind by spilling in cases of high wind
• Upcoming results consider the impact of storage and deferrable load on wind utilization
Next steps

• Inclusion of storage and deferrable load
• Analysis of ramping costs
• Further sensitivity testing on the correct k to use for
  – Stochastic cost
  – Low probability scenarios
  – High “ramp” scenarios
Questions?